

Qu'est-ce qu'une belle photo ?

Essai sur l'esthétique en photographie numérique

version 3 : 25 septembre 2019

Henri Maître

Professeur émérite, LTCI - Télécom Paris, 46 rue Barrault 75013 Paris
henri.maitre@telecom-paris.fr

RÉSUMÉ

Si la notion de qualité d'image est un sujet d'étude des traiteurs d'images depuis plus de quarante ans, les notions liées à l'esthétique des photographies ne sont apparues que depuis une dizaine d'années dans cette communauté. Mais avec l'automatisation des fonctions d'archivage et de recherche dans les très vastes bases de données numériques (big data), les études qui sont consacrées à ce thème se multiplient aujourd'hui, mettant à profit le développement des techniques d'apprentissage (machine learning) ainsi que la prolifération des sites spécialisés dans l'archivage des photos sur internet. Se pose alors fréquemment le problème de choisir parmi de nombreuses photos et le critère esthétique est alors fréquemment invoqué pour justifier un choix. Est-ce que ce critère est pertinent ? Est-il judicieusement mis en algorithme par les techniques automatiques ?

Quel espoir peut-on mettre dans des solutions qui ne font appel que très marginalement au jugement humain ? Quel profit tirent-elles de vingt siècles de littérature sur l'Esthétique, le Beau et l'Art tant en philosophie qu'en sociologie et en psychologie expérimentale ? Comment se relient-elles aux travaux menés parallèlement en physiologie de la perception et en neuro-biologie qui tentent d'expliquer le fonctionnement du cerveau et dont certains se sont dégagés vers la compréhension de notre jugement esthétique, se regroupant sous le nom de « neuro-esthétique » ?

On fait ici une présentation des diverses approches de l'esthétique des images telles qu'elles ont pu être proposées au cours des 20 derniers siècles, en dégageant les critères les plus opérationnels qui peuvent se prêter à une analyse automatique. Après ce bilan dans le champ de la philosophie, de la neuro-biologie et la psychologie de la perception, on aborde les méthodes qui cherchent à se substituer à l'expert humain. On examine en particulier les deux grandes voies qui ont conduit à des algorithmes : l'approche algébrique, et en particulier sa formulation par Birkhoff, s'appuyant sur une conception objectiviste de la beauté, et l'approche par apprentissage (machine learning), d'inspiration nettement plus subjectiviste. On insiste sur la seconde qui focalise aujourd'hui les attentions et on fait un bilan sur ses résultats, alors même que les travaux utilisant l'outil puissant des réseaux de neurones profonds n'ont encore fourni qu'un petit nombre de conclusions. On discute le choix de techniques à base de détection de paramètres et de classifieurs vs. les techniques à base de réseaux neuro-mimétiques profonds. On dresse un tableau critique des divers résultats aujourd'hui disponibles. On questionne la pertinence des bases de données utilisées ainsi que des « expertises » servant à l'apprentissage. On examine les pistes pour le futur et on donne quelques recommandations à ceux qui s'engagent dans cette voie.

ABSTRACT

Image quality has been introduced in image processing community studies at least 40 years ago, on the contrary to notions linked with aesthetical qualities of photographs which only appeared in the computer science literature in the last decade. However, because of the evolution of Information and Communication Technologies and its consequences on business and societal practices (storage on very large and distributed data bases, archival and retrieval functions, automatic learning processes) and because of the flourishing of photographic exchange sites, the domain received a growing attention in the last decade. Digital aesthetic image assessment is now rich of some tens of papers and "Image Beauty" is more and more chosen as a major criterion to select images in retrieval operations. Is this criterion mimicking human judgement relevant ? Is it correctly translated into algorithms ?

How reliable may be approaches which are only moderately supported by human expertise ? Which benefit is hired from 20 centuries of literature on aesthetics, beauty and art, in Philosophy, in Sociology and in experimental Psychology as well ? Which relations are weaved with the ongoing studies in Physiology and in Neuro-Biology that intend to analyze how our brain is working, some of them directed towards explaining our aesthetical judgement, now recognized under the name of "neuro-aesthetics" ?

This text proposes a review of the several ways to come up to image aesthetics, as proposed in the last 20 centuries. It tries to focus on the most operational criteria which may lead to automatic

implementation. This review being done in the field of philosophy, neuro-biology and psychology of perception, methods are presented which have been proposed to oust human expert. We separate the two main trends : from one side the algebraic way, following Birkhoff's formulation, based on an objectivist concept of beauty, and, on the other side the machine learning approach, much more subjectivist in its inspiration. We pay a greater attention to the later, which retains today most of the computer science community efforts. We present the major works on the subject along with their performances and we discuss their premisses, aware that the most efficient techniques, based on deep neural networks are still under development and far from having deployed their full potential. Methods based on handcrafted feature detection and classifiers are discussed versus end-to-end multi-layer neural nets. We present a critical landscape of the available results today. Relevance of the used data bases is questioned, as well as the associated expertise. Possible tracks for the future are examined and some recommendations are given to those who enter the domain.

Table des matières

1	Introduction	9
1.1	Pourquoi regardons-nous une image?	9
1.1.1	Les mécanismes liés à l'intérêt	10
1.1.2	Les mécanismes liés à la surprise	11
1.1.3	Plaisir <i>vs.</i> excitation	12
1.2	Art, Beauté, Esthétique . . . quel rapport?	13
1.3	Notre projet	15
2	L'héritage des philosophes	17
2.1	L'approche objectiviste	19
2.2	L'approche subjectiviste	25
2.3	Subjectivisme et Objectivisme : un débat toujours vivant	29
3	Neuro-esthétique	33
3.1	Les protocoles d'IRMf et la neuro-esthétique	34
3.2	IRMf et processeur de beauté	35
3.2.1	Le rôle du cortex préfrontal	35
3.2.2	Le rôle du cortex insulaire	36
3.2.3	Le rôle des aires visuelles	38
3.2.4	Le rôle de la mémoire et de la cognition	38
3.2.5	Le rôle de l' <i>embodiment</i>	39
3.3	Que nous dit l'imagerie cérébrale fonctionnelle électrique?	40
3.4	Un schéma cognitif global pour le jugement esthétique?	41
3.4.1	Le modèle neuro-géométrique de J. Petitot	42
3.4.2	Le modèle d'émotion esthétique de A. Chatterjee	42
3.4.3	Le modèle de S. Brown <i>et al.</i>	43
3.4.4	Le modèle de H. Leder	44
3.4.5	Le modèle de C. Redies	45
3.4.6	Le modèle d'émotions de S. Koelsch <i>et al.</i>	45

3.4.7	Le modèle d'émotions de Li-Hsiang Hsu d'après A. Damasio	46
3.4.8	D'autres modèles	47
3.5	L'approche neuro-esthétique critiquée	48
4	Les règles de la photo	51
4.1	Le vocabulaire de l'évaluation esthétique	52
4.2	Le format ou rapport d'aspect	54
4.2.1	Que disent les statistiques?	54
4.2.2	Le nombre d'or	54
4.3	La composition	55
4.3.1	La règle du tiers (<i>Rule Of Third</i> , ROT)	55
4.3.2	Le centre de l'image	57
4.3.3	Quelles autres règles de composition?	57
4.4	Mise au point et spectre de densité de puissance	58
4.5	Texture et histogramme	59
4.5.1	Niveaux de gris	59
4.5.2	Textures	60
4.6	La couleur	61
4.6.1	Du concept de couleur	62
4.6.2	Les préférences des couleurs isolées	62
4.6.3	Les préférences concernant les palettes colorées	64
4.7	Que dit la psycho-sociologie?	68
4.7.1	Mesurer la beauté	68
4.7.2	Images de la nature ou scènes artificielles?	69
4.7.3	Le rôle du contexte	71
4.7.4	La perception globale : unité et prototypalité	72
5	Approches algorithmiques	73
5.1	Les premiers essais : Ch. Henry	73
5.2	L'approche mathématique de G. D. Birkhoff	73
5.3	A la suite de G. Birkhoff	75
5.3.1	L'après-guerre : les <i>designers</i>	75
5.3.2	Une approche dynamique : P. Machado et A. Cardoso	75
5.3.3	Les travaux de J. Rigau <i>et al.</i>	75
5.4	Une approche algorithmique d'IA	76
6	Le contexte de l'IA	79
6.1	L'IA, pourquoi faire?	79

6.1.1	Les résultats attendus	80
6.1.2	Les expertises	82
6.2	Les bases de données	83
7	<i>Hand-crafted features</i>	89
7.1	Primitives multimedia : le système Acquine et les travaux de R. Datta <i>et al.</i>	90
7.2	Contours et distance chromatique : l’approche de Y. Ke <i>et al.</i>	91
7.3	Objet et fond	92
7.4	Des primitives de haut niveau : l’approche de S. Dhar <i>et al.</i>	93
7.5	La bibliothèque de primitives esthétiques de H.-H. Su <i>et al.</i>	94
7.6	Le système OSCAR de L. Yao <i>et al.</i>	95
7.7	Un système embarqué : l’approche de K.Y. Lo <i>et al.</i>	96
7.8	Descripteurs génériques de la vision : l’approche de L. Marchesotti <i>et al.</i>	97
7.9	Harmonie chromatique : les travaux de P. Lu <i>et al.</i>	97
7.10	Réseaux sociaux et <i>crowdsourcing</i> : l’approche de R. Schifanella <i>et al.</i>	99
7.11	Prise en compte de commentaires : les travaux de J. San Pedro <i>et al.</i>	99
8	Réseaux de neurones profonds	101
8.1	Les RNP dédiés à l’évaluation esthétique	102
8.1.1	Le système RAPID de X. Lu <i>et al.</i>	102
8.1.2	L’architecture multi-voies DMA-Net de X. Lu <i>et al.</i>	104
8.1.3	S’adapter à la taille de l’image : les travaux de L. Mai <i>et al.</i>	105
8.1.4	La beauté du web : travaux de M. Redi <i>et al.</i>	105
8.1.5	Le réseau siamois de S. Kong <i>et al.</i>	106
8.1.6	Utiliser l’information sémantique : les travaux de Kao <i>et al.</i>	107
8.2	Variantes autour de l’architecture de base des RNP	108
8.2.1	Une approche qui distingue la nature de la photo : les travaux de Y. Kao <i>et al.</i>	109
8.2.2	Comparer les photos entre-elles : la proposition de K. Schwarz <i>et al.</i>	109
8.2.3	A mi chemin entre classification et RNP : la méthode de Z. Wang <i>et al.</i>	110
8.2.4	Porter attention à la construction de l’image : le système A-Lamp-CNN	111
8.2.5	Extraire une image “dramatique” d’un panorama : le système <i>Creatism</i>	112
8.2.6	Exploiter la distribution des évaluations	113
8.2.7	Formuler une appréciation littérale	114
8.2.8	Personnaliser l’évaluation	115
9	Analyse critique de l’apprentissage	117
9.1	Popularité des études sur l’esthétique	117
9.2	Bilan des méthodes d’apprentissage	117

9.2.1	Quelles architectures?	118
9.2.2	Quelles performances?	119
9.2.3	S'interroger sur les hypothèses	121
9.2.4	De quelle image une belle image est-elle l'image?	122
9.3	Comment introduire de la "subjectivité" dans les RNP?	127
9.3.1	Multiplier les avis	127
9.3.2	Les systèmes de recommandation	128
10	En forme de conclusion	131
	Annexe I : Petite revue des esthétiques	133
	Le monde antique	133
	La Renaissance	137
	Le monde moderne	138
	La bascule romantique	143
	Annexe II : L'esthétique en Chine	145
	L'image du lettré chinois	145
	Objective ou subjective?	149
	Que pourra faire l'IA pour l'esthétique chinoise?	150
	Bibliographie	153
	Index	169

Chapitre 1

Introduction

1.1 Pourquoi regardons-nous une image ?

La prolifération des images dans le monde moderne nous place de plus en plus devant l'obligation de sélectionner des photos dans un vaste ensemble. Cette tâche était réservée aux professionnels de l'édition, de la communication, de l'archivage, aux photographes professionnels, aux collectionneurs, pour choisir les illustrations d'un article, la couverture d'un ouvrage, une affiche, pour résumer un événement, un voyage. Elle incombe maintenant à chacun de nous, pour décider de ce qu'il conserve ou qu'il oublie, de ce qu'il envoie, de ce qu'il poste sur internet, de ce qu'il détruit irréversiblement ou de ce qu'il archive pour une improbable revisite. La gestion de ses archives photographiques est un casse-tête pour beaucoup d'entre nous. L'étape de sélection a toujours été réputée importante et délicate. En milieu professionnel, elle est confiée à des personnes distinguées soit pour leur expertise, soit pour leur autorité.

Arrêtons nous un instant pour nous interroger sur les mécanismes qui sont activés lors de ces décisions de tri et de sélection d'images et voyons la place qui peut être faite alors à la beauté de l'image.

Il nous apparaît que l'on peut regrouper les mécanismes de préférence lors de la sélection de photographies selon trois grandes classes :

- l'intérêt du document, c'est-à-dire sa capacité à mobiliser nos capacités d'attention en attachant le document à nos domaines familiers,
- la surprise, c'est-à-dire, au contraire du point précédent, sa capacité à renouveler notre expérience visuelle par une contribution inattendue,
- enfin la beauté, c'est-à-dire l'agrément qu'il nous procure, indépendamment de son contenu, par la disposition de ses composantes.

C'est ce dernier aspect qui nous intéresse dans ce texte, c'est lui que nous examinerons exclusivement. Dans ce cadre, les deux premiers relèvent des *impressions associatives* au sens de G.T. Fechner qui a soulevé cette distinction dès le *XIX^e* siècle [Fechner, 1871]. Il arrive fréquemment qu'un même document associe plusieurs de ces registres, renforçant l'attrait du document, mais rendant moins clair l'apport de chacun des registres d'attention. Les contributions d'intérêt, de surprise et de beauté semblent cependant évoluer, pour notre décision, dans des espaces indépendants ou orthogonaux, au sens de P. Gärdenfors [Gärdenfors, 2000], c'est-à-dire sans intime influence de l'un sur l'autre. Ils sont donc évalués séparément par notre conscience, puis probablement combinés en un score unique qui nous fait finalement préférer telle image à telle autre dans un choix heuristique difficilement exprimable mais qui pourrait suivre par exemple les schémas de décision empirique proposés par A. Tversky et D. Kahneman [Tversky and Kahneman, 1981].

Examinons brièvement, avant de les abandonner, les deux premiers domaines. Dans le domaine du traitement des images ils sont parfois couverts par des concepts aux noms différents de ceux que nous

employons ici : *interestingness*¹, *memorability*², *unusualness*³ ou *popularity*⁴; on lira à ce propos par exemple [Isola et al., 2011, Gygli et al., 2013, Amengual et al., 2015]).

1.1.1 Les mécanismes liés à l'intérêt

Ils ont été particulièrement étudiés depuis les années 1920 par les psychologues expérimentaux qui ont produit des travaux souvent regroupés aujourd'hui sous le nom de "théorie de la pertinence" (*relevance theory*) [Sperber and Wilson, 2004].

L'intérêt ou la pertinence⁵ est une émotion⁶, c'est-à-dire une expérience psycho-physiologique complexe qui peut résulter :

- de stimulus extérieurs (environnementaux se transmettant par nos sens) ou intérieurs (biochimiques, créés par notre cortex), elle est alors passive ;
- ou de raisonnements, c'est-à-dire de processus cognitifs qui permettent de déduire de nouveaux éléments de connaissance à partir d'un contexte particulier ; elle est alors active.

Les travaux de W. James [Lange and James, 1922] ont ouvert les études sur les relations entre émotion et évaluation (*appraisal*) d'une situation par l'homme. L'importance de l'incitation (*arousal*) et l'ordre des diverses étapes (incitation, émotion, évaluation) ont ainsi été bien documentés. Ces travaux ont servi de guide à de nombreuses études conduites un demi-siècle plus tard sur l'esthétique.

Mais c'est surtout D. Sperber et D. Wilson [Sperber and Wilson, 2004] qui ont permis de construire une théorie de la pertinence mobilisant l'ensemble du bagage cognitif du récepteur et non plus seulement ses connaissances linguistiques comme le suggérait W. James. Dans ce cadre, J.L. Dessalles [Dessalles, 2008] a proposé une mesure plus quantitative de la pertinence et l'a dotée de capacités prédictives par une modélisation mathématique originale.

Lorsqu'on cherche à appliquer cette théorie de la pertinence à l'intérêt éveillé par une collection d'images, il est utile de distinguer deux cadres extrêmes qui seront peut-être traités de façon différente.

L'intérêt universel fait référence aux thèmes particulièrement exposés dans la société et les médias et transmis par le contexte et la culture : tel acteur ou tel sportif, telle voiture, monument ou évènement explicitement et régulièrement couvert. Leur popularité pourrait se mesurer aujourd'hui avec les outils de médiamétrie [Hsieh et al., 2014, Fu et al., 2014] : fréquence d'exposition à la télévision, nombre d'instances sur internet, de " clics " ou de "*likes*" sur les réseaux sociaux, etc. (figure 1.1 à gauche).

L'intérêt personnel il rapproche des images en raison des liens étroits avec les thèmes de la vie individuelle et plus particulièrement personnelle : "ma" famille, "mon" travail, "ma" ville ... La mesure de la pertinence prend alors les formes plus classiques proposées par exemple dans [Dessalles, 2013] mettant en jeu des variables d'espace et de temps à décroissance plus ou moins rapide et des relations d'ordre de proximité (par exemple dans un arbre de famille ou d'entreprise) (figure 1.1 à droite).

1. *Interestingness* est souvent assez proche de ce que nous décrivons par intérêt, mais le mot recouvre chez de nombreux auteurs également l'attrait de l'esthétique [Gygli et al., 2013]. Il peut aussi prendre un sens particulier dans le cas de réseaux sociaux où il mesure toute forme d'intérêt indépendamment des causes mises en jeu [Dhar et al., 2011].

2. *Memorability* distingue les images sur un critère différent de leur impact immédiat sur l'observateur. Il reflète la capacité d'une image à imprégner notre mémoire à long terme qui stocke les souvenirs installés dans notre esprit. C'est une qualité qui, même si elle est partagée, est mesurée au niveau individuel [Khosla et al., 2012, Kim et al., 2013].

3. *Unusualness* recouvre une propriété des images que nous caractériserons par leur effet de surprise : à ce titre, les images concernées se distinguent des autres images proposées, en particulier dans le cas de séquences animées. Cette propriété est spécifique d'un contexte donné et peu disparaître dans un autre contexte [Schuster et al., 2010].

4. *Popularity* est le terme réservé pour caractériser des images postées sur les réseaux sociaux. La popularité reflète un intérêt partagé par une grande quantité de personnes sur des critères qui peuvent être très variables [Amengual et al., 2015].

5. Le texte anglo-saxon utilisé est le mot *relevance* de préférence à *interestingness* plus général comme nous l'avons vu ci-dessus.

6. Ce point est cependant encore discuté dans la littérature.



Figure 1.1 – La pertinence d’une image peut avoir pour référence le cadre général des connaissances d’une communauté (à gauche : des acteurs, des sportifs, des objets fréquemment présentés par exemple dans les journaux ou sur les réseaux sociaux), ou au contraire un cadre personnel lié à l’observateur (à droite : les membres d’une même famille, les lieux de vacances, les passe-temps, etc.).

1.1.2 Les mécanismes liés à la surprise

Ces mécanismes fonctionnent un peu à l’inverse des précédents car ils sont d’autant plus importants que le document qui nous est présenté est plus éloigné de ce que nous connaissons (figure 1.2). La surprise⁷ peut se décliner selon des formes variées : l’humour, l’effroi, la perplexité, la tendresse, le dégoût ...



Figure 1.2 – La surprise est un autre mécanisme qui nous amène à porter une attention particulière à une image. On s’attache alors à bien identifier les éléments particuliers de l’image qui sortent des schémas connus de notre représentation de l’univers.

On voit bien que la pertinence d’une part, la surprise d’autre part, sont des motivations différentes des qualités esthétiques pour retenir une photographie. Qu’elles opèrent conjointement pour expliquer un choix est aussi très probable⁸, comme nous le constatons sur les deux photos de la figure 1.3. Nous

7. Le terme anglais consacré pour la surprise est *amazement*. Il couvre un champ plus large que le mot français et c’est bien toute l’étendue qu’il couvre que l’on souhaite représenter ici.

8. Il est en particulier très intéressant de voir comment les rétrospectives régulièrement faites par les journaux professionnels sous des titres comme : “*Les 100 plus belles photos de l’année*” mêlent très largement aux photos esthétiquement remarquables les évocations de phénomènes exceptionnels du sport, de la vie politique ou de la marche

nous efforcerons cependant par la suite de les séparer, c'est-à-dire que nous essaierons d'évaluer la qualité esthétique à pertinence et surprise égales. Ce ne sera pas facile et il faudra retenir que dans de nombreuses circonstances les avis donnés lors de tests subjectifs n'auront pu échapper à une certaine confusion sur ce point [Gygli et al., 2013]. Cela sera particulièrement vrai lorsque l'on sollicitera des avis de personnes inconnues et distantes comme on le fait par exemple par des mesures sur le web.

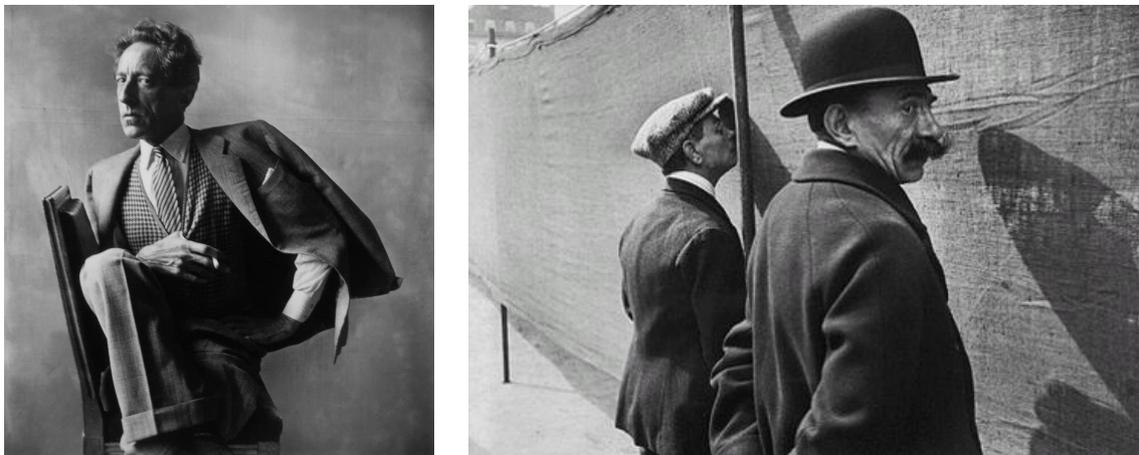


Figure 1.3 – Ces deux photographies sont indéniablement remarquables pour leurs qualités esthétiques. Celle de gauche (Portrait de Jean Cocteau, 1948 par Irving Penn) retient également notre attention parce que nous y reconnaissons un homme célèbre (stimulus d'intérêt de type universel). Celle de droite (Bruxelles, 1932 de Henri Cartier-Bresson) nous interpelle par l'interrogation qu'elle suscite : "que regardent-ils ?" (stimulus d'amazement).

1.1.3 Plaisir vs. excitation

Les trois moteurs de notre attention lors de l'examen d'une photo : sa *beauté*, notre *intérêt* pour son sujet et la *surprise* qu'elle éveille en nous sont le résultat d'une observation longue, attentionnée et raisonnée. Mais on peut s'intéresser aux effets immédiats sur notre organisme de la vision très brève d'une photographie. On recherchera alors à identifier les effets biologiques les plus élémentaires qui ne mettent pas en œuvre les fonctions élaborées de la cognition et du raisonnement mais les activations physiologiques de base. Pour cela on présente une photo à un observateur pendant une fraction de seconde et l'on mesure quelques indicateurs physiologiques qui reflètent deux émotions réflexes indépendantes : le *plaisir* d'une part et l'*excitation*⁹ d'autre part. Ces expériences ont été initiées par les psychologues du *National Institute of Health*, puis repris par de nombreux autres auteurs. Pour cela, ils ont constitué une bibliothèque d'un millier de photos de référence (la bibliothèque IAPS (*International affective picture system*¹⁰) [Lang et al., 1999]. Ils ont ensuite conduit des expériences de psycho-physiologie qui ont permis de placer chaque image dans un graphe *plaisir* en fonction de l'*excitation*, dégageant deux orientations très nettes qu'ils ont baptisées la direction de *motivation appétitive* et la direction de *motivation défensive* (figure 1.4).

Une photo qui est sélectionnée pour ses capacités de surprise peut éveiller tout autant du plaisir ou de l'excitation. Une photo sélectionnée pour l'intérêt qu'elle éveille en nous l'est surtout pour le plaisir, et plus encore une photo que l'on distingue par sa beauté et nous verrons que ce plaisir est même aujourd'hui un déterminant de la Beauté. Cependant la littérature des *XVIII^e* et *XIX^e* siècles plaçait, parfois, au dessus du Beau, le Sublime¹¹ qui comportait viscéralement une composante d'effroi

du monde, retenues essentiellement pour leurs qualités mémorielles.

9. Excitation : le terme anglais déjà rencontré qui est consacré est *arousal*, qui exprime la sortie de l'état d'indifférence.

10. Une bibliothèque concurrente de la base de données IAPS, la bibliothèque OASIS, a été créée avec le même objectif par l'Université d'Harvard [Kurdi et al., 2017]. Elle offre l'avantage d'être disponible en accès libre.

11. Le Sublime est l'objet du premier texte philosophique d'Edmond Burke (en 1757) : *A philosophical enquiry into*

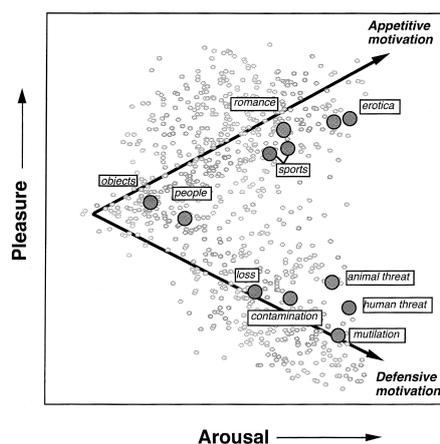


Figure 1.4 – L’espace affectif de Lang et al. Les photos de la bibliothèque IAPS [Lang et al., 1999] sont distribuées selon deux axes, l’un porte l’excitation (arousal), qui ne présente que des valences positives, l’autre porte le plaisir qui dispose de valences positive et négative. Quelques mentions sont portées isolant des thématiques particulières et permettant de juger de la gradation des affects le long des deux axes (hy-po-thétiques) de motivation appétitive d’une part et défensive d’autre part [Schupp et al., 2004].

(et donc d’excitation). Le sublime n’est plus beaucoup évoqué en ce sens aujourd’hui et n’exprime le plus souvent qu’un superlatif du beau.

1.2 Art, Beauté, Esthétique ... quel rapport ?

Tournons donc maintenant notre attention sur les seules images que nous jugeons belles et tâchons d’écarter les motivations plus complexes qui peuvent mériter que l’on soit attiré par elles. Ce texte se destine à dégager des éléments objectifs nous permettant d’expliquer pourquoi on attribue ce qualificatif à une photo. Mais pour ce faire, il nous faut au préalable partager quelques définitions pour fixer un vocabulaire qui est souvent assez élastique.

« **L’Esthétique** est la science qui a pour objet de rechercher et de déterminer les caractères du beau dans les productions de la nature ou de l’art » (définition de l’Académie française de 1835). En termes plus modernes, c’est « la branche de la philosophie concernée par :

- la nature de l’art, de la beauté et du goût,
- la création et l’appréciation de la beauté. »

« **La Beauté** est la qualité de ce qui est beau [...]. C’est le caractère de ce qui provoque l’admiration et l’émotion, par ses formes, ses proportions, ses rythmes, son harmonie » (définition de l’Académie française de 1986).

On préfère souvent aujourd’hui des définitions qui mettent plus l’accent d’une part sur l’origine du stimulus, d’autre part sur la forme de ses effets : « La Beauté est la caractéristique d’une personne ou d’un objet qui procure une expérience perceptuelle de plaisir ou de satisfaction ».

the origin of our ideas of the Sublime and Beautiful”. Il est apparu en France au cœur de la polémique des Anciens et des Modernes après la traduction, en 1674, par Nicolas Boileau de l’ouvrage retrouvé du grec Pseudo-Longin : Περὶ υψους (*Du sublime*) et il apportait le lyrisme de l’enthousiasme en contrepoids de la juste mesure et de l’harmonie. Il a été très longuement discuté par Emmanuel Kant en 1790 dans *La critique de la faculté de juger* ([Kant, 1985], §29) qui écrit en particulier “Le beau nous prépare à aimer quelque chose, même la nature, d’une façon désintéressée ; le sublime à l’estimer hautement, même contre notre intérêt sensible.”. Et il ajoutait : “La satisfaction prise au sublime de la nature est en outre, (...) uniquement négative (alors que celle qui s’attache au beau est positive)”. Le Sublime est chanté par Byron lors de ses voyages dans les vallées étroites des Alpes par temps d’orage. On trouvera dans [Saint Girons, 2005] une étude historique du sublime.

« **L’Art** est la gamme des activités de l’homme conduisant à la fabrication d’artéfacts dans le but qu’ils soient appréciés pour leur beauté ou pour leur pouvoir émotionnel ».

Rapprochant ces trois termes : Esthétique, Beauté et Art, nous voyons qu’ils couvrent bien notre champ d’étude, et comme ils ont les uns et les autres conduit à une abondante littérature, il serait tentant de les mobiliser tous les trois. Il faut cependant être plus prudent avec le troisième : “Art”. Si, pendant de nombreux siècles, l’objectif de l’Art a été très majoritairement tendu vers la recherche de la Beauté¹² il faut cependant examiner avec circonspection son évolution à partir du *XIX^e* siècle et en particulier considérer la “fracture du *XX^e* siècle” comme un changement définitif de paradigme. Il est assez clair que le second terme de la définition ci-dessus (“leur pouvoir émotionnel”) est devenu l’enjeu premier de l’Art, bien avant “la Beauté”. Rétrospectivement, il est aisé de retrouver des prémices de cette préoccupation bien avant les Surréalistes, puisque de nombreuses œuvres éminemment artistiques (Bosch, Chardin, Goya . . .) ne prétendent pas à être premièrement “belles” (figure 1.5).



Figure 1.5 – Nombre d’œuvres d’art ont été conçues sans que le critère de “Beauté” soit mis en avant. L’artiste recherche alors avant tout la “charge émotionnelle” de son message. Ces œuvres “à message” sont très minoritaires dans la période classique (représentée ici par J. Bosch, F. Goya et J.S. Chardin). Elles deviennent rapidement majoritaires au *XX^e* siècle (représenté ici par M. Duchamp, A. Warhol, E. Munch, F. Bacon) et il serait bien difficile d’évaluer les œuvres du *XX^e* siècle à la seule aune de leurs qualités esthétiques.

Cette rupture dans les objectifs mêmes de l’Art doit nous faire prendre pour guides avec une grande précaution certains textes dans notre quête du Beau. Cela est très vrai pour ce qui concerne la peinture, mais ça l’est aussi pour ce qui concerne la photographie qui aspire à juste titre au rang d’Art¹³. Il nous faudra suivre avec une grande prudence les recommandations des photographes, surtout les plus chevronnés, lorsqu’ils s’engageront dans des conseils où ne seront pas clairement démêlées les intentions à atteindre une plus grande esthétique et celles visant à une plus forte charge émotionnelle. Nous renvoyons le lecteur aux remarques faites plus haut en section 1.1. Nous noterons aussi qu’une définition claire de ce qu’est l’Art dans notre société (voir par exemple [Danto, 1992, Danto and Goehr, 2014]) nous sera particulièrement utile pour répondre à deux questions souvent

12. Il nous faut ici restreindre notre propos à l’art occidental issu de l’école grecque. Les arts des autres sociétés, et en particulier ceux de l’Orient, relèvent probablement de tout autres intentionalités. On trouvera ainsi dans l’Annexe 2 une brève étude sur la conception esthétique de la Chine antique et dans [Jullien, 2003] une étude exhaustive de ce sujet.

13. Que la photographie puisse accéder au rang d’art est aujourd’hui bien admis parmi les amateurs d’art. Une enquête d’octobre 2017 de la revue *Beaux Arts Magazine* indique qu’à la question *A l’aide de quelle forme d’expression les artistes réussissent-ils le plus souvent à incarner la beauté ?*, la photo est placée au premier rang pour 32 % des Français interrogés, tandis que la peinture ne recueille que 31 % des avis et la sculpture 15 %. On est bien loin de la situation constatée par P. Bourdieu [Bourdieu, 1965] qui qualifiait, au milieu du siècle dernier, la photographie en France d’ “art moyen”, non seulement pour la qualité de sa production mais également pour l’estime dont elle bénéficiait.

entendues qui empoisonnent l'art contemporain :

- une copie parfaite de la Joconde est-elle aussi belle que l'original ¹⁴ ?
- est-ce que la boîte de soupe Campbell d'A. Warhol est plus belle que celle de l'épicerie du coin et qu'est-ce qui en fait une œuvre d'art ?

1.3 Notre projet

Il nous faut expliquer tout d'abord pourquoi nous abordons ce sujet aujourd'hui. Les premières lignes de ce texte rappelaient combien l'image est aujourd'hui abondante dans notre société. Elles évoquaient la complexité de la tâche nécessaire de tri des masses de photographie et le rôle de l'esthétique, à côté d'autres émotions, l'intérêt et la surprise, pour assurer ce tri. Il nous faut alors aussi rendre compte des progrès qui ont été faits pour comprendre comment on juge de la qualité esthétique d'une photographie, puis des tentatives qui ont été faites pour nous faire aider dans cette tâche par des outils du traitement de l'information. Ainsi, pour la première fois, en 2018, un concours photographique international est arbitré parallèlement par deux jurys, l'un constitué d'experts photographes, l'autre d'un ordinateur doté d'un logiciel d'intelligence artificielle ¹⁵.

Il nous semble que ces deux aspects : la compréhension des mécanismes mis en jeu par l'homme d'une part, le développement des outils d'assistance d'autre part, vont nécessairement de pair. C'était du moins la conviction profonde qui nous animait au lancement de cette étude. Mais alors même que ce travail avançait (soit de septembre 2016 à 2019), des résultats n'ont cessé de s'accumuler qui semblent contredire cette idée : les réseaux de neurones profonds, quoiqu'encore balbutiants, font fi d'une compréhension des liens entre l'observateur et la photo, mais conduisent pourtant à d'étonnament bons jugements sur les qualités esthétiques de celle-ci (s'ils ne se substituent nullement aujourd'hui à l'œil d'un expert, ils surclassent nettement les autres outils préalablement développés, comme cela sera longuement discuté plus loin : voir chapitre 9.).

Notre sujet a ainsi évolué alors que nous le traitions. Nous expliquons moins que nous le voulions initialement, mais nous témoignons plus précisément des démarches adoptées, nous rapportons les évolutions dans les approches, nous comparons les performances et nous soulignons les avancées et les faiblesses.

Notre travail est scindé en deux grandes parties.

Dans la première, nous essayons de faire un bilan des travaux entrepris dans divers domaines de la connaissance pour identifier les mécanismes de la perception esthétique. C'est tout d'abord l'énorme corpus de la littérature philosophique que nous chercherons à cartographier de façon très schématique au chapitre 2. Nous nous efforcerons de simplifier les très nombreux travaux qui courent depuis 25 siècles de façon à dégager des pistes universelles. Nous examinerons ensuite ce que nous disent les biologistes, et surtout les neurologues, des mécanismes de la perception esthétique. Le développement des outils modernes d'imagerie cérébrale a bouleversé notre connaissance de la cartographie des aires impliquées dans l'élaboration de nos émotions et de nos jugements. Les travaux qui relèvent du domaine de l'art, actuellement regroupés sous le terme de neuro-esthétique, seront présentés au chapitre 3. La photographie, au cours des 150 années de son histoire, a également développé un important corpus de recommandations et d'enseignements, souvent issus de la pratique de la prise de vue, mais fréquemment aussi étayés par des expérimentations et des démarches scientifiques s'appuyant sur les connaissances issues de la psychologie expérimentale. Ces résultats que les traiteurs d'image ont souvent suivis seront présentés très en détail au chapitre 4. On verra comment ils sont utilisés aujourd'hui par les algorithmes de la seconde partie.

14. Cette question fut longuement débattue par G.T. Fechner à propos de la Madone de Dresde (voir [Vidal, 2011]). A contrario de la position exprimée par exemple dans [Danto, 1992] qui exprime un consensus souvent partagé aujourd'hui, Fechner avait soutenu l'étroite dépendance entre *Schönheit* (beauté) et *Echtheit* (authenticité), refusant toute beauté à la copie, aussi parfaite fût-elle.

15. Il s'agit de la compétition "Spark" (*SPARK, a Renaissance photographic contest*) de Huawei, construite autour de son modèle P20 Pro de téléphone portable, <https://consumer.huawei.com/uk/campaign/sparkrenaissance/> (juillet 2018).

Dans cette seconde partie, nous examinerons les travaux mettant à contribution les outils issus des mathématiques pour mesurer la beauté. Tout d'abord, au chapitre 5 nous présenterons les techniques, parfois assez anciennes, qui proposent une formule complète (*closed form*) de la beauté (comme celle proposée par G. Birkhoff). Ces travaux, quoique très rares, se poursuivent encore de nos jours. Les trois chapitres suivants sont consacrés aux méthodes actuelles issues de l'intelligence artificielle. Nous ferons tout d'abord un examen global du contexte qui rend ces démarches possibles, en particulier des bases de données d'images disponibles sur le réseau, puis des expertises que l'on peut mobiliser pour assurer l'indispensable étape d'apprentissage automatique (chapitre 6). Le chapitre 7 détaillera la première famille des travaux conduits à partir d'une détection de primitives sensées contribuer à la qualité esthétique et leur classification par des techniques de *machine learning*. Ces travaux ont pris naissance vers l'an 2000. Le dernier chapitre couvrira les travaux les plus actuels, conduits essentiellement depuis 2014, effectués à l'aide de réseaux de neurones profonds et se passant donc de toute expertise humaine, hormis dans la constitution des sources d'images annotées (chapitre 8).

Bien sûr nous apporterons une conclusion à ce travail, en remettant en perspective les contributions des divers chapitres et en montrant comment les techniques d'intelligence artificielle s'appuient sur des prémisses différentes des autres travaux conduits sur l'esthétique. Nous proposerons quelques pistes pour réconcilier les diverses communautés. Nous avons le sentiment, cependant, que nous ne sommes qu'à l'aube d'une époque nouvelle et qu'il est plus que probable que l'avenir se fera un malin plaisir de déjouer nos prévisions.

Enfin deux annexes complètent le chapitre 2. La première retrace le parcours de l'esthétique issue des Grecs telle qu'elle s'est déployée jusqu'au siècle des Lumières, conduisant à l'apogée de ce que l'on a appelé l'Ecole Classique. La seconde aborde l'esthétique chinoise dont les critères échappent totalement aux règles mises en place par l'esthétique occidentale et qui, pour cela ne semble pas concernée par les approches algorithmiques que nous traitons ici.

Chapitre 2

L'héritage des philosophes

“Tout le monde raisonne du beau : on l’admire dans les ouvrages de la nature : on l’exige dans les productions des Arts : on accorde ou l’on refuse cette qualité à tout moment ; cependant si l’on demande aux hommes du goût le plus sûr & le plus exquis, quelle est son origine, sa nature, sa notion précise, sa véritable idée, son exacte définition ; si c’est quelque chose d’absolu ou de relatif ; s’il y a un beau essentiel, éternel, immuable, règle & modèle du beau subalterne ; ou s’il en est de la beauté comme des modes : on voit aussitôt les sentimens partagés ; & les uns avoient leur ignorance, les autres se jettent dans le scepticisme. Comment se fait-il que presque tous les hommes soient d’accord qu’il y a un beau ; qu’il y en ait tant entr’eux qui le sentent vivement où il est, & que si peu sachent ce que c’est ?”

Denis Diderot : Encyclopédie, entrée : Beau

L’importance accordée par les philosophes à la notion de Beau depuis l’Antiquité se lit à travers l’abondance des œuvres majeures que cette notion a engendrées. Les “travaux théoriques” de Platon, de Socrate et d’Aristote, appuyés sur les “travaux pratiques” d’Homère, de Phidias ou de Praxitèle fournissent le cadre général de réflexion que la Grèce a transmis au monde romain, puis au Moyen-Âge et enfin à la Renaissance où il s’est vu accorder une place prééminente dans une vie artistique et intellectuelle particulièrement exubérante. Au Siècle des Lumières, c’est le monde des idées qui s’en empare à l’occasion de disputes très vigoureuses (Boileau, Locke, Diderot ...) mais sans que le domaine des Arts le remette vraiment en cause [Tatarkiewicz, 1970]. A la fin du XVIII^e et au XIX^e siècle, il n’est pas de grand philosophe qui ne se mêle au débat : Kant, Hegel, Schopenhauer, Nietzsche ... Tous produisent sur ce sujet les traités essentiels de nos bibliothèques, cependant que les artistes entament de leur côté une révolution copernicienne dans la conception de leur art, emblématiquement portée par les Romantiques en littérature et par les Impressionistes en peinture. Dans le domaine de l’Art, le XX^e siècle s’ouvre sur un indescriptible champ de bataille où s’affrontent les théoriciens arc boutés sur des positions passionnées : Heidegger, Adorno, Wittgenstein,

Deleuze, Derrida, Foucault, Eco, Danto ...¹, tandis que le monde des œuvres d'art est livré à une rare anarchie².

Il n'est pas question bien sûr de rapporter ici fidèlement toutes ces contributions. Nous en ferons un raccourci extrêmement simplifié (et donc extrêmement critiquable), qui aura pour but de situer l'enjeu de notre étude et d'en placer certains termes en perspective. Pour le besoin de notre cause, nous privilégierons ainsi deux points de vue de la conception esthétique, aux antipodes l'un de l'autre : l'approche "objectiviste" qui attribue le mérite de la beauté au seul objet qui la porte, et l'approche "subjectiviste" qui au contraire situe la beauté dans le ressenti de l'observateur³. Nous montrerons comment la première a laissé la place à la seconde au cours des derniers siècles, puis comment les deux approches se sont tant bien que mal hybridées. Nous montrerons surtout quelles implications ces deux théories induisent dans notre projet de déterminer automatiquement la beauté d'une photographie et pourquoi elles sont inmanquablement présentes dans les attendus des logiciels développés aujourd'hui. Nous adoptons le parti-pris de mettre un accent plus marqué sur l'approche objectiviste car il apparaît que toutes les démarches qui visent à remplacer le jugement humain par un jugement par ordinateur adoptent aujourd'hui un point de vue objectiviste.

Introduisons tout d'abord l'enjeu de cette classification en deux familles opposées, et pour cela nous citons encore D. Diderot qui en fixe clairement les termes dans l'Encyclopédie : *"Il est évident que S. Augustin avoit été beaucoup plus loin dans la recherche du beau que le philosophe Leibnitien⁴ : celui-ci semble prétendre d'abord qu'une chose est belle, parce qu'elle nous plaît ; au lieu qu'elle ne nous plaît que parce qu'elle est belle ; comme Platon & S. Augustin l'ont très-bien remarqué."* Est-ce le Beau qui donne du plaisir, ou le plaisir qui crée le Beau ? Diderot, se référant aux Anciens, a choisi, mais il existe bien deux façons de prendre le problème comme on le verra ci-dessous ; par exemple, on lit dans Luc Ferry la contraposée exacte de la phrase de Diderot⁵.

Munis de ces deux référentiels philosophiques, nous nous tournerons vers les résultats récents de la neurobiologie pour voir s'il est possible de les départager ou de les valider par des expérimentations en imagerie cérébrale. Nous examinerons également leur compatibilité avec des résultats d'études en psychologie expérimentale, en psychosociologie et enfin en sociologie menées depuis un siècle. Nous disposerons alors d'un faisceau de travaux qui nous permettront d'interpréter les démarches adoptées dans les approches automatiques de détermination de la beauté des photographies.

1. On consultera avec intérêt : [Gombrich, 1960, Danto, 1964, Zemach, 1987, Kemp, 1990, Solso, 1996, Ferry, 1990]

2. On ne peut manquer d'en appeler ici à "la fin de l'histoire de l'art" affichée par A. Danto en 1964 [Danto and Goehr, 2014], mais déjà annoncée auparavant par Hegel, Spencer, Heidegger et bien d'autres ...

3. D'autres classifications seraient possibles et certaines permettraient aussi d'éclairer notre objet : "monisme/dualisme", "empirisme/rationalisme", "physicalisme/spiritualisme" ... Ces grandes conceptions de la nature de l'esprit et du corps sont souvent des clés d'entrée dans la philosophie de l'esthétique dans le débat contemporain.

4. Diderot parle ici du philosophe et mathématicien Christian Wolff (1679-1784).

5. *"Ce n'est plus parce que l'objet est intrinsèquement beau qu'il plaît, mais, à la limite, parce qu'il procure un certain type de plaisir qu'on le nomme beau."*, ([Ferry, 1990] p. 25).

2.1 L'approche objectiviste

A la source : la Grèce

L'approche objectiviste⁶ est celle qui fut formalisée dès l'Antiquité grecque, celle de Platon⁷ et Socrate⁸, celle qu'Aristote⁹ a également adoptée, en l'humanisant un peu. C'est celle qui fut transmise au monde occidental par les Romains, qu'a reprise Augustin d'Hippone¹⁰ pour l'adapter au christianisme, qui alors s'est épanouie à la Renaissance, prospéra dans l'époque Classique et tint le haut du pavé jusqu'au milieu du *XIX*^e siècle dans le milieu artistique, si ce n'est dans le milieu philosophique.

L'approche objectiviste affirme que la Beauté est une qualité de la personne, de l'objet, de la musique concernés. Cette Beauté est le reflet, dans l'objet observé, de propriétés universelles d'harmonie dans les proportions et dans les formes, propriétés qui ne nous sont pas exactement connues mais qui sont préexistantes à notre observation et s'imposent à notre jugement¹¹. La Beauté est à ce titre perçue également par tout observateur¹², en tous temps et en tous lieux. Ainsi, parmi d'autres, la Vénus de Milo, l'Apollon du Belvédère, le David de Michel-Ange ont été parés par les experts et historiens d'art les plus réputés de ces qualités éternelles qui les imposent à l'humanité. Les règles établies à Athènes au 5^e siècle avant notre ère et transcrites dans le domaine de la perception résident dans l'harmonie des formes¹³ et des proportions et en particulier dans la consonance du tout et des parties¹⁴. Les proportions les plus simples sont les plus agréables (1/2, 1/3, 3/4 ...), mais l'œil apprécie des fractions plus subtiles dans les constructions plus complexes comme celles de l'architecture. L'écho des proportions à diverses échelles et en diverses parties de l'objet d'art est important et contribue à la "symétrie"¹⁵. Tout l'art grec est soutenu par une trame mathématique très

6. L'objectivisme est une pensée philosophique reposant sur 3 prémisses :

- l'existence existe, (c'est-à-dire qu'il y a un monde en dehors de nous),
- la conscience existe (c'est-à-dire que les résultats de notre activité cérébrale sont réels),
- la conscience peut accéder à l'existant (par le truchement de la perception).

L'objectivisme partage ces prémisses avec le réalisme, mais l'objectivisme s'autorise le recours à la transcendance. Le terme a été revendiqué depuis 1950 par Ayn Rand pour recouvrir une forme radicale d'objectivisme appliqué à la politique et à la société qui vise à l'établissement de gouvernements libertariens et prône l'égoïsme social. Cette acception de l'objectivisme est devenue très populaire outre Atlantique depuis la diffusion à très grand succès de *Atlas Shrugged*, en 1957, (en français : *La Grève*). Ce n'est pas dans le sens d'A. Rand que nous parlons ici de l'objectivisme esthétique. Nous écartons deux attendus qu'elle a ajoutés à ceux que nous donnons : "la finalité de l'homme est son bien-être" et "le meilleur système politico-économique est le capitalisme libéral".

7. Platon (428-348 av. JC) aurait écrit deux Traités du Beau, l'un dans *Hippias Majeur*, l'autre dans *Phèdre*, [Platon, 2011].

8. Socrate n'a laissé aucun texte et n'est connu que par ce que les autres auteurs rapportent de lui : Platon, Aristophane, Xénophon, Aristote.

9. Aristote a exprimé son point de vue sur l'esthétique dans sa *Poétique*, écrite vers 335 avant notre ère, qui concerne surtout la tragédie et l'épopée, [Aristote, 1990].

10. L'œuvre d'Augustin d'Hippone (Saint Augustin) qui traite de l'esthétique est principalement *De l'origine du Bien*, écrite vers 554 de notre ère, mais ce thème est également abordé dans de nombreux autres textes.

11. Dans une forme extrême, la beauté est même parfois expliquée comme une volonté active de l'objet à se manifester auprès de l'observateur. On trouve cette idée dans Schopenhauer : "il est curieux de voir avec quelle insistance le monde végétal en particulier nous sollicite et pour ainsi dire nous contraint à le contempler", (in [Schopenhauer, 1966], p. 259), et Schopenhauer note à ce point qu'il a retrouvé la même idée dans Augustin d'Hippone : "Les plantes offrent leurs formes diverses, qui embellissent notre monde visible, à la perception de nos sens ; (...) elles semblent en quelque sorte vouloir être connues.", Saint Augustin (*De Civitate Dei contra paganos*, XI, 27).

12. Selon les auteurs, cette universalité de perception de la beauté par tout observateur doit être modérée : dans la démocratie grecque, tous les hommes ne sont pas égaux et le discours des philosophes s'adresse aux seuls "καλὸς καγαθός" (*kalos kagathos*), l'homme "beau et bon" digne d'être citoyen d'Athènes ; à la Renaissance les textes s'adressent au petit nombre des lettrés et le temps des Lumières distingue l'Honnête Homme. Le langage de l'art est accessible à tout observateur pour peu qu'il soit "assez sensible" (Kant) pour recevoir le message du philosophe, ce que C. Batteux exprime par : "Tous les hommes sont presque à l'unisson du côté du cœur... Qu'un homme qui ait le goût exquis soit attentif à l'impression que fait sur lui l'Ouvrage de l'Art, qu'il sente distinctement & qu'en conséquence il prononce : il n'est guères possible que les autres hommes ne souscrivent à son jugement." [Batteux, 1747].

13. "... les formes les plus hautes du Beau sont l'ordre, la symétrie, le défini ...", Aristote, Métaphysique M 3.

14. "La Beauté est une certaine harmonie régulière entre toutes les parties d'une chose, harmonie d'une espèce telle que rien ne pourrait être ôté ou ajouté, ou changé en elle sans qu'elle ait aussitôt moins de charme." L. B. Alberti, *L'Art d'Edifier*, Florence, 1485 [Alberti, 1992].

15. La symétrie, au sens grec, (*συμμεθρία*, *summetria*), désigne le contraire de l'*ametria* (sans mesure) et exprime

visible qui reflète l'omniprésence du nombre dans la cosmogonie athénienne. Mais, si le beau absolu est affaire des dieux, les règles d'harmonie sont essentiellement héritées de la nature et en particulier des proportions du corps humain ; la nature, reflet des constructions divines, est le guide premier de l'esthétique.

Tous les auteurs prônent la nécessaire simplicité des œuvres, qui doivent se lire au premier regard.

Par contre, dans le domaine des couleurs, bien des règles anciennes sont perdues et la plupart des œuvres ont disparu. On sait peu de choses des accords chromatiques qui semblaient utiliser une gamme du noir au blanc, partagée harmoniquement par le rouge et le jaune.

Les règles grecques s'expriment à travers les "canons" qui s'appliquent de façon assez rigide aux formes, aux lignes, aux harmonies. Les canons limitent la création artistique et l'innovation n'est pas une vertu de l'œuvre d'art. D'ailleurs, à Athènes, l'artiste est artisan avant d'être artiste et très peu créateur, si ce n'est dans l'art plus noble de la tragédie ou à moins qu'il ne touche à la perfection, comme le firent Praxitèle, Polyclète ou Xéusis.

L'approche objectiviste justifie pleinement que l'on s'attache à déterminer, sans l'aide d'un observateur, les qualités esthétiques d'un objet puisque celles-ci sont immanentes et s'expriment également, avec ou sans observateur. Cette intention est d'autant plus légitime que les règles de l'esthétique sont universelles et largement indépendantes de la forme prise par l'*artefact* : peinture, musique, architecture, sculpture. Et c'est bien ainsi que les conçoivent les penseurs grecs de l'époque hellénistique¹⁶.

Après la Grèce

Rome adopte l'esthétique hellénistique et, si elle l'assouplit, la naturalise, l'immerge dans la société, elle offre aux règles antiques l'immensité du bassin méditerranéen pour galerie. Elle y impose la statuaire et l'architecture grecque pour idéaux et les canons pour règles. Vitruve¹⁷ nous transmettra ces enseignements dont les sources grecques ont disparu.

Mais l'Empire est grand et éclate en deux pôles entre les 6^e et 7^e siècles. Les influences orientales de Byzance la tirent loin de l'esthétique athénienne vers des arts plus convenus tant dans leurs inspirations que dans leurs formes. L'humain le cède au spirituel ; les canons se figent dans des conventions abstraites, réglant les manières, les matières, les formes, les expressions et les couleurs. La nature s'éloigne, les ornements apparaissent, les symboles, les codes reflètent le mysticisme et le religieux. Cette branche de l'esthétique diverge du tronc athénien et ne rejoindra plus le grand courant de l'art occidental.

Au contraire, l'Empire d'Occident s'empare du legs athénien. En Afrique du Nord, Augustin d'Hippone, nourri de l'esthétique hellénistique affirme l'objectivité et l'universalité de la beauté antique et sa capacité à exprimer l'idéal divin : simplicité, harmonie, respect des proportions. La beauté de l'Olympe est offerte au temple de Jérusalem et à l'Église de Rome. C'est la transcendance divine qui justifie l'universalité du Beau, reflet de la grandeur du Dieu unique comme il l'était du Panthéon. La chrétienté pétrie de ces enseignements fera traverser le Moyen Âge aux préceptes de Platon et d'Aristote, à travers les moines-copistes et les ateliers d'artistes. Parallèlement, c'est le même enseignement que l'on professe aux rives du Tigre et de l'Euphrate ou aux lointaines contrées Samanides, où l'on dispute de poésie, de cosmologie et d'algèbre. Au delà de l'empire byzantin, la tradition antique est

donc l'immanence d'une mesure commune et non pas la simple reproduction du motif à l'identique comme aujourd'hui. Comme le dit l'architecte Claude Perrault qui utilise dans son traité de 1673 *proportion* en place de *symétrie* : "Bien que le mot *symétrie* soit devenu français, je n'ai pu m'en servir ici, parce que *symétrie*, en français ne signifie point ce que *Symmetria* signifie en grec et en latin, ni ce que Vitruve entend par *symmetria*, qui est le rapport que la grandeur d'un tout a avec ses parties où la grandeur est différente.", rapporté dans [Dézarnaud-Dandine and Sevin, 2007]. Plus encore, au delà de ces proportions géométriques, la symétrie "est la trace, l'empreinte que le Démenteur a laissé lors de la création du monde.", ([Dézarnaud-Dandine and Sevin, 2007] p. 8).

16. Nous renvoyons le lecteur à l'Annexe 1 pour une analyse détaillée de l'évolution des réflexions sur l'esthétique en Occident au cours des 25 derniers siècles.

17. "Or il n'y a ny Temple ny autre edifice qui puisse avoir grace de bonne structure sans symmetrie et proportion, et si la convenance n'est gardee en toutes ses parties aussi bien qu'en un corps humain parfaitement formé." est traduit (avec son orthographe) en 1545 par Jean Martin, le texte de l'*Architecture* de Vitruve, de la fin du 1^{er} siècle av. JC.

maintenue vivante et sera transmise en retour à Bologne, Montpellier, Coimbra ou Paris au début du nouveau millénaire.

Au tournant de l'an 1000, si les règles esthétiques n'évoluent pas beaucoup, la philosophie athénienne est revisitée, en particulier dans les rapports entre vision et perception. Les penseurs anglais et écossais amorcent le grand bouleversement empiriste qui prône la fin de l'inné et la prééminence du pensé dans la conscience et l'émotion. Le Beau, comme le Bien, n'est pas donné à la naissance, il est acquis par l'expérience et la pratique, par l'exposition à l'exemple. Il faudra plusieurs siècles pour que les pensées d'Okham, de Locke, d'Hutcheson, puis de Hume, imposent un relativisme sceptique (voir par exemple [Ferry, 1990] et [Tatarkiewicz, 1970] pour l'impact de ces penseurs sur la notion d'esthétique).

Pendant ce temps, de ce côté de la Manche, Descartes¹⁸ et Spinoza, affirment la prééminence de la raison, la séparation entre le corps et l'esprit qui préparent l'épanouissement de l'individu, seule référence de la pensée : "*cogito ergo sum*". On est entré dans une période complexe pour l'esthétique où les bases de l'objectivisme sont ouvertement sapées, mais pourtant ses murs tiennent toujours et des argumentaires sont échafaudés pour maintenir les conclusions que la société veut entendre : le beau est universel, intemporel et immédiatement perçu par tous. Pour l'un, c'est une part réservée à Dieu, placée en nous pour juger également du Bien et du Juste, à côté de notre libre arbitre. Pour un autre, c'est une heureuse finalité de la nature qui garantit l'harmonie des jugements. Pour d'autres encore, c'est le fait d'un organe biologique encore inconnu qui "perçoit" le beau comme l'œil perçoit la lumière. C'est le temps de traités savants entièrement dévolus à l'esthétique¹⁹ : Jean-Baptiste Dubos, Charles Batteux, Dominique Bouhours, Jean-Pierre de Crousaz, William Hogarth . . .

C'est aussi le moment où le débat des esprits redescend dans la société : la Querelle des Anciens et des Modernes est l'occasion de confronter les idées qui défendent le classicisme et celles qui prennent le parti de "l'esthétique du sensible", et pour cela toute œuvre se transforme en terrain de bataille : la poésie, la musique, la peinture, la décoration et l'architecture. Devant la montée de l'individu et du subjectif, la position objectiviste est de moins en moins tenue : Johann Joachim Winckelmann en sera le dernier rempart²⁰, tandis que dans le champ de la philosophie, Emmanuel Kant tourne l'argumentaire de sa pensée vers ce qui formera les fondements modernes de l'esthétique tels qu'ils sont encore très largement discutés aujourd'hui [Schaper, 1964, Aquila, 1970, Crowther, 1976, ?, Hopkins, 2001]. Kant est réputé comme la référence première au subjectivisme²¹ ; pourtant sa vision du Beau est profondément empreinte de préceptes objectivistes. Nous allons nous expliquer dans les lignes qui suivent sur nos raisons de placer Kant comme le dernier des objectivistes et le premier des subjectivistes.

Kant et l'esthétique moderne

La complexité de la pensée universelle sur l'esthétique se reflète dans l'œuvre d'Emmanuel Kant et en particulier dans sa *Critique de la faculté de juger* [Kant, 1985] qui prend place peu de temps après sa *Critique de la raison pure* et en utilise de multiples résultats.

Pour Kant, la Beauté est le produit du "jugement du goût" et relève donc à ce titre du jugement et non de la perception. Nous dirions aujourd'hui que c'est un acte cognitif et non perceptif (comme l'est "rouge", par exemple, pur produit de la perception). Mais Kant distingue plusieurs types de jugements cognitifs. "Beau", pour lui, est dérivé de la perception, comme le serait, par exemple, "carré", et l'un comme l'autre sont des produits de l'esprit. Mais "carré" n'est pas un jugement du goût, c'est le résultat raisonné de l'application de "concepts", tandis que le jugement du goût ne fait appel à aucun concept qui "expliquerait" le jugement. Ce n'est que le produit du plaisir éprouvé lors

18. En 1637, "La Dioptrique" nous explique comment voit l'œil [Descartes, 1991], mais c'est l'esprit qui juge du Beau, à partir des indices que transmet l'œil, un Beau toujours universel, intemporel, immuable.

19. On se rapportera à l'annexe 1 pour des détails sur cette période.

20. Johann Joachim Winckelmann est au XVIII^e siècle le grand théoricien du néoclassicisme, défenseur éclairé de l'esthétique platonicienne dans ses traités *Réflexions sur l'imitation des œuvres grecques dans la sculpture et la peinture*, de 1755 et *Histoire de l'Art de l'Antiquité*, de 1764. Il est l'un des ultimes défenseurs de l'objectivisme, en particulier dans son ouvrage : *Traité sur la capacité à ressentir le Beau*, de 1763.

21. On lit par exemple : "*Et, dans la Critique de la faculté de juger, que certains considèrent pourtant comme l'apogée du subjectivisme moderne, (...)*" dans [Ferry, 1990], p. 21.

de l'observation. "Beau", "bon", "juste" partagent cette propriété d'être produits du jugement du goût. D'autre part, "rouge" ne se discute pas. Si vingt personnes me disent que cet objet est rouge que je perçois brun, il me faut questionner mon système visuel. Si vingt personnes me disent que cette silhouette au loin est un chien que je prends pour un loup, nous nous expliquons sur les indices et sur les raisonnements (les "concepts") qui me conduisent à en décider et je m'alignerai probablement sur leur avis, ou ils changeront le leur. Mais si je juge laide une peinture que vingt personnes jugent belle, Kant pense que l'avis des autres ne peut me faire changer d'avis car "*ce qui a plu aux autres ne peut jamais servir de base à un jugement esthétique.*" ([Kant, 1985], §33), tout au plus peut-il me rendre un peu moins confiant dans mon avis²².

Le jugement esthétique est-il objectif ou subjectif, selon Kant ? Même pour Kant la chose est délicate. Il écrit : "*La couleur verte des prairies relève de la sensation objective en tant que perception d'un objet du sens ; mais ce qu'elle a d'agréable relève de la sensation subjective*" (*idem* §3), mais il précise : "*la satisfaction ne suppose pas ici le simple jugement sur l'objet, mais le rapport de son existence à mon état, dans la mesure où ce dernier est affecté par un tel objet. De là vient que l'on dit de l'agréable, non seulement qu'il plaît, mais encore qu'il fait plaisir. Ce n'est pas simplement que j'y applaudis, mais il engendre par là une inclination*". Il complète cela un peu plus loin : "*le jugement du goût est simplement contemplatif, c'est-à-dire qu'il s'agit d'un jugement qui, indifférent à l'existence d'un objet, met seulement en liaison la nature de celui-ci au sentiment de plaisir et de peine.*" (*idem* §5). Le rôle actif de l'objet est ici clair. Kant emprunte également d'autres arguments aux objectivistes : l'œuvre d'art est universellement belle, et, comme nous l'avons vu plus haut, la sensation de beau n'est ni explicable ni démontrable. Il dit d'ailleurs "*Il parlera donc du beau comme si la beauté était une propriété de l'objet.*" (*idem* §6).

Pour Kant, le jugement du goût n'est influencé par aucun intérêt personnel²³ ou émotion, il est antérieur au plaisir donné à l'observation de l'objet qui ne peut donc en être le principe déterminant (*idem* §13).

Kant distingue deux types différents de beauté : la beauté "libre", inconditionnelle et la beauté "adhérente" qui accompagne une fonction particulière (un beau vêtement, un beau bouclier, un bâtiment). L'art est concerné par la beauté libre tandis que dans le cas de beauté adhérente, le jugement de goût est condamné à être impur, c'est-à-dire entaché de motivations qui ne relèvent pas de critères esthétiques (*idem* §16).

Il est assez clair que la pensée de Kant tient une place à part, à cheval entre objectivisme et subjectivisme. Elle tient cette position d'équilibre parce qu'est introduit un *deus ex machina* commode qui permet de faire partager une sensibilité commune à tous les hommes : le "sens commun" (*sensus communis*), un jugement partagé par tous, non sur des fondements rationnels comme le sens commun de Descartes, non sur des critères de majorité d'avis comme dans l'interprétation des stoïciens, mais par une évidence spontanée, universelle et transcendentale qui le rend perceptible immédiatement et par tous²⁴. Il dit ainsi : "*La nécessité de l'adhésion universelle à laquelle il est fait référence dans un jugement de goût est une nécessité subjective qui, sous la supposition d'un sens commun, est représentée comme objective.*" (*idem* §22).

22. Lire sur ce point [Hopkins, 2000].

23. Cette idée d'un art gratuit a trouvé son avocat en Théophile Gautier, qui la défend sous la formule *l'Art pour l'Art* dans son roman *Mademoiselle de Maupin* en 1834. L'idée sera très largement diffusée (et partagée) par O. Wilde, W. Blake ou E.A. Poe. Puis le XX^e mobilisera contre elle, de nombreux auteurs, comme W. Benjamin [Lamarque, 2010] ou artistes, comme W. Kandinsky [Kandinsky, 1954].

24. On lira avec intérêt sur ce sujet le blog de S. Manon, <http://www.philolog.fr/le-sens-commun/>.

L'objectivisme après Kant : du pseudo-subjectivisme au réalisme esthétique

Après le monumental ouvrage de Kant, et avec l'émergence progressive de la phénoménologie²⁵ comme cadre d'intégration des concepts de la perception et de la pensée, la réflexion philosophique s'inscrit très majoritairement dans un cadre subjectiviste. Néanmoins de nombreux penseurs partagent encore certains concepts issus de l'héritage objectiviste.

Georg Wilhelm Friedrich Hegel, comme Kant, s'installe au carrefour des deux systèmes. Mais, lui considère que l'esthétique relève non du monde sensible, mais du monde raisonné, en une sorte de retour au cartésianisme. Pour lui aussi, le Beau existe mais il est le résultat de notre réflexion produit par l'influence directe et objective de l'objet²⁶. Hegel partage aussi l'idée platonicienne que le Beau est universel et atemporel²⁷. Pourtant, nous le verrons plus loin, lorsqu'il parle de son expérience, Hegel a très largement recours à des éléments très personnels et hautement subjectifs qui laissent bien peu de place aux arguments platoniciens.

Arthur Schopenhauer assume également l'héritage de Kant, qu'il croise avec un platonisme et un empirisme sincères. Il soutient que notre connaissance est la conjonction de deux processus aux fondements différents : ce qu'il nomme "intuition", un accès immédiat et spontané à la nature profonde du monde qui nous entoure, et ce qu'il nomme "volonté" et qui relève du principe de raison. Le principe de raison est l'outil majeur de notre construction cognitive. Mais, dans le domaine esthétique, il doit céder la priorité à l'intuition qui seule peut permettre l'accès à la nature vraie et à l'intention de l'artiste si celui-ci a fait preuve de génie. Si la volonté et le principe de raison sont des manifestations subjectives de notre moi, l'intuition est naturellement objective, conformément aux préceptes de Platon²⁸.

Avec l'importance grandissante du rôle de la conscience dans la prise de décision humaine, les interprétations objectivistes du beau se font plus rares. A la fin du XIX^e siècle, c'est la communauté scientifique et positiviste qui défendra l'approche objectiviste : le beau existe et c'est la "machine humaine" qui le traite par une adaptation particulière des circuits mentaux. Cette voie singulière est défendue en particulier par Charles Henry²⁹ [Henry, 1885] ou par Gabriel Séailles³⁰. Mais cette voie originale est plutôt minoritaire dans son siècle.

Il faudra attendre le XX^e siècle pour qu'un important courant esthétique remette en lumière ces théories, prenant à contrepied l'approche subjectiviste très majoritaire. Sa trace se retrouve en France

25. La phénoménologie est une relecture critique de la métaphysique qui essaie d'associer perception et représentation du monde par un retour au concret et à l'intuition. Dans ce cadre de pensée il n'y a plus de séparation entre la nature "vraie" d'un objet et la connaissance qu'on en a. Un objet est constitué de l'ensemble des représentations dont on dispose, qu'elles soient perçues ou élaborées par notre conscience : les *phénomènes*. Initiée par Lambert et Kant au XVIII^e siècle, la phénoménologie moderne a été formalisée par Husserl et dispose de nombreuses variantes dues à Fichte, Schopenhauer, Ricœur, Levinas, Sartre ou Merleau-Ponty.

26. "... l'idée seule est la véritable réalité. [...] Cette vérité n'est pas du tout à prendre dans le sens subjectif selon lequel il y aurait un accord entre une existence et mes idées, mais dans le sens objectif selon lequel moi ou un objet extérieur, une action, un événement, une situation dans sa réalité constituent en eux-mêmes une réalisation du concept." ([Hegel, 1997], p. 176).

27. On lit par exemple "La Beauté est goûtée de toutes les nations, à toutes les époques" dans *L'Esthétique* écrite entre 1818 et 1829.

28. "J'espère avoir montré clairement par ces considérations la nature et l'importance de la condition subjective du plaisir esthétique ; cette condition, nous l'avons vu, consiste à affranchir la connaissance que la volonté asservirait, à oublier le moi individuel, à transformer la conscience en un sujet connaissant pur et affranchi de la volonté, du temps, de toute relation. En même temps que ce côté subjectif de la contemplation esthétique, son côté objectif, c'est-à-dire la conception intuitive de l'Idée platonicienne, se manifeste toujours à titre de corrélatif nécessaire." (in [Schopenhauer, 1966], p. 256.)

29. Nous reparlerons plus loin de Charles Henry car il est très probablement le tout premier à avoir imaginé que la beauté soit détectée par les sciences "dures", cf section 5.1.

30. "La beauté formelle des accords musicaux et des figures géométriques, vient de ce que les rapports logiques et internes du contenu idéal trouvent une représentation sensible. Le fond en tant que tel n'apparaît pas à la conscience, mais c'est cependant sa richesse qui attire l'oreille suspendue, le regard fixé, et occupe l'esprit d'une arithmétique inconsciente." G. Séailles, [Séailles, 1877]

chez Raymond Bayer³¹, Mikel Dufrenne³², puis René Pouivet³³, Jean-Pierre Cometti³⁴, Jacques Morizot³⁵.

On doit cependant au subjectiviste Clive Bell [Bell, 1914] la notion de “forme signifiante” pour expliquer comment certaines conjonctions de lignes et de couleurs nous conduisent objectivement au sentiment du Beau. Ces formes signifiantes sont composées des éléments physiques captés par notre système visuel, mais aussi d'éléments plus subtils que l'on recouvrira plus tard des termes “prédicats esthétiques” (Zemach) ou “expressivité perceptive” (Arnheim). Ils font qu'une scène nous semble joyeuse, poétique ou inquiétante. Ces éléments, que l'on ne saurait traduire formellement en signaux optiques, s'adressent directement à notre sensibilité. Bien établie dans le champ de l'esthétique musicale (un accord de mineur n'est-il pas “évidemment” triste?³⁶), cette idée que des configurations “invisibles” entraînent objectivement des sentiments chez l'observateur, est aussi défendue dans le domaine de l'image, où elle a cependant plus de mal à s'imposer. L'émergence de ressentis psychologiques à la suite d'une expérience de perception mobilise diverses théories de philosophie de l'esprit en particulier autour de la notion de “survenance” telle que proposée par D. Davidson³⁷.

C'est l'école “réaliste” qui accordera la place la plus importante à ces formes signifiantes dans l'élaboration de l'évaluation esthétique. Les réalistes épistémologiques s'accordent sur l'existence d'une réalité extérieure, indépendante de la pensée, sur laquelle se forme notre connaissance. Le réalisme esthétique, s'attache à montrer, tout au long du XX^e siècle, que des expressions esthétiques sont transmises à notre conscience par nos canaux de perception et portent l'information esthétique. Pour eux ces expressions esthétiques, issues de l'objet, sont objectives.

Franck Sibley [Sibley, 1959] et Eddy Zemach³⁸, [Zemach, 1991, Zemach, 2005], affirment ainsi que les prédicats esthétiques sont réels, observables et dépendants des corrélats non-esthétiques³⁹. Zemach en déduit que la beauté est une propriété réelle, objective, de certaines œuvres. Il soutient que les désaccords sur l'appréciation esthétique sont dûs aux différences dans les conditions d'observation. Ces conditions d'observation sont bien sûr celles de l'environnement matériel (lumière, espace, ambiance), mais elles couvrent aussi les conditions épistémiques attachées à l'observateur (son humeur, sa disponibilité, sa connaissance de l'œuvre...). Il affirme que si l'on dispose de ce qu'il appelle des conditions standart d'observation (SOC)⁴⁰, on est amené à toujours juger de façon semblable la beauté d'une œuvre. Zemach affirme ainsi l'invariance du jugement esthétique. Il reconnaît qu'il est difficile de garantir ces SOC, ce qui explique des divergences de jugement. En cas de contestation il propose de recourir aux experts en esthétique qui disposent des SOC par vocation.

Jerrold Levinson⁴¹ s'accorde au réalisme esthétique de Zemach mais interprète différemment les écarts de jugement. Il les explique par la diversité de psychologie des observateurs qui induit une diversité de leurs sensibilités. Il s'inscrit donc résolument dans le subjectivisme, [Hsu, 2009]).

Roger Pouivet, accompagne la position de Zemach, anti-relativiste et non-subjectiviste, adoptant cependant un réalisme plus modéré et met en doute la possibilité de dégager des SOC dans certains

31. Raymond Bayer : *Traité d'esthétique*, Collin, Paris, 1956

32. Mikel Dufrenne, *Esthétique et philosophie*, Klincksieck, Paris, 1980

33. Roger Pouivet : *Le réalisme esthétique*, PUF, Paris, 2006.

34. Jean-Pierre Cometti : *Art, représentation, expression*, Colin, 2006

35. Jacques Morizot, *La philosophie de l'Art de Nelson Goodman*, Ed. Jacqueline Chambon, 1996.

36. On lit sur le web : “Pourquoi cette tristesse du mineur ? Il existe une réponse assez simple à cette question : les intervalles mineurs sont des intervalles plus petits, plus “resserrés” que les intervalles majeurs... ce qui donne cette impression de mélancolie, d'introversion, de mystère.”, (<http://www.musiclodge.fr/article-la-tristesse-du-mineur-37419635.html>), explication qui devrait assurer l'adhésion des convaincus.

37. En esthétique, la survenance des prédicats esthétiques sur les prédicats non-esthétiques se traduit par une corrélation univoque de la façon suivante : si deux objets ont des prédicats esthétiques différents, ils ont nécessairement des prédicats non-esthétiques différents, mais l'inverse n'est pas vrai (voir la discussion de [Bréhin, 2007]).

38. Eddy Zemach a publié de nombreux textes sur l'esthétique, en particulier : *Analytic Aesthetics* en 1970, *Aesthetics* en 1976 et *Real Beauty* en 1997.

39. Ainsi, F. Sibley écrit : “Aesthetic words apply ultimately because of, and aesthetic qualities ultimately depend upon, the presence of features which, like curving or angular lines, color contrasts, placing of masses, or speed of movement, are visible, audible, or otherwise discernible without any exercise of taste or sensibility.”, [Sibley, 1959].

40. SOC = *standart observation conditions*.

41. Jerrold Levinson a publié ses travaux dans *The pleasures of aesthetics*, en 1996.

cas. Il appuie sa réflexion réaliste par l'examen de l'expression du langage courant et en particulier de la forme : "*A est beau*"⁴², montrant ainsi que notre héritage linguistique n'est pas compatible avec une justification subjectiviste de l'évaluation esthétique⁴³.

Nous voyons donc que l'objectivisme rassemble des attitudes philosophiques assez différentes : la position la plus extrême fait dépendre le beau d'un transcendant hors de l'objet lui-même (Dieu, ou le Panthéon, ou la Nature) ; l'objet n'est alors que le réceptacle d'un beau partagé qui s'impose à l'observateur s'il est initié. Plus tard, le Beau est rattaché à l'objet, à sa forme à ses proportions. Puis le rôle de l'observateur est progressivement reconnu. Il est aussi le témoin du Beau qui n'est reconnu que par le plaisir de l'observateur. L'observateur est "touché" par le beau et en atteste l'existence. Par quelles voies ? Descartes et Hegel le justifient par le raisonnement, Kant le ressent par une intuition première, antérieure à tout concept mais distincte des perceptions sensorielles primaires, interprétation qui sera souvent adoptée dans des approches postérieures se revendiquant de la phénoménologie (Bergson, Merleau-Ponty). Enfin, dans d'autres démarches très modernes, la réalité du beau est subtilement tramée dans le signal perçu et affecte de façon objective notre sensibilité lorsque les conditions de son observation sont réunies. Pour les objectivistes, le beau existe, il fait objet d'un quasi-consensus, il perdure dans le temps, on peut en attester.

2.2 L'approche subjectiviste

L'approche subjectiviste, qui ramène tout jugement de valeur à un acte de conscience individuel de l'observateur, ne dépouille pas nécessairement l'objet de sa responsabilité esthétique, comme nous l'avons illustré à travers le cas particulier de Kant. Mais en définissant la beauté par le plaisir qu'éprouve l'observateur, on offre à la grande diversité des "sciences de l'esprit" des possibilités très variées de s'appliquer, avec des conséquences très différentes sur l'interprétation, les plus extrêmes pouvant quasiment exclure l'objet considéré de tout rôle dans la décision esthétique. Précisons ce dilemme. Dans l'interprétation kantienne, l'observateur fournit le "capteur de beauté", rendant le jugement subjectif puisque personne ne peut remplacer l'observateur dans sa décision (on ne convainc pas de la beauté pour Kant). Mais ce que l'observateur kantien met en œuvre est bien peu personnel puisque n'intervient aucun intermédiaire dans cette décision : ni mémoire, ni réflexion, ni raisonnement. Relisons au contraire les raisons⁴⁴ qui décident G.W.F. Hegel à attribuer les plus grandes qualités esthétiques à la Madeleine de Corrège. Il est clair qu'un tel jugement convoque massivement la culture, la sensibilité, l'imagination et la mémoire de l'observateur. Prenez dans la rue le prochain passant et il ne saura retrouver aucune de ces raisons qui ont décidé Hegel. Le subjectivisme de Hegel est ainsi beaucoup plus marqué par l'individu et pourra difficilement se satisfaire de SOC adaptées.

Le paradigme de la beauté universelle et intrinsèque à l'objet a donc été remis en cause au *XVI^e* siècle. En fait, les critiques n'avaient jamais été totalement absentes des disputes philosophiques,

42. Pouivet tient le raisonnement suivant. L'expression usuelle face à un bel objet "A" est de dire "*A est beau*", en attribuant naturellement le sentiment que nous ressentons à A, et non pas à nous, ce que nous ferions si nous avions un sentiment véritablement subjectiviste de l'esthétique. On aurait alors dit : "*J'ai beau*". A l'opposé, on dit "*J'ai mal*" sous l'impulsion du sentiment de douleur en heurtant un meuble, on ne dit pas "*Le meuble est mal*", [Pouivet, 2006].

43. A cet argument, Gérard Genette répond : "*nous nous livrons spontanément à "l'objectivisation" de nos sentiments. C'est une tendance naturelle de psychologie empirique qui conduit à attribuer à un objet comme une propriété objective, la valeur qui découle du sentiment qu'on éprouve à son endroit.*", (cité par [Pouivet, 2006] p. 153)

44. Dans [Hegel, 1838], il écrit : "*Dans les oeuvres du genre le plus élevé, on ne doit pas donner place à d'autres pensées qu'à celles que doit éveiller la situation même. C'est d'après ce principe que la Madeleine de Corrège, à Dresde, me paraît si digne d'admiration, et sera, en effet, éternellement admirée. C'est bien la pécheresse repentante ; mais on voit en elle que le péché n'est pas sérieux, qu'elle avait une âme naturellement noble, et qu'elle n'a pu être capable d'une mauvaise passion ni d'une mauvaise action. Aussi n'a-t-elle pas à se dépouiller de son vrai caractère, mais à rentrer plus profondément en soi. Cette conversion, ce n'est qu'un retour à elle-même. Ce n'est donc pas là une situation momentanée, c'est sa nature entière. Ainsi, dans l'ensemble de la représentation, dans l'extérieur et les traits du visage, l'habillement, le maintien, les objets environnants, l'artiste n'a laissé aucune trace, aucun souvenir qui pût rappeler le péché ou la faute.*".

par exemple celles des sophistes grecs⁴⁵, mais ces critiques prennent au XVI^e siècle une forme très affirmée. En 1588, William Shakespeare dénonce l'universalité du Beau : “*Beauty is bought by judgement of the eye*”⁴⁶ et, en 1591, Giordano Bruno son unicité : “*Alia est pulchritudo unius speciei, alia alterius, alia unius generis, alia alius*”⁴⁷. En 1630, René Descartes exprime sans ambiguïté qu'il ne tient pas le Beau comme réel et objet de rationalité partageable⁴⁸.

C'est au philosophe John Locke que l'on doit l'élaboration de la doctrine “subjectiviste”. Il exprime de la façon la plus nette qu'il n'y a pas d'acquis dans la pensée humaine, que toute pensée naît d'une “réception sensible”, donc de nos observations. Quoiqu'il ne traite pas directement du Beau, on comprend que celui-ci comme les autres *Qualités Secondes* d'un objet (celles qui ne sont pas immédiatement déduites de la perception comme la taille et la couleur, mais résultent de la conjonction de nombreuses qualités premières qui rattachent ainsi une “idée complexe” à l'objet observé), n'existe nulle part en dehors de celui qui le perçoit⁴⁹.

De nombreux arguments qu'il a utilisés ont alimenté la querelles des Anciens et des Modernes, qui fait rage en ce temps dans toute l'Europe. Ses idées sont largement reprises au XVIII^e siècle dans des textes de Denis Diderot⁵⁰ ou de Johann Wolfgang von Goethe⁵¹. Elles affirment que la Beauté est une expérience et non un attribut, que cette expérience est largement le fruit de l'émotion de l'individu, qu'elle est donc personnelle et non universelle, contingente et non immanente. Elle est éminemment fluctuante dans le temps, dépendant des “humeurs” de l'individu. Ce qui est beau ne le sera pas toujours et ce qui ne l'est pas aujourd'hui sera peut-être beau demain. L'idée est résumée dans la formule célèbre : “*La Beauté est dans l'œil de l'observateur*”⁵² qui laisse cependant la porte ouverte à de multiples interprétations.

Parmi les philosophes qui composent avec les interprétations du subjectivisme, on peut placer ceux qui ne quittent l'objectivisme qu'à regret. Luc Ferry par exemple écrit : “*L'esthétique moderne est certes subjectiviste en ce qu'elle fonde le beau sur des facultés humaines, la raison, le sentiment, ou l'imagination. Elle n'en reste pas moins animée par l'idée que l'œuvre d'art est inséparable d'une certaine forme d'objectivité*” ([Ferry, 1990], p. 20), ou : “*Il y a moins de désaccord sur la grandeur de Bach ou de Shakespeare que sur la validité de la physique d'Einstein*” (*idem*, p. 42).

Les tenants de la psychologie expérimentale, suivant la trace de G.T. Fechner au XIX^e siècle, adoptent également un subjectivisme pragmatique. Ils s'attachent à tester systématiquement sur des populations d'observateurs l'universalité et la permanence du Beau ; ils montrent l'étendue du relatif lié au contexte, à l'éducation, à la culture mais peinent à généraliser leurs conclusions à des populations très variées. Ils ont surtout beaucoup de mal à dégager des constantes de critères esthétiques

45. On lit ainsi dans les *Dialexeis*, recueil de controverses du temps de Platon, “*Si l'on demandait à tous les hommes de jeter en pile les choses qu'ils trouvent laides, puis de retirer de la pile les choses qu'ils trouvent belles, je pense que rien ne resterait sur la pile, car tous les gens diffèrent dans leurs opinions.*” (*Dialexeis*, 2, 8, rapporté dans [Tatarkiewicz, 1970]).

46. “*La beauté est achetée par le jugement de l'œil.*” Shakespeare dans *Love's labour lost*, 1588.

47. “*Les divers genres et types de beauté diffèrent les uns des autres*”, G. Bruno, *De vinculis in genere*, III, 638, cité par [Tatarkiewicz, 1970].

48. “*Généralement ni le beau, ni l'agréable, ne signifient rien qu'un rapport de notre jugement à l'objet ; et parce que les jugements des hommes sont si différents, on ne peut dire que le beau, ni l'agréable, aient aucune mesure déterminée*” - *Lettres à Mersenne* du 18 mars 1630 in *Œuvres et Lettres*, Gallimard, Paris, 1953, rapporté dans [Malinowski-Charles, 2004].

49. Ces idées de Locke font l'objet de l'*Essai sur l'entendement humain* de 1689 [Locke, 2009] : “*From whence I think it easy to draw this observation, that the ideas of primary qualities of bodies are resemblances of them, and their patterns do really exist in the bodies themselves ; but the ideas, produced in us by these secondary qualities, have no resemblance of them at all. There is nothing like our ideas existing in the bodies themselves.*” (“Il est aisé, je pense, de tirer de là cette conclusion, que les idées des premières qualités des corps ressemblent à ces qualités, et que les formes de ces idées existent réellement dans les corps, mais que les idées, produites en nous par les secondes qualités, ne leur ressemblent en rien. Il n'y a rien qui correspond à nos idées dans les objets eux-mêmes.”) J. Locke, *An Essay concerning Human Understanding* - Livre II, Chapitre 8 § 15.

50. Les textes de Diderot concernant la critique esthétique sont regroupés dans la rubrique *Les Salons* de [Diderot, 1769].

51. Les idées de Goethe se retrouvent dans “*De l'expérience considérée comme médiatrice entre l'objet et le sujet*” de 1792 et plus encore dans “*Le traité des couleurs*” de 1810.

52. “*Beauty is in the eye of the beholder* (La beauté est dans l'œil de celui qui regarde.)”, phrase de Margaret Wolfe Hungerford dans *Molly Bawn* (1878) souvent attribuée à Oscar Wilde - on ne prête qu'aux riches !

qui convainquent par leur généralité. D.E. Berlyne [Berlyne, 1970, Berlyne, 1971] identifie ainsi des qualités particulières de l'objet esthétique : la nouveauté, l'incongruité et la complexité, les deux premières ne relevant que de l'observateur. Il met aussi en évidence la puissance émotive de la forme par des expérimentations contrôlées (répondant en cela aux questionnements de C. Bell) en s'appuyant sur la notion d'activation physiologique (*arousal*). Il s'attache à montrer la dépendance de l'activation (de l'observateur) en fonction de l'ampleur des paramètres de la forme (de l'objet)⁵³.

D'autres travaux [Reber et al., 2004, Moshagen and Thielsch, 2010], empruntant le formalisme de la Gestalt, mettent en avant la *Prägnanz*⁵⁴ pour expliquer que certaines formes signifiantes activent préférentiellement notre système perceptif. A ce titre, la *Prägnanz* devrait permettre de partager des critères universels et soutenir une approche objectiviste, mais il apparaît qu'elle dépend en pratique d'un grand nombre de paramètres internes à l'observateur, certains relevant de son humeur, d'autres de son tempérament, de sa culture, de son entraînement, d'autres enfin du contexte, tous facteurs qui agissent concurremment sur ses émotions par des effets de masquage ou au contraire d'amplification.. Certains mettent l'accent sur le rôle particulier du système perceptif [Molnar, 1974, Molnar, 1997], d'autres tissent le lien entre esthétique et sémiologie, associant le message de l'artiste au ressenti de l'observateur et conditionnant l'expérience esthétique à la connaissance de "codes" partagés à l'occasion d'une éducation culturelle par nature très subjective [Solso, 1996].

En opposition à ces interprétations modérées du beau, le XIX^e siècle et la période romantique ont donné à l'interprétation subjectiviste des pages très éloignées du platonisme. La beauté a besoin d'un objet, mais sa relation à l'objet est ténue et peut éventuellement se reporter sur un autre objet si le premier s'éloigne ou disparaît⁵⁵. Si poètes et musiciens sont les premiers à suivre cette voie, les peintres s'en emparent également. Dans la société, la rupture avec l'Ecole Classique est brutale et les deux courants sont très largement parallèles et concurrents, chacun "tenant" sa vérité. Le succès des Impressionnistes sur les Pompiers en France consacre le renversement des tendances dans le domaine des arts picturaux. L'éclosion ensuite à travers le monde de nombreux autres mouvements artistiques : Pointillisme, Fauvisme, Cubisme, Surréalisme... installe durablement cette approche centrée sur le spectateur au cœur de la démarche de l'artiste, tandis que se maintiennent ici et là des îlots de Classicisme.

Emblématique de cette démarche romantique, se place Friedrich Nietzsche⁵⁶. Nietzsche nie toute existence au beau comme au laid⁵⁷. L'esthétique se réduit à une acte organique⁵⁸. Nietzsche veut arracher le beau aux règles platoniciennes comme au monde cartésien et clame l'importance des critères dionysiaques (éminemment subjectifs dans leurs excès) devant les critères apolliniens faits de mesure et de réflexion. Dans la foulée, il rejette l'universalité du beau et son intemporalité. Il n'accepte que le beau des élites, celui que la foule (la "*plèbe sensuelle*") ne percevra pas : "*Pulchrum est paucorum hominum*"⁵⁹. La beauté n'est pas faite de canons et d'harmonies, mais de violence et de fange, aux antipodes des préceptes aristotéliens⁶⁰. Mais Nietzsche est complexe et changeant et il renie souvent en fin de vie ce qu'il a encensé dans sa jeunesse⁶¹.

53. Pour Berlyne, la dépendance entre forme et activation suit une loi en "U", passant par un optimum avant de se dégrader pour des formes trop marquées.

54. *Prägnanz* est le terme choisi par M. Wertheimer pour exprimer la facilité de perception d'une figure en raison des relations particulières entre ses composantes (symétrie, parallélisme, proximité, etc.) [Wertheimer, 1922].

55. Les expériences amoureuses du Romantisme fleurissent de tels exemples, mais c'est aussi sur cet argument décliné dans de multiples modalités que se construit avec subtilité *La recherche du temps perdu* (Proust, 1911).

56. C'est surtout dans "*Le crépuscule des idoles*" de 1888 que F. Nietzsche s'exprime sur l'esthétique [Nietzsche, 2005], mais on trouve de nombreuses autres références également dans *La volonté de puissance* ou dans *L'Antéchrist*.

57. "*Rien n'est beau, il n'y a que l'homme qui soit beau : sur cette naïveté repose toute esthétique, c'est sa première vérité. Ajoutons-y dès l'abord la deuxième : rien n'est laid si ce n'est l'homme qui dégénère, - avec cela l'empire des jugements esthétiques est circonscrit. - Au point de vue physiologique, tout ce qui est laid affaiblit et attriste l'homme.*", dit-il dans [Nietzsche, 2005].

58. L'art comme fonction physiologique : "*Livrer l'art à la physiologie revient à ravalier l'art au niveau des sécrétions gastriques.*", ou l'évaluation du beau comme pratique du sport : "*On peut mesurer au dynamomètre l'effet de la laideur.*", [Nietzsche, 2005].

59. "*La beauté est pour un petit nombre*", F. Nietzsche, *L'Antéchrist*, 1895.

60. "*La beauté est difficile. Gardons nous de la beauté ! Osons, mes amis, osons être laids. Wagner l'a osé ! Remuons devant nous sans crainte le limon des harmonies les plus rebutantes !*", [Nietzsche, 2005].

61. A titre d'exemple, Nietzsche défendra Corneille contre Hugo, en rupture avec les lignes ci-dessus.

On retrouve de nombreux éléments de la pensée nietzschéenne dans la très moderne *Posthuman Aesthetics* de Robert Pepperell qui réclame de la part de l'art une rupture avec les compromis de la société⁶².

A la fin du XIX^e siècle, l'école psychanalytique apporte sa pierre au subjectivisme, rejette tout argument objectiviste du Beau. Elle adopte une interprétation rigoureuse de la Beauté qui n'en appelle qu'à l'individu et à son expérience passée, enfonçant au plus profond ses ressorts dans le brouillard de l'inconscient. Sigmund Freud écrit : "... la beauté dérive du domaine de la sensibilité sexuelle"⁶³, en écho à Nietzsche qui affirmait aussi l'origine sexuelle du sentiment du beau : "Sans un certain surchauffement du système sexuel, on ne saurait imaginer un Raphaël"⁶⁴.

Il faudra attendre la fin du XX^e siècle pour qu'une pierre nouvelle soit ajoutée au subjectivisme avec l'avènement de la neuro-biologie. S'appuyant sur les nouvelles techniques d'imagerie cérébrale (IRM fonctionnelle et MEG), de très nombreux travaux permettent de mieux comprendre le fonctionnement de notre pensée. Les résultats obtenus en physiologie, au milieu du XX^e siècle, avaient permis d'expliquer le rôle des voies visuelles et du cortex postérieur dans la perception. La neuro-biologie nous éclaire aujourd'hui sur le rôle du néo-cortex et du centre limbique lors de ces tâches de perception, d'évaluation et de prise de décision. Elle souligne la place des émotions dans l'élaboration de notre interprétation. Sous cet éclairage, un certain nombre de philosophes, d'historiens d'art ou d'artistes, associés à des neuro-biologistes, se sont regroupés dans un volet nouveau de l'esthétique, la neuro-esthétique. qui a suscité une très abondante littérature depuis 30 ans, parmi laquelle les travaux séminaux de [Zeki, 1999, Livingstone, 2002, Kawabata and Zeki, 2004], ceux plus récents de [Cupchnik et al., 2009, Di Dio and Vittorio, 2009] ou ceux de [Di Dio et al., 2011, Ishizu and Zeki, 2011, Kuehn and Gallinat, 2012]. Ces travaux feront l'objet des développements du chapitre 3.

Ces résultats ont durablement ancré l'interprétation subjectiviste de la Beauté pour une large part des penseurs. Ils soutiennent des conclusions qui, dans leurs formes extrêmes, suggèrent que :

- la Beauté est une émotion personnelle, c'est-à-dire une expérience psycho-physiologique complexe en réaction aux influences biochimiques et environnementales transmises par nos sens ;
- le jugement esthétique (*esthetic appraisal*) est distinct des autres formes de jugement fournies par notre cortex (par exemple le jugement sur la nourriture) car la beauté n'est pas vitale pour l'être humain (à l'inverse par exemple de la nourriture) et nous n'avons aucun "intérêt" physiologique à l'exprimer [Beardsley, 1966] ;
- en cela, la matière artistique pourrait être traitée par des circuits différents et spécialisés auxquels est parfois donné le nom "d'aires hédoniques"⁶⁵.

En suivant J.P. Changeux [Changeux, 2016], on pourrait dire que le subjectivisme a lentement migré d'un sensualisme (à la façon de Condillac) et d'un empirisme issu des philosophes écossais et britanniques (Locke, Hume), vers une cérébralisation, domaine où se retrouvent les neuro-biologistes. Ils rejoignent sur ce terrain Martin Heidegger qui affirmait : "Du moment que l'art n'est plus que du ressort de la physiologie, l'essence et la réalité de l'art se dissocient en des états nerveux et en des processus dans les cellules nerveuses."

C'est aussi avec ce point de vue que nous retrouverons nombre d'informaticiens cognitivistes, bien

62. Dans son "Manifeste Posthumain", R. Pepperell affirme, entre autres :

- " (...)
- *Rich aesthetic experience is generated by the perception, simultaneously, of continuity and discontinuity in the same event.*
- *All stimulating design relies on balancing the relative quotients of order and disorder in the object. This also goes for the composition of music and literature. However, such judgements cannot be made in isolation from the fact that values of order and disorder are largely prescribed by social agreement.*
- *Posthuman art uses technology to promote discontinuity. Healthy societies tolerate the promotion of discontinuity since they understand that humans need exposure to it, in spite of themselves. Unhealthy societies discourage the promotion of discontinuity.* [Pepperell, 2005]

63. in S. Freud, *Le Malaise dans la culture*, Presses Universitaires de France, Paris, 1985.

64. in F. Nietzsche, *La volonté de puissance*, textes inachevés, de 1888, regroupés dans les *Kritische Gesamtausgabe*.

65. Ainsi, Francis Hutcheson, vers 1730, considère que le Beau est ce qui est saisi par notre sens interne du beau, comme le visible est ce qui est saisi par notre œil. Il attribue ainsi à l'Homme des organes de perception hédonique que de nombreux travaux ultérieurs chercheront en vain à localiser (voir section 3.2).

loin des présupposés platoniciens adoptés par les tenants du “big data”. Nous aborderons ce dilemme d’informaticiens dans les derniers chapitres.

2.3 Subjectivisme et Objectivisme : un débat toujours vivant

L’importance des phénomènes subjectifs d’appréciation des qualités esthétiques d’une image ne peut être aujourd’hui écartée et le rôle de l’observateur, doté d’une sensibilité particulière, d’une éducation, d’une culture, est incontournable. Mais par ailleurs, de nombreux témoignages attestent de l’existence de consensus objectifs sur la beauté de certains objets, ce consensus indiquant que des “indices de beauté” émanent de cet objet qui sont très largement reconnus et actifs sur une large part de la population. Devant cette double évidence, nous sommes au cœur du dilemme que tente de cerner ce texte et nous écarterons d’autres approches du Beau qui nous éloigneraient de ce débat particulier⁶⁶.



Figure 2.1 – Ces photos sont-elles belles ? C’est à cette question qu’essaie de répondre ce texte. Plaçant un vaste jury devant cette interrogation, un certain consensus se manifeste souvent, qui conforte la démarche de ceux qui pensent mesurer cette beauté par ordinateur. Si tout le monde perçoit dans les deux images de gauche des critères de beauté plus évidents que dans l’image de droite, généralement perçue comme ordinaire, pourquoi l’ordinateur (capable de remarquables performances de reconnaissance des formes) ne les détecterait-il pas ? La photo de gauche (Poppy d’Irving Penn) est signée d’un photographe célèbre et l’on pourrait objecter qu’elle bénéficie de cette référence (c’est l’effet de “signature” analysé par exemple dans [Changeux, 2016]), mais la photo du centre est anonyme (issue du web) et n’est appréciée que par ses qualités propres. Par ailleurs ces trois photos représentent des fleurs et héritent donc du même contexte culturel et sémantique pour la plupart des observateurs. Enfin, elles ne font appel à aucun des ressorts qui mettent en jeu de l’intérêt ou de la surprise, autres ingrédients potentiels de différenciation comme nous l’avons vu en section 1.1.

Il semble aujourd’hui assez largement partagé que les objets reconnus comme “beaux” possèdent des propriétés intrinsèques qui favorisent cette reconnaissance pour tout observateur (voir par exemple la figure 2.1), mais que l’observateur reste celui qui décide *in fine* de ce qu’il juge beau ou non. On remet ainsi au centre de l’interprétation du Beau des visions que l’on a déjà rencontrées chez Crousaz ou chez Kant qui tentent d’accomoder le subjectivisme qui s’impose depuis cinq siècles et l’objectivisme qui a régné pendant vingt siècles. Ces approches n’émergent pas au temps des Lumières,

66. Parmi les courants de pensée les plus importants de l’esthétique contemporaine, nous écarterons l’esthétique fonctionnelle qui a pris une place considérable au cours du XX^e siècle, en particulier pour sa relation avec le *design* industriel et l’architecture. Suivant une idée déjà avancée par les stoïciens, elle n’attribue de qualité esthétique qu’en fonction de la capacité de l’objet à remplir une fonction. Elle a soutenu de grands mouvements artistiques (Blau Reiter, Suprématistes, ...). Elle a été largement exposée dans les œuvres de Jacques Viénot, Gilbert Simondon ou Mikel Dufrenne, en France. L’esthétique fonctionnelle n’est pas adaptée pour juger de la beauté des photographies.

A également été écartée l’école esthétique dite contextualisée ou historique, illustrée par Ernst Gombrich ou Arthur Danto, qui suggère qu’une part importante des qualités esthétiques de l’œuvre d’art est fournie par le contexte artistique, politique, social et familial qui a entouré sa naissance. La connaissance du contexte permettrait seule d’apprécier l’intention artistique d’où découlerait le jugement de beau. Dans [Bullot and Reber, 2013], une synthèse est tentée entre l’interprétation contextualisée et une approche purement subjective de la beauté faisant appel à la neuro-esthétique. Néanmoins, il nous semble que cette démarche n’est pas adaptée pour justifier du jugement esthétique dans une approche de traitement par ordinateur.

car elles ont toujours été présentes dans les débats sur l'esthétique, mais bien souvent leur caractère contradictoire n'est pas relevé et prête rarement à discussion. Citons le seul exemple de Guillaume d'Auvergne, qui, au début du XIII^e siècle, écrit successivement : “*En ce qui est indubitablement essentiellement beau et élégant, l'essence profonde et la qualité est la beauté*⁶⁷” et “*Une chose belle en elle-même peut être laide près d'un autre objet, ou avec un autre objet ou dans un autre objet*⁶⁸”. Des positions semblables, difficiles à concilier et qui appellent des solutions de compromis, peuvent être relevées dans de nombreux textes depuis l'antiquité grecque.

De nos jours, les approches phénoménologiques (que nous avons évoquées précédemment) qui associent intimement l'objet observé à l'esprit observant à travers les “phénomènes”, proposent parfois de telles synthèses. Si l'objet est en charge de fournir les “prédicats esthétiques”, ceux-ci seraient transformés en “*qualia*”⁶⁹. Ces qualia seraient alors des porteurs de l'appréciation esthétique, personnels, ineffables (intransmissibles), immédiats et liés à l'expérience perceptive seule et appréhendés directement par la conscience [Dennett, 1993]. Mais ces interprétations, qui s'opposent fortement à la séparation esprit/conscience défendue par les rationalistes sont encore l'objet de nombreuses discussions.

Les indices de beauté seront alors plus ou moins efficaces pour créer l'émotion et le plaisir qui accompagnent le sentiment de beauté. C'est là un point central de l'étude que nous ferons. De quoi dépend cette efficacité? On peut tout d'abord invoquer l'objet qui a justifié la photographie. Est-il beau? Ou ne l'est-il pas? Il est évident que répondre à cette première question nécessite que l'on réponde à tout notre problème, puisque, invoquant la beauté de l'objet pris pour modèle, on reporte sur cet objet des critères que l'on pourrait attribuer à sa photographie (voir la figure 2.2). Ensuite, évoquer des “prédicats esthétiques” nous invite à en faire l'inventaire. On aura alors recours aux arguments issus de la pensée platonicienne : simplicité, symétrie, harmonie, éventuellement associés aux nuances aristotéliennes pour les humaniser. On pourra aussi exploiter les critères de la Gestalt (symétrie, parallélisme, proximité, etc.) [Reber et al., 2004, Moshagen and Thielsch, 2010] et les fondements de la *Prägnanz* [Wertheimer, 1922], on pourra en appeler aux critères plus complexes dégagés par Berlyne : la nouveauté, l'incongruité et la complexité [Berlyne, 1970]. Mais on s'attend aussi à faire appel à un grand nombre de paramètres internes à l'observateur, certains relevant de son humeur, d'autres de son tempérament, de sa culture, de son entraînement, d'autres enfin du contexte, tous facteurs qui agissent concurremment sur ses émotions par des effets de masquage ou au contraire d'amplification [Dufrenne, 1980, Solso, 1996, Changeux, 2016].

67. “*Hoc indubitanter essentialiter pulchrum est et decorum, et ujus essentia et quidditas ejus est pulchritudo.*” Du bien et du Mal, 206, (Pouillon, 317), cité par [Tatarkiewicz, 1970].

68. “*Ipsum quod est in se decorum, est turpe ad aliquid, aut cum alio, aut in alio.*”, Du bien et du Mal, 207, (Pouillon, 317), cité par [Tatarkiewicz, 1970].

69. Pour les philosophes qui les ont adoptés, les qualia (au singulier “*quale*”) sont la trace de l'expérience phénoménale. Les qualia ne peuvent exister hors de l'expérience sensible et sont incommunicables (je ne peux expliquer le “rouge” à un non-voyant).

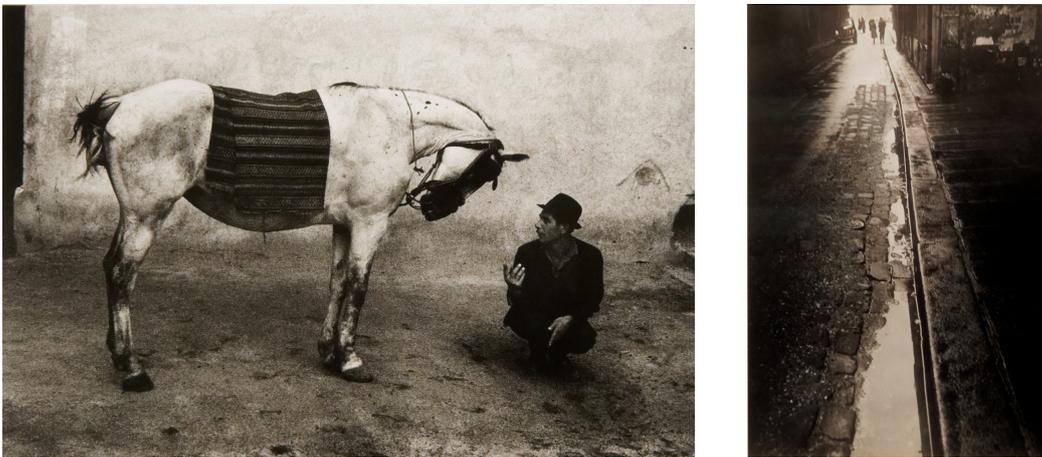


Figure 2.2 – *Pourquoi une photo est-elle belle ? On pourrait dire de la photo de gauche (Romania de Josef Koudelka) que sa beauté est issue de son objet ; ici le cheval, lui-même est beau. Mais outre que cette remarque ne fait que déplacer notre interrogation sur les attributs de la beauté, elle ne supprime pas l'essence esthétique de la photo qui s'impose nettement. // Sur la photo de droite (Rive gauche de Gordon Parks), l'argument précédent ne tient pas car l'objet lui-même (une ruelle à contre-jour) ne tire ses qualités esthétiques que de l'image qui en est faite, de sa composition et de ses jeux de lumières. Les peintres ne s'y sont pas trompés qui ont progressivement remplacé les odalisques et les phaëtons par des cruches et des compotiers.*

Chapitre 3

La neuro-biologie ou l'arbitre des consciences

Les études qui ont permis de connaître les mécanismes de la vision humaine sont multiples. Les études cliniques portant sur les blessés lors d'interventions chirurgicales, complétées parfois des études anatomiques *post mortem*, ont ouvert la voie permettant de localiser des régions importantes pour certaines étapes de perception, d'analyse ou d'interprétation du signal visuel. Puis des travaux d'électrophysiologie sur l'homme (électro-rétinographie, électro-oculométrie, potentiels évoqués), mais surtout ceux conduits par des techniques plus invasives sur les oiseaux, les chats et les singes, ont permis de vérifier *in vivo* la validité des modèles de fonctionnement des voies optiques et des aires primaires de la vision.

Ces voies sont représentées sur la figure 3.1.

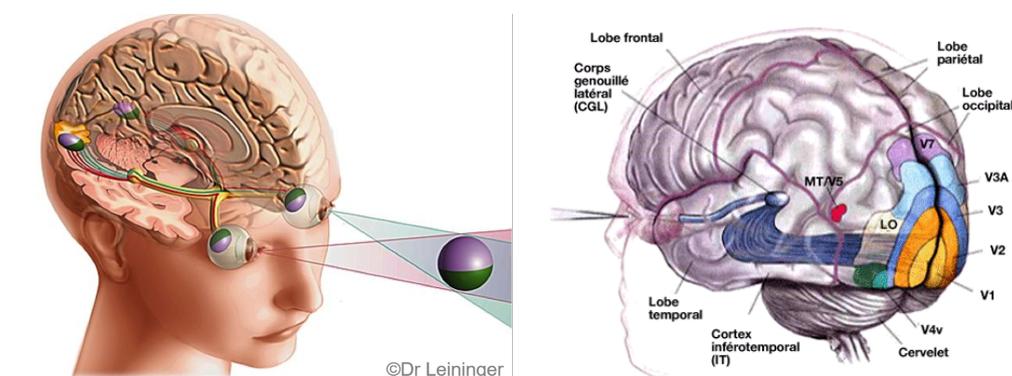


Figure 3.1 – A gauche : schéma des voies optiques qui conduisent de l'œil aux aires visuelles en passant par le chiasma optique, puis les corps genouillés latéraux (crédit : Dr Leininger). A droite, les aires visuelles. Certaines aires ont des fonctions bien identifiées : V1 pour le traitement des contrastes, V2 pour le contraste et les mouvements, V3A pour l'orientation et le mouvement, V3B pour la profondeur, V4 pour la reconnaissance des visages, de l'écrit, des couleurs, etc. (crédit : Wikipédia)

Il a fallu cependant attendre la disponibilité des techniques d'électro-encéphalographie (EEG), de magnéto-encéphalographie (MEG), puis de l'imagerie par résonance magnétique (IRM) et de la tomographie par émission de positons (TEP), pour accéder au fonctionnement des aires supérieures de la vision, en temps (presque) réel.

Les voies montantes visuelles sont bien évidemment les plus concernées pour acheminer le message issu de l'observation d'une photographie vers les aires supérieures, mais elles ne sont pas les seules.

M. Imbert a montré très tôt que des voies directes permettent de court-circuiter les aires visuelles de façon à transmettre l’observation vue au cortex moteur et constituer notre réponse proprioceptive. Le rôle de ces voies à la perception du monde qui nous entoure est cependant encore mal connu et elles sont peu intégrées à nos schémas de vision.

3.1 Les protocoles d’IRMf et la neuro-esthétique

Ce sont surtout les travaux menés en résonance magnétique fonctionnelle (IRMf) qui ont permis de faire les progrès les plus décisifs dans ce domaine. Rappelons que cette technique, mise en place dans les années 1990, repose sur la détection de la consommation d’oxygène (que l’on rattache à l’activité circulatoire) par les zones cérébrales actives. Cette consommation est faible et se traduit par une très faible variation du signal détecté par IRM (de l’ordre du pourcent). Pour cela, les protocoles employés nécessitent que le sujet examiné se livre plusieurs fois de suite au même exercice, de façon que l’on puisse intégrer plusieurs mesures. Comme il faut aussi que ces protocoles permettent d’isoler les signaux recherchés, il convient d’alterner les étapes de détection avec des étapes conçues pour identifier la consommation d’oxygène au repos, ou celle due à d’autres activités cérébrales du sujet, indépendantes du stimulus de test. La réponse recherchée est alors obtenue par différence entre ces deux séries d’expériences. Voyons sur l’exemple d’une expérience fictive comment fonctionnent ces protocoles. Supposons que l’on cherche à identifier si une couleur, le rose, dispose d’un circuit particulier de traitement dans notre cerveau. On produit une série d’images très variées, certaines contenant des plages roses, d’autres n’en contenant pas. On les présente successivement, dans un ordre aléatoire, pendant quelques secondes, à un observateur placé dans un imageur par résonance magnétique. S’il existe dans le cerveau un “processeur de rose”, alors on s’attend à voir apparaître une zone active dans toutes les images IRMf produites lors de la présentation d’une image contenant du “rose”. Ce processeur sera spécifique du rose si la même expérience répétée avec les plages roses converties en violet “n’allument” pas les zones précédemment actives.

Pour généraliser ces résultats, il convient de répéter chaque expérience sur plusieurs sujets pour s’assurer d’un consensus sur la réponse, puis de moyenner les réponses obtenues. Mais ces moyennes devant prendre en compte la variabilité des morphologies cérébrales, on procède à un recalage de chaque carte cérébrale dans un repère de référence (par exemple le repère de Talairach), avant de procéder au moyennage.

Ces protocoles complexes sont bien sûr l’objet de nombreuses discussions et critiques, et tous les résultats doivent être examinés avec une grande circonspection. Cela explique que de nombreuses études aient été menées avec des objectifs assez proches mais des protocoles différents. C’est en particulier le cas des études qui s’attachent à suivre les processus cérébraux associés à des signaux en fonction de leurs qualités esthétiques. Ces travaux ont été regroupés depuis une dizaine d’années dans une branche des sciences spécifique, dénommée “neuro-esthétique” [Skov, 2009, Vidal, 2011] et dédiée spécifiquement aux bases neurales de la beauté dans les diverses modalités de la perception : la vision bien sûr (avec la peinture, la photographie, l’architecture, la statuaire), mais aussi l’audition (avec la musique et le chant) et aussi, dans de très rares cas, le goût et l’olfaction. Un objectif important de la neuro-esthétique pour de nombreux auteurs a été de “naturaliser” les circuits d’évaluation de la beauté, c’est-à-dire de montrer qu’ils utilisent des fonctions du cortex déjà établies dans le cerveau pour d’autres fonctionnalités et en particulier des fonctionnalités cruciales pour la survie de l’espèce (ce que les jugements esthétiques ne sont pas) [Brown et al., 2011], ces fonctionnalités cruciales pouvant justifier l’existence de ces structures dans l’espèce dans un cadre darwinien de l’évolution.

3.2 Que nous dit l'IRMf des “processeurs de beauté” du cerveau ?

Nous ne saurons évoquer ici tous les travaux consacrés à la neuro-esthétique. Une revue de 2009, [Di Dio and Vittorio, 2009], en citait une trentaine, une autre [Chatterjee and Vartanian, 2016] près de 200, en 2016, mais une recherche en 2017 sur le mot *neuroaesthetics* en rapporte près de 3000. Nous ne visiterons pas ces travaux séparément. Nous isolerons les principales zones cérébrales concernées en renvoyant le lecteur à quelques articles qui expliquent comment elles ont été mises en évidence.

3.2.1 Le rôle du cortex préfrontal

Le cortex préfrontal (en couleur sur la figure 3.3) a été le premier candidat à la fonction de “centre hédonique” dans des travaux qui ont souligné très tôt son implication lors de l'observation d'œuvres d'art [Kawabata and Zeki, 2004, Jacobsen et al., 2006, Cela-Conde et al., 2011]. S'il s'est vu adjoindre ensuite d'autres aires très importantes, il demeure l'une des zones maîtresses de la vision esthétique. L'OFC (cortex orbito-frontal) est connu pour son rôle important dans les tâches de décision. C'est un point de passage de la plupart des voies perceptuelles et un élément important du système en charge de l'évaluation (*appraisal*) [Kringelbach, 2005, Wallis, 2007] ainsi que de la régulation des “récompenses” (*rewards*). A ces titres, il pourrait être une pièce centrale de la fonction de plaisir [Berridge and Kringelbach, 2013].

Nous illustrerons ce rôle par les travaux d'Ishizu et Zeki [Ishizu and Zeki, 2011] qui constituent un socle pour les études de neuro-esthétique et ont permis à leurs auteurs de proposer une “théorie de la beauté” qui sert souvent de référence. Dans leurs expériences, des sujets sont exposés dans un imageur IRMf à une série de stimulus (des photos et des musiques), classés préalablement en “beaux”, “indifférents” ou “laid”. Le dépouillement consiste alors à comparer les zones activées selon la “qualité” du stimulus.

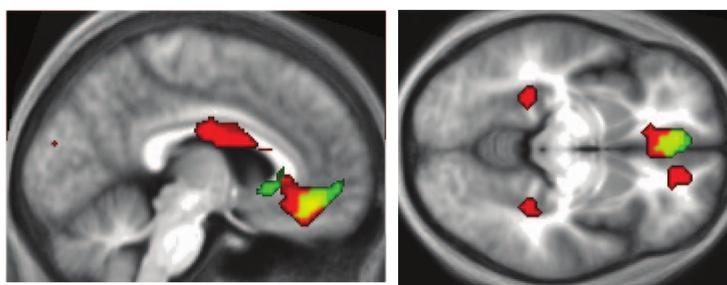


Figure 3.2 – Dans l'expérience en référence [Ishizu and Zeki, 2011], Ishizu et Zeki ont montré que seule la partie centrale du cortex orbito-frontal (la tache à droite dans l'image gauche) était systématiquement activée lors d'expériences avec des signaux visuels ou acoustiques porteurs de qualités esthétiques. Les noyaux caudés peuvent également être actifs dans certaines expériences. Expériences conduites avec 21 sujets. Les zones en rouge sont excitées par un signal visuel, les zones en vert par un signal acoustique tandis que les zones en jaune sont excitées par les deux types de signaux.

Les résultats ont montré une activité systématique et préférentielle de la partie centrale du cortex orbito-frontal (OFC) chaque fois qu'un document “beau” est présenté plutôt qu'un document “indifférent” ou “laid” (figure 3.2). Ces résultats, déjà cités par ailleurs, ont été régulièrement retrouvés (voir la synthèse de [Brown et al., 2011]).

Par ailleurs, pour autant qu'il est possible de le déduire d'un petit nombre d'expériences, l'ampleur de l'activation est directement fonction de la beauté du document telle qu'elle a été appréciée lors de la phase préalable de calibrage.

Ces expériences semblent donc ancrer biologiquement les notions d'esthétique et de plaisir, un

point régulièrement abordé par la philosophie du Beau.

Enfin, il apparaît qu’aucune activation significative ne permet de distinguer la “Laideur”. Ce point tente de répondre à une question importante en philosophie de l’esthétique concernant l’existence ou non d’une valence négative à la Beauté¹.

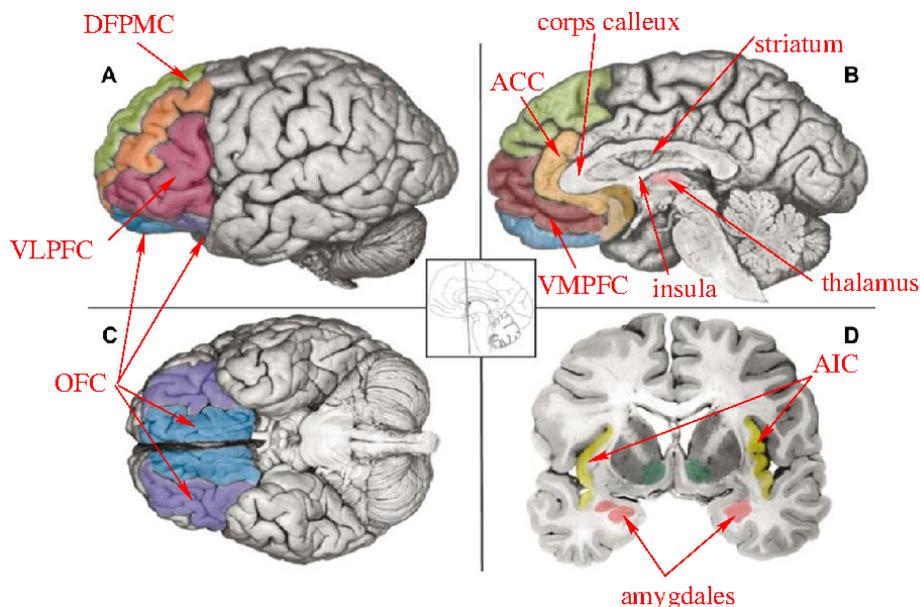


Figure 3.3 – Les aires concernées lors de l’expression d’émotions, d’après [Chatterjee and Vartanian, 2016]. Le cortex orbito-frontal (OFC) est la zone en violet des vues A et C. Le cortex insulaire (AIC) est en jaune sur la vue D. Il est relié de façon très étroite aux aires latérales de l’OFC. Les circuits viscéro-moteurs comportent la zone ventrale du cortex ventro-médian préfrontal (VMPMC) (en bleu sur A, B et C). Le VMPMC est étroitement lié à l’amygdale (en rose dans la vue D). Le cortex cingulaire antérieur (ACC) est en beige sur B, en avant du corps calleux. La zone dorsale du cortex préfrontal est associée au contrôle des états mentaux : partie dorsale du VMPMC, zone frontale (en marron), cortex dorso-médian préfrontal (DMPFC) (en vert en A et B). Le cortex ventro-latéral préfrontal (VLPFC) est en rouge sur la vue A. Le thalamus est en rose au centre de la vue B et le gyrus frontal moyen est en orange.

L’étude citée identifie d’autres régions également activées, en particulier les noyaux caudés (dont le rôle pour réguler les émotions est connu) ou le gyrus supra-marginal (impliqué dans la perception de l’espace), mais de façon non systématique. Nous les retrouverons par ailleurs.

3.2.2 Le rôle du cortex insulaire

Les travaux de S. Brown *et al.* [Brown *et al.*, 2011] ont fait la synthèse des études disponibles en neuro-esthétique en 2011 de façon à retrouver, si elles existent, les aires communes à l’évaluation esthétique à travers les diverses modalités de perception (vision, audition, goût et olfaction). C’est donc là aussi la recherche de la “fonction hédonique” qui est en jeu, mais que l’on cherche à “factoriser” pour toutes les modalités. Les résultats de ces travaux sont bien corroborés par des études plus récentes qui les précisent un peu [Boccia *et al.*, 2016]. La synthèse a porté sur 93 études de neuro-imagerie (IRMf ou TEP exclusivement) concernées par l’étude des zones du cerveau actives lors de l’exposition

1. Beaucoup de traités de l’esthétique ne reconnaissent pas de valence négative à la Beauté, tandis que le sens commun accepterait la candidature de la “Laideur” dans ce rôle. Mais il faudrait alors admettre des attributs de “Laideur”, ce qui serait distinct de l’absence de critères de “Beauté”, position la plus communément adoptée. La notion de “laid” est très peu couverte dans la littérature de l’esthétique (surtout si on la compare à la notion de “mal” dans la littérature de l’éthique) proposons cependant deux références [Rozenkranz, 1853, Bénard, 1877]. Notons donc ici que la neuro-biologie mettant en évidence des zones d’évaluation aussi sensibles aux valences positives que négatives (notamment le cortex orbito-frontal) tente d’apporter des réponses à cette question.

à des stimulus esthétiques positifs dans les quatre domaines de perception. Les articles concernant la vision étaient nettement majoritaires et nous insisterons ici sur les résultats qui les concernent. Toutes ces études portaient sur des personnes saines et fournissaient des résultats pouvant être rapportés dans un même atlas de référence (Talairach ou MNI)².

L'étude a consisté à examiner la conjonction des activations entre auteurs et entre modalités dans l'ensemble des études, à l'aide de protocoles bien établis³.

Bien que l'on retrouve le rôle important de l'OFC dans la plupart des observations (figure 3.4), il n'est pas engagé dans le jugement esthétique en olfaction ou en goût et, pour les auteurs de [Brown et al., 2011], il n'est donc pas systématiquement associé au jugement esthétique. Les auteurs ne lui attribuent donc pas la primauté de la fonction hédonique.

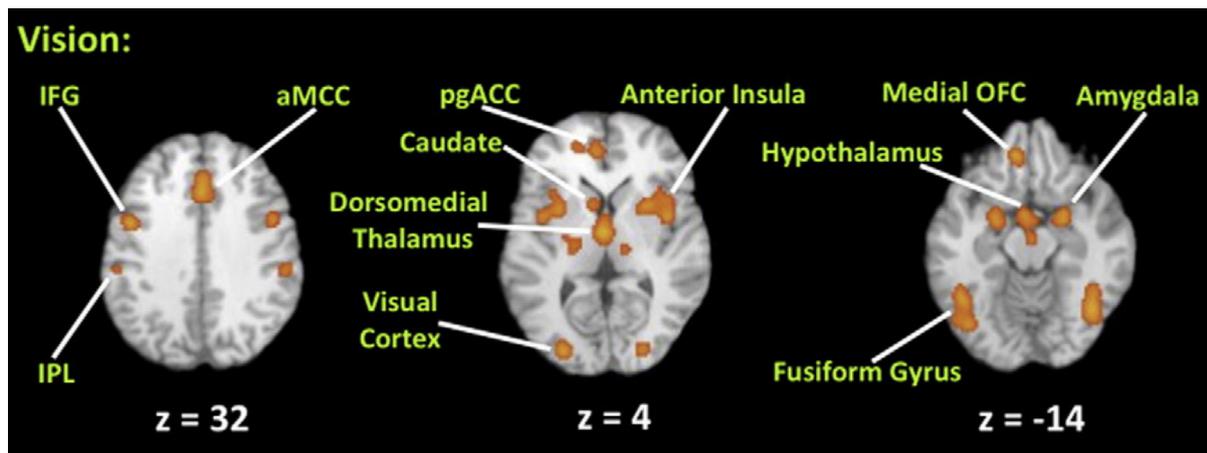


Figure 3.4 – Les principaux foyers d'activité de la vision lors des expériences rapportées dans [Brown et al., 2011]. Le repère choisi est celui de Talairach. La présentation est faite dans des coupes axiales dont la coordonnée verticale est indiquée par z . Les zones actives ont une probabilité seuillée à 5%. Les abréviations signifient : IFG = gyrus frontal inférieur, IPL = lobe pariétal inférieur, aMCC = cortex cingulaire médian antérieur, pgACC = cortex cingulaire antérieur prégenouillé, OFC = cortex orbito-frontal. Caudate désigne le noyau caudé.

Cette étude conclut au contraire à l'importance première du lobe droit du cortex insulaire antérieur, siège d'une très importante activation lors d'un stimulus esthétique quelle que soit la modalité (audition, vision, goût, olfaction). Ce serait la seule aire spécifiquement active pour toutes les modalités. Seule cette aire serait systématiquement activée par les stimulus esthétiques positifs et il semblerait que ce soit là que se situent les mécanismes les plus spécifiques liés à l'appréciation esthétique.

On retrouve des conclusions concordantes dans [Cupchnik et al., 2009] (figure 3.5).

Cette aire est cependant activée également dans une grande variété de situations où la dimension esthétique n'est pas engagée et son rôle exact est encore très confus. Elle est en particulier réputée pour être active dans les situations de jugement négatif (dégoût, douleur, empathie, etc.), ce qui nous éloigne sensiblement de nos conclusions. Elle participe aussi à la conscience introspective (maintien de l'état du corps humain, homéostasie, conscience de soi, etc.) activités qui ne nous rapprochent pas beaucoup du jugement esthétique. Enfin elle semble avoir un rôle régulateur dans l'expression des émotions par l'amygdale, ce qui est probablement plus proche de notre sujet.

Dans les travaux de [Lacey et al., 2011], on précise l'importance des zones en charge des "récompenses" :

2. En raison de la diversité géométrique des anatomies cérébrales, il n'est pas possible de superposer les aires fonctionnellement identiques d'individus différents sans procéder à des transformations géométriques. L'atlas de Talairach est la référence la plus ancienne [Talairach and Tournoux, 1988] ; elle est très adaptée pour les structures superficielles. On lui préfère parfois l'atlas MNI (du nom du Montreal Neurological Institute de l'université McGill) plus représentatif de la variété des populations et plus adapté aux structures profondes [Evans et al., 1992].

3. Le protocole choisi est GingerALE : ALE=Activation Likelihood Estimation.

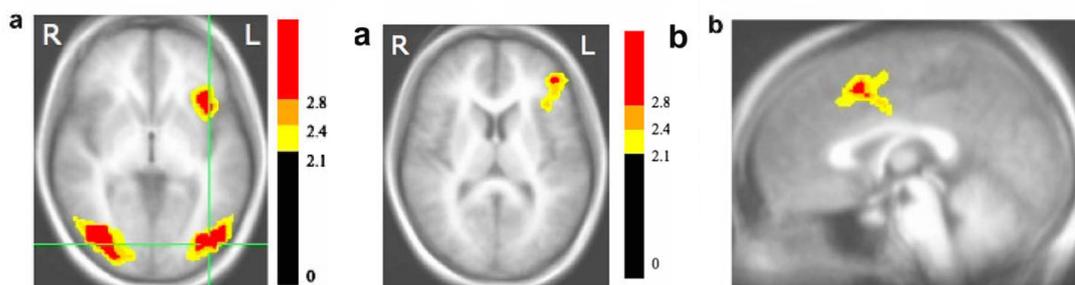


Figure 3.5 – L’expérience conduite dans [Cupchnik et al., 2009] offre des résultats d’expérimentation avec trois types de stimulus différents : i) un champ visuel quasiment vide (fond uniforme marqué de 4 croix), ii) des images quelconques et iii) des reproductions de peintures. On voit à gauche les aires activées spécifiquement en contrastant les images (ordinaires et artistiques réunies) par rapport au champ vide : on distingue, dans la partie inférieure du cerveau, l’activation des aires visuelles tertiaires (gyrus fusiforme inférieur), et dans la partie supérieure le cortex insulaire antérieur. Dans les deux images de droite, on distingue les aires cérébrales activées de façon discriminante lors de l’observation d’images “esthétiques” versus d’images ordinaires.

striatum ventral, hypothalamus et OFC lors de l’examen d’œuvres d’art, et l’on propose quelques schémas de connectivité entre ces zones.

Selon [Cupchnik et al., 2009], la coordination entre OFC et AIC (cortex insulaire antérieur) serait alors complexe, l’OFC, en charge des “récompenses” liées à la perception (en particulier visuelle), garderait la “mémoire” de ces récompenses au cours de l’expérience de façon à maintenir l’attractivité esthétique à travers l’AIC jusqu’à la prise de décision.

3.2.3 Le rôle des aires visuelles

Ces études ne passent pas sous silence l’activité des aires supérieures visuelles et l’on dispose de très nombreux travaux pour attester de leur engagement dans la vision des œuvres d’art (par exemple sur la figure 3.5 à gauche on a noté l’engagement des aires visuelles tertiaires).

Dans [Di Dio et al., 2007], un observateur était placé dans les conditions d’observation de statues dans un musée. Ces statues avaient soit les proportions d’origine, soit des proportions déformées. L’étude a mis en évidence le rôle des aires visuelles : les aires temporales ainsi que le cortex occipital latéral, l’un et l’autre connus pour traiter des images du corps humain. Dans [Cupchnik et al., 2009], une comparaison a été faite entre l’observation des peintures à contours nets de celle des peintures à contours doux. La motivation de cette étude était de mettre en évidence le possible rôle de l’attention active dans l’interprétation de l’œuvre d’art, puisque l’information de contour, importante pour la reconnaissance des formes, n’est pas disponible à la sortie des voies optiques en l’absence de bords nets. L’étude montre alors une distinction au niveau du lobule pariétal supérieur, aire en charge, entre autres, de l’orientation spatiale, indiquant, conformément à l’hypothèse des auteurs, que le cortex est bien engagé dans une tâche de complétion du message pictoral si l’image source ne porte pas toutes les informations (figure 3.6).

Dans [Vartanian and Skov, 2014] une synthèse de 15 études confirme l’enrôlement systématique des lobes antérieurs temporaux, des gyrus fusiformes et du gyrus parahippocampique qui joue un rôle dans la reconnaissance de l’environnement.

3.2.4 Le rôle de la mémoire et de la cognition

Ces mêmes études concluent souvent à l’activation d’aires impliquées dans la mémorisation. Ainsi, dans l’étude déjà citée [Di Dio et al., 2007], les auteurs identifient l’action des aires pariétales médianes et du lobe préfrontal, zones impliquées dans la fonction de mémoire. Cette expérience plaide



Figure 3.6 – Dans cette expérience issue de [Cupchuk et al., 2009], on distingue les réponses des aires spécifiquement activées lors de la vision d’images artistiques “à bords doux” (dont un exemple est au centre) versus d’images à “bord raides” (dont un exemple est à droite). C’est le lobule pariétal supérieur qui est préférentiellement actif dans le premier cas. Cette aire est en charge de fonctions de localisation spatiale dans l’interprétation de scènes.

donc aussi pour un enrôlement des fonctions cognitives lors du traitement des statues de référence. De plus, l’examen des enregistrements en présence des objets “beaux” a montré une suractivation de l’amygdale, région aux rôles multiples mais souvent engagée pour son rôle de valuation positive ou négative lors des épreuves d’apprentissage et pour son rôle dans l’expression des sentiments.

De même, la méta-étude de [Vartanian and Skov, 2014] conclut à l’activation très générale du lobe antérieur temporal dont le rôle a été identifié comme important pour la mémorisation sémantique mais aussi pour l’intégration conceptuelle lors de la reconnaissance des formes. De leurs observations, les auteurs de [Di Dio et al., 2007] déduisent que l’interprétation des œuvres d’art est d’une part “poussée par les données” (rôle des voies optiques), d’autre part “tirée par les émotions” (rôle de l’AFC et de l’AIC).

3.2.5 Le rôle de l’*embodiment*

Si les travaux précédents plaident pour la prise en compte de la mémoire, de l’expérience et *in fine* de la cognition dans l’évaluation esthétique, des travaux ont au contraire suggéré la mise en œuvre de processus très différents.

C’est le cas des travaux présentés dans [Freedberg and Gallese, 2007]. Les auteurs partent de l’hypothèse que des œuvres de contenus très différents peuvent mobiliser des aires cérébrales très différentes. Cette approche semble particulièrement adaptée à la photographie qui reproduit des scènes réelles, au risque de mêler des phénomènes que nous avons tenté de distinguer en section 1.1.

Les auteurs montrent effectivement la variété des aires concernées, souvent liées à la cognition ou à la mémorisation (comme ci-dessus), mais de façon plus originale ils trouvent que l’observation de l’œuvre peut affecter des sites qui, chez l’observateur, feraient écho à ceux qui ont pu être touchés par l’artiste soit lors de son inspiration soit lors de la réalisation de l’œuvre⁴. La réponse envoyée par ces sites va contribuer activement à l’élaboration de notre émotion, suivant une démarche connue sous le nom d’*embodiment*, d’incarnation ou de personnification.

Cette interprétation s’appuie sur la théorie encore débattue des neurones miroirs qui permettent de maintenir en soi une représentation de son propre corps et, par transition, du monde, aboutissant au développement de l’empathie classiquement exprimée lors de l’observation de ces œuvres. Ainsi les aires de notre cortex ventral prémoteur en charge des membres seraient sollicitées en observant une photographie d’athlète et interviendraient alors dans notre sensation autant, et peut-être plus, que les zones en charge des connaissances ou des souvenirs qui sont aussi invoquées (lobes temporaux, amygdale et hippocampe). Cette interprétation, testée sur des œuvres de Goya, Michelange et Le Caravage, a également été vérifiée sur des œuvres abstraites (Pollock, Fontana). Elle a été soutenue

4. Par exemple un paysage maritime peut évoquer les mêmes souvenirs de vacances à l’observateur qu’au peintre.

expérimentalement dans [Di Dio et al., 2007] et plus récemment dans [Ticini et al., 2014], et corroborerait des propositions avancées par ailleurs sur le fonctionnement de notre jugement [Damasio, 1994].

3.3 Que nous dit l'imagerie cérébrale fonctionnelle électrique ?

A la différence des approches par résonance magnétique nucléaire fonctionnelle, les techniques de mesure des potentiels électro-magnétiques des activités cérébrales permettent d'accéder à des phénomènes très rapides, mais au prix d'une localisation médiocre des sources, en particulier profondes [Renault, 2004]. Ces techniques peuvent être étendues par tomographie⁵, permettant de recalculer une information de volume.

Des études en EEG ont ainsi permis de mettre en évidence les temps de réponses des aires sensibles aux orientations, aux couleurs, ou capables de détecter la présence d'un texte écrit ou d'un visage dans une photographie même observée pendant un temps très bref (typiquement 20 ms). Ces temps s'échelonnent de 150 à 300 ms, montrant une remarquable capacité de notre circuiterie cérébrale à identifier des primitives importantes de l'image [Thorpe et al., 1996, Fize, 2004]. D'autres études ont permis de mesurer les temps nécessaires à la mise en œuvre de fonctions plus élaborées comme la catégorisation d'objets, impliquant des capacités de mémoire [Schendan and Kutas, 2007]. Ces temps s'échelonnent typiquement entre 300 et 850 ms suivant la présentation d'un objet très brièvement. Ces activations sont appelées ERP (*event-related potentials*)⁶.

Ce qui distingue la vision des œuvres d'art ne semble pas résider dans des capacités réflexes, mais dans des activations ERP potentiellement assez tardives, faisant appel à la mémoire mais aussi mettant en jeu des émotions. Ces activations apparaissent après des temps relativement longs (de l'ordre d'une seconde ou plus) et parfois très longs (de 5 à 10 s)⁷ [Hillyard and Anllo-Vento, 1998, Schupp et al., 2004]

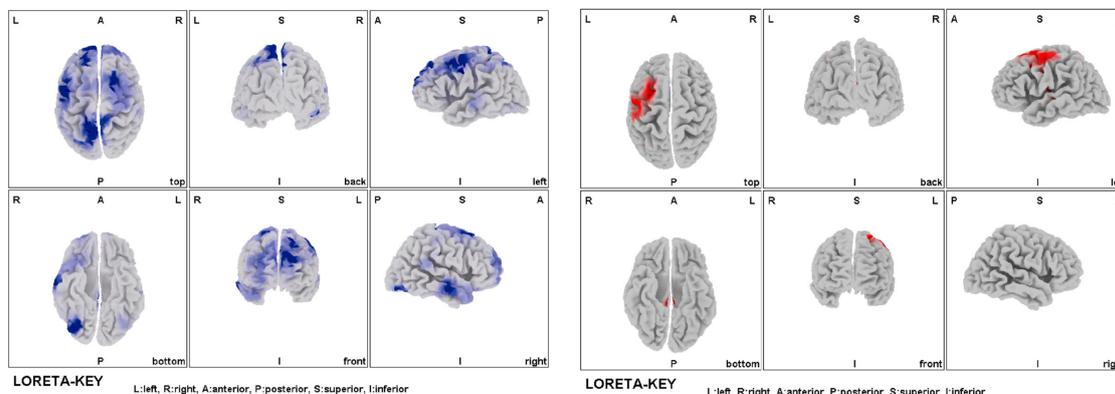


Figure 3.7 – Deux résultats d'analyse par EEG mettant en évidence le rôle de diverses aires cérébrales lors de l'observation d'œuvres d'art. À gauche, dans une expérience où l'on présente à l'observateur des œuvres d'art contemporain figuratives ou abstraites, on met en évidence les zones qui sont significativement plus actives en présence d'art figuratif : c'est en particulier le gyrus précentral gauche, le gyrus lingual temporal médian droit, le gyrus supérieur central gauche et le lobule paracentral. À droite, l'expérience consiste à exposer l'observateur à une œuvre d'art, accompagnée ou non d'un commentaire préalable. La figure présente les zones significativement plus activées en l'absence de commentaire. Ce sont surtout : le cortex insulaire, le gyrus frontal supérieur gauche et précentral gauche, (d'après [Lengger et al., 2007]).

Dans les travaux de [Lengger et al., 2007], on s'est ainsi intéressé aux différences d'activité cérébrale en présence d'une œuvre d'art abstraite ou figurative. Ces travaux, ont été menés à l'aide d'œuvres

5. Un protocole fréquemment adopté de reconstruction tomographique électromagnétique s'appelle LORETA *low resolution electromagnetic tomography analysis* [Pascual-Marqui, 1999].

6. *Event-related potentials* ou potentiel évoqué par un événement.

7. Pour cela on les appelle parfois SLP *slow cortical potentials*, potentiels corticaux lents.

contemporaines (donc relevant de critères esthétiques assez homogènes), ils étaient constitués d'un relevé des activités cérébrales et d'une enquête visant à déterminer 2 propriétés de l'expérience de l'observateur : son niveau de plaisir et sa compréhension de l'œuvre.

Concernant le plaisir ressenti, l'étude n'a pas permis d'associer de résultats significatifs : peinture abstraite ou figurative procurent des gammes de plaisir semblables qui ne se distinguent pas de façon claire dans les activations cérébrales.

Les études sur la compréhension ont conduit aux résultats présentés sur la figure 3.7 à gauche. Ils font apparaître que, après un délai de 6 à 10 secondes, des zones différentes sont actives, semblant démontrer, d'après les auteurs, que des traitements parallèles sont effectués par le cerveau pour ces deux types de stimulus. Ils constatent par ailleurs que les images figuratives provoquent de plus nombreuses associations que les images abstraites, en particulier dans le lobe frontal gauche et dans les lobes pariétal et limbique. Le lobe frontal est connu pour son rôle d'intégration et de combinaisons d'informations diverses ainsi que son implication dans le maintien de l'attention et la mémoire, implication qu'il partage avec le lobe pariétal (on retrouve certains de ces résultats dans [de Tommaso et al., 2008]).

Une autre contribution de cette étude a été de tester le rôle possible de l'apport d'information contextuelle préalable à l'observation. Pour cela certains tests ont été conduits après qu'une explication a été donnée sur le style et l'intention de l'artiste. Les résultats sont présentés sur la figure 3.7 à droite. Ils montrent que l'absence d'information conduit à un accroissement de l'activité du lobe frontal gauche, du lobe pariétal et de l'insula, ce qui est interprété par les auteurs comme une action supplémentaire entreprise par l'observateur pour rechercher une interprétation et pour catégoriser le signal reçu.

Il apparaît ainsi que toute information aisément disponible sur l'œuvre, qu'elle soit trouvée à l'intérieur de l'œuvre, ou dans une information adjointe, rend l'interprétation plus aisée. Il n'est pas mis en évidence que cette information aide à l'expérience esthétique.

Dans des travaux récents, [Van Dongen et al., 2016], les auteurs ont porté leur attention sur le rôle du contexte artistique dans l'expression de l'émotion ressentie à la vue d'une image. Pour cela ils ont présenté à un observateur des images émotionnellement fortes, issues de la collection IAPS, (images de test utilisées en psychiatrie, voir la section 1.1.3), en les prétendant ou non, issues d'une production artistique. Ils ont alors mesuré une activité différente entre œuvres prétendues d'art ou non (par exemple les images désagréables apparaissent moins désagréables et les images à contenu sexuel moins excitantes). Ils interprètent ces différences par la mobilisation différente des aires cérébrales (les zones centro-pariétales sont moins sollicitées) ; dans un contexte artistique, l'observateur serait amené à s'intéresser aux distributions de formes et couleurs (caractéristiques des œuvres d'art) plus qu'à l'interprétation du contenu.

3.4 Un schéma cognitif global pour le jugement esthétique ?

On ne peut manquer de garder une vision un peu confuse de ces travaux complexes qui sont encore très largement débattus [Cela-Conde et al., 2011, Boccia et al., 2016]. On aimerait disposer d'un schéma fonctionnel exprimant d'une part le trajet des signaux et des stimulus à travers les aires cérébrales, et d'autre part l'élaboration progressive des émotions, des jugements et des actions qui découlent de cette observation. Ces modèles n'existent pas qui couvriraient l'étendue des fonctions impliquées dans la perception esthétique.

On rencontre des modèles de perception de "bas niveau" qui décrivent de façon très précise les aires visuelles, s'appuyant sur les travaux pionniers de Hubel et Wiesel et débouchant sur des schémas fonctionnels assez élaborés et relativement bien confirmés par l'expérimentation, dont le plus célèbre est le modèle de David Marr [Marr, 1982]. Nous donnerons ici pour référence de ces travaux de modélisation de la vision précoce le modèle de Jean Petitot, d'une part pour l'extrême sophistication mathématique de son modèle, d'autre part pour les conclusions qu'il en tire sur les

ancrages innés/acquis de la représentation visuelle [Petitot, 2008].

On rencontre, par ailleurs, plusieurs modèles de vision esthétique de “haut-niveau” qui se gardent de faire le lien entre signal transmis et perception ressentie et démarrent leur modélisation par les activations neuro-physiologiques supérieures, donc, en quelque sorte, par les résultats de l’étape de perception. Elles sont discutées dans [Liu et al., 2017] et résumées ci-dessous. On est cependant surpris par la diversité des approches adoptées qui tantôt placent au centre du schéma les aires cérébrales, tantôt les fonctions cognitives, tantôt les systèmes biologiques.

3.4.1 Le modèle neuro-géométrique de J. Petitot

Jean Petitot, propose un modèle mathématique original des voies visuelles primaires (en particulier de l’aire V1) qu’il appelle modèle neuro-géométrique [Petitot, 2008]. Il s’intéresse particulièrement à l’architecture des connexions neuronales au sein des colonnes de l’aire V1, colonnes qui conduisent à un filtrage du signal visuel en orientation d’une part et en fréquences spatiales d’autre part. Il fait le parallèle de ces opérations avec la fibration du plan rétinien par les droites projectives des orientations du plan telle qu’elle pourrait être définie en géométrie différentielle. Développant cette analogie, il montre que l’invariance par translation conduit à introduire une géométrie sous-riemannienne qui permet d’expliquer des caractéristiques spécifiques de la vision comme les groupements perceptifs décrits par la théorie de la Gestalt ou la fermeture à distance des contours illustrée par les illusions de Kanizsa.

Ce modèle mathématiques trouve de bonnes vérifications tant dans la confrontation aux architectures neuronales que dans l’interprétation de nombreuses expériences vécues. L’auteur montre également qu’il apporte une explication à des problèmes épistémologiques très profonds de la perception visuelle comme la notion de l’intuition transcendentale qui permet de résoudre de délicats problèmes de conflit entre inné et acquis, problèmes qui sont au cœur de l’esthétique kantienne, mais qui irriguent également toute la philosophie phénoménologique de Husserl à Merleau-Ponty, en passant par Bergson. Ces travaux de neuro-géométrie vont donc en partie à contre-courant de la plupart des approches de neurobiologies qui ne reconnaissent pas l’importance de l’architecture cérébrale dans notre représentation “précablée” de l’espace et confient à l’apprentissage le rôle premier de construction de nos représentations.

3.4.2 Le modèle d’émotion esthétique de A. Chatterjee

Le modèle de A. Chatterjee [Chatterjee, 2004] est l’un des plus anciens qui s’attache à représenter la perception de haut-niveau et son impact sur le ressenti de l’observateur. Il subdivise les tâches perceptives et cognitives engagées dans une observation visuelle en quelques grandes catégories qu’il rattache aux structures cérébrales et il tente de suivre les transmissions temporelles entre ces tâches (figure 3.8). L’évaluation esthétique prend deux formes : une réponse émotionnelle d’une part, une décision anticipant l’action d’autre part. Suivant ce schéma :

- les aires visuelles, zones occipitales, zones latérales inférieures, cortex insulaire, lobule pariétal supérieur, ... sont bien sûr actives par la tâche précoce de vision, agissant directement sur le flot optique mais aussi en extrayant les trames sémantiques selon des aires spécifiques : formes, couleurs, mouvements, visages ;
- le cortex orbito-frontal en charge de l’évaluation, régulateur des “récompenses” et du plaisir, est également une pièce maîtresse du contrôle de notre décision esthétique⁸ ;
- le cortex insulaire, qui contrôle nos émotions, est lui aussi une pièce importante et quasiment toujours impliquée lors de l’observation d’une œuvre d’art ; il apporte une contribution nécessaire à l’évaluation. Cette contribution est-elle antérieure à la “récompense” ou en est-elle la conséquence ? Cela ne semble pas tranché ;

8. Rappelons que la définition la plus fréquemment utilisée aujourd’hui pour la Beauté est “la caractéristique d’un objet dont la vision procure une expérience perceptuelle de plaisir” (voir le Chapitre 1).

- les zones engagées dans les opérations de cognition (amygdale) et de mémoire (aires pariétales médianes, lobe préfrontal) sont également sollicitées. Elles fournissent les contributions de haut-niveau qui seront invoquées pour “justifier” l’évaluation, en particulier chaque fois que l’on sollicitera de notre conscience un rapport explicatif du jugement esthétique. Ces justifications pourront être des références à notre vécu ou les déductions d’un raisonnement ;
- des aires en charge du contrôle pré moteur sont convoquées de façon spécifique dans des situations de forte empathie et d’*embodiment* (cortex ventral pré moteur, lobes temporaux, hippocampe). Elles interviennent aussi en réflexe dans le parcours visuel de l’image, ou sur la sollicitation des conclusions des aires de cognition, à la recherche d’indices de confirmation ou d’explication.

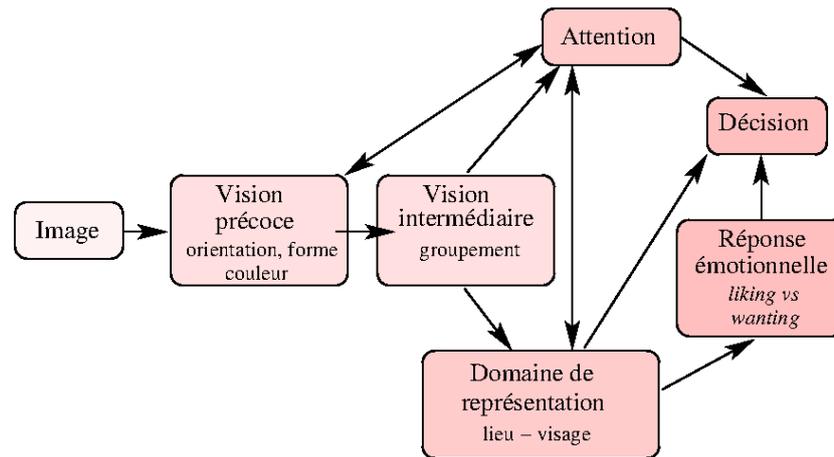


Figure 3.8 – *Modèle fonctionnel de A. Chatterjee pour la vision avec jugement esthétique. Deux fonctions différentes concernent l’une l’attention, l’autre la représentation. Elles sont en interaction univoque. L’attention réagit sur l’étage de vision précoce en dirigeant l’exploration. Elle est en partie influencée par la réponse émotionnelle qu’elle contribue à élaborer et elle agit sur notre décision en particulier à travers les réactions réflexe. La représentation est en relation avec nos modèles du monde et intervient également dans la prise de décision (la justification du jugement esthétique) et la réponse émotionnelle (adapté de [Chatterjee, 2004]).*

3.4.3 Le modèle de S. Brown *et al.*

Ce modèle a été présenté dans [Brown *et al.*, 2011]. Il est schématisé sur la figure 3.9 et reflète les observations qui ont été rapportées dans la section 3.2. L’évaluation esthétique, dans ce modèle, résulte de la conjonction de deux traitements : l’un selon des voies intéroceptives (exprimant l’état des divers indicateurs de nos fonctions internes), l’autre extéroceptif reflétant la perception, visuelle dans le cas d’une expérience esthétique sur une photo.

La voie intéroceptive provient des zones sous-corticales : essentiellement le thalamus, l’hypothalamus et le cortex insulaire dorsal postérieur. Les stimulus ainsi élaborés agissent sur le réseau neuronal intéroceptif, et particulièrement sur les cortex insulaires antérieurs droit et gauche. Toutes ces zones sont connues pour intervenir en cas d’émotion, d’empathie et en règle générale pour exprimer le goût (tant positif que négatif).

La voie extéroceptive est issue du cortex orbito-frontal (OFC) qui jouerait le rôle de porte d’entrée pour l’évaluation des récompenses pour les divers types de modalité (vision, ouïe, goût, ...). Dans l’OFC, il semble que les diverses modalités disposeraient d’aires adjacentes plutôt que communes. Ainsi, comme la voie intéroceptive, elle serait également impliquée dans l’élaboration de la récompense attachée au message des canaux visuels, mais la voie à travers l’OFC ferait plus appel à la mémoire et au raisonnement. Par ailleurs, elle semblerait maintenir l’évaluation faite dans le temps de façon à la faire intervenir dans les évaluations futures.

Dans le modèle de Brown, l’OFC est également en interaction avec la partie la plus antérieure du cortex cingulaire (l’ACC) qui semble impliquée dans le monitoring de la sillance visuelle.

La connectivité entre cortex insulaire antérieur et OFC conduirait alors à une “émotion homéostatique” (c’est à dire biologiquement stabilisée), reflétant notre jugement esthétique.

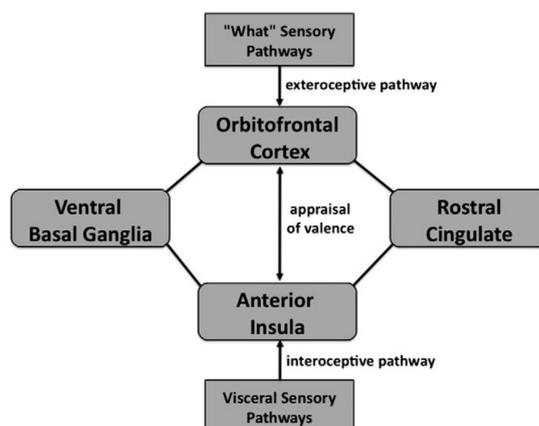


Figure 3.9 – Modèle fonctionnel de connectivité de S. Brown et al. pour la vision avec jugement esthétique. Dans ce modèle, l’évaluation est faite par la comparaison récurrente d’informations externes (extéroceptives) passant par le cortex orbito-frontal (OFC) et d’informations interoceptives passant par les régions sous-corticales (insula antérieure). On voit également l’influence du cortex cingulaire (rostral cortex) en charge des émotions et du ganglion ventral de la base en charge de la réponse hédonique. (d’après [Brown et al., 2011]).

3.4.4 Le modèle de H. Leder

A peu près en même temps que A. Chatterjee, H. Leder propose un autre schéma. Il se place sur un plan très différent pour proposer une modélisation des actions de perception et d’évaluation esthétique pour laquelle il propose une catégorisation très fine des opérations conduites par l’opérateur, en choisissant une modélisation psychologique et non physiologique [Leder et al., 2004]. De plus, il s’intéresse à l’évaluation esthétique des œuvres d’art, tâche qui fait probablement beaucoup plus appel à des fonctions cognitives et sémantiques (reconnaissance du style, reconnaissance de l’artiste) que l’évaluation de photographies. Cette modélisation nous offre néanmoins un cadre intéressant pour notre propos. Selon H. Leder *et al.*, le mécanisme se diviserait en cinq étapes (figure 3.10) : les deux premières concernent essentiellement les voies visuelles et sont surtout réflexes (pour nous donc nous pourrions les qualifier d’“objectivistes”), les trois suivantes sont en charge de la reconnaissance du contenu, de l’interprétation globale et de l’évaluation ; celles-ci sont cognitives et pourraient être qualifiées de “subjectivistes” : elles utilisent notre expertise du domaine, nos centres d’intérêt, nos goûts, mais aussi les éléments de contexte qui accompagnent l’expérience et qui sont fort importants pour estimer une œuvre d’art (l’ambiance d’un musée ou l’effet “signature” déjà évoqué). Les grandes lignes de la figure 3.10 pourraient certainement inspirer un schéma à peine différent qui serait adapté à la photographie.

Comme dans le travail de Chatterjee, H. Leder propose pour sorties de ce schéma d’une part une émotion, d’autre part une évaluation esthétique, mais ces deux résultats sont élaborés simultanément mais selon des circuits séparés tout au long d’une évaluation affective qui convoque progressivement des éléments d’expertise de l’œuvre d’art, expertise résultat d’une expérience antérieure, de goûts personnels et d’intérêt. Enfin H. Leder fait intervenir un contexte propre à l’observation dans un musée qui agit en dehors des circuits perceptifs sur le seul jugement esthétique et non sur l’émotion.

Le modèle de Leder a été complété et un peu aménagé dans [Leder and Nadal, 2014], en particulier pour mieux prendre en compte le rôle de l’émotion dans l’évaluation esthétique.

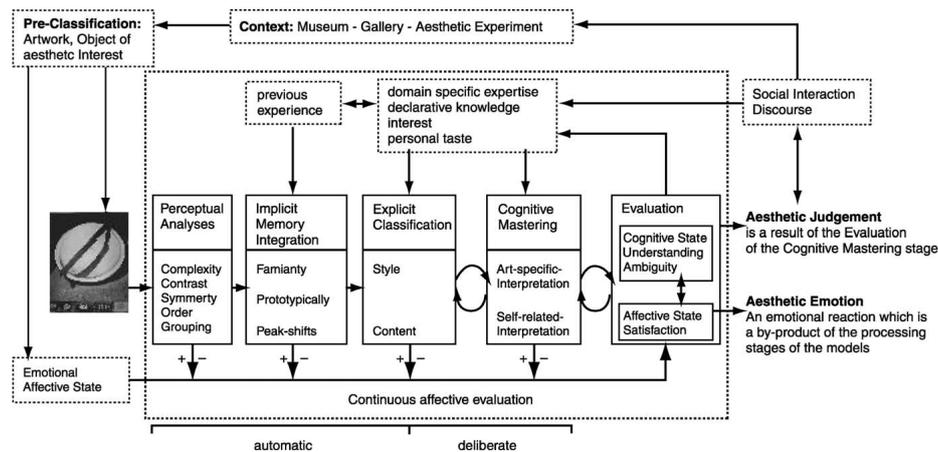


Figure 3.10 – Un modèle de la perception esthétique de H. Leder (d’après [Leder et al., 2004]). On distingue dans ce modèle deux sorties distinctes : l’une correspond à l’évaluation esthétique, l’autre à l’émotion esthétique. Selon les auteurs, les deux opérations sont élaborées en parallèle, suivant des circuits distincts et ne sont donc pas réductibles l’une à l’autre. Ils considèrent cependant que le plaisir esthétique (au cœur de l’émotion) est d’autant plus important que la compréhension de l’œuvre (un des points majeurs de l’évaluation) est plus forte. L’étude est conduite sur des œuvres de l’art contemporain et il n’est pas évident que les mêmes arguments s’appliquent aux photographies.

3.4.5 Le modèle de C. Redies

C. Redies a proposé un modèle de jugement esthétique qui mêle les aspects perceptifs et cognitifs au sein d’un même flot de signaux, mais les séparant selon deux mondes (voir la figure 3.11) : l’un externe (le signal d’image d’une part et le contexte de l’observation), l’autre interne rassemblant toutes les fonctionnalités biologiques, qu’elles soient perceptives ou cognitives [Redies, 2015]. Dans ce monde interne, le flux traité par les voies visuelles primaires est dirigé vers les zones de la perception et fournit les éléments de décision de la beauté ainsi qu’une contribution à l’établissement de la réponse émotionnelle (partie haute du schéma). Envoyé vers les zones de codage cognitif et de mémoire, il élabore une réponse contextuelle par filtrage personnel qui contribue également à la réponse émotionnelle (partie basse du schéma). Ce filtrage est le résultat, à long terme, de l’exposition à l’environnement culturel de l’observateur (apprentissage, expérience), à la différence du monde de l’information externe qui est sous l’influence du contexte immédiat de l’observation (ambiance, humeur, ...).

L’émotion partage les deux flots de données, perceptuel d’une part, cognitif d’autre part, ce qui est une différence avec le modèle précédent qui ne reconnaissait pas d’influence des étapes cognitives sur l’émotion. Enfin l’expérience esthétique est ici la boîte de sortie du système qui n’envisage pas explicitement le jugement esthétique comme produit de l’observation.

3.4.6 Le modèle d’émotions de S. Koelsch et al.

Ce modèle ne se limite pas au domaine de l’art. Il souhaite couvrir tout le champ de l’émotion [Koelsch et al., 2015], et particulièrement de l’émotion humaine avec sa très forte et très complexe composante sociale. Il repose sur une séparation fonctionnelle du cerveau en quatre noyaux (d’où le nom de quatuor (*quartet*) donné à ce modèle), qui régulent les activités affectives et qui se sont progressivement constitués au cours de l’évolution biologique : les aires orbito-frontales, l’hippocampe, le diencephale et le tronc cérébral (voir figure 3.12). Les auteurs détaillent très précisément les aires cérébrales concernées par chacun de ces quatre noyaux, leur rôle et leur action ainsi que leur contribution à l’élaboration de l’émotion. Ils distribuent sur ces aires de très nombreux concepts psycholo-

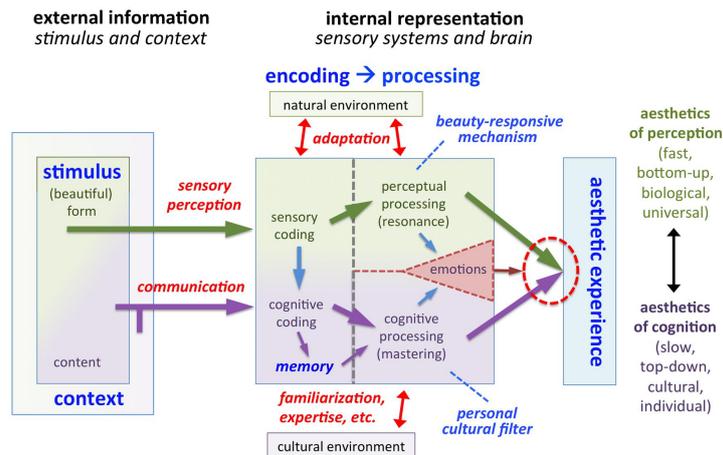


Figure 3.11 – Modèle fonctionnel de E. Redies. Ce schéma distingue un monde externe d’un monde interne. Le monde externe est constitué de l’objet observé, déjà doté de ses attributs perceptifs (dont la beauté) et du contexte de son observation (dans un musée, par exemple). Le monde interne sépare le traitement du signal (codage, traitement) dans la partie supérieure, du traitement cognitif dans la partie basse. L’émotion procède de ces deux traitements. Le résultat est une “expérience esthétique” qui comporte à la fois les dimensions émotionnelles et les dimensions cognitives (justification, explication). (d’après [Redies, 2015]).

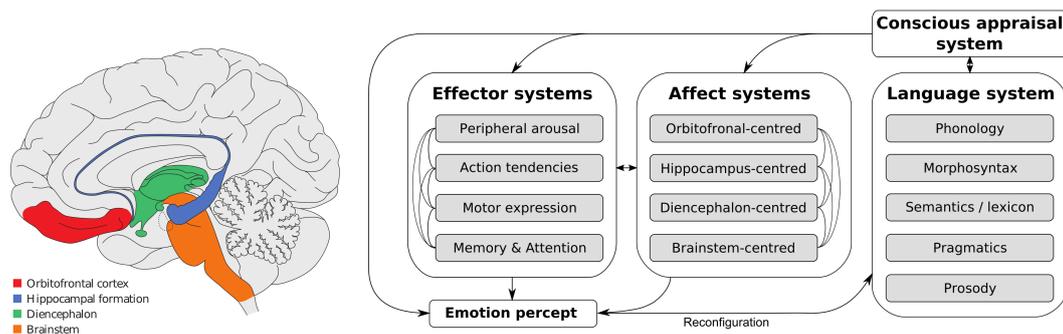


Figure 3.12 – Modèle fonctionnel (dit “du quatuor”) de S. Koelsch et al. Le système émotionnel est constitué de quatre centres actifs (appelés noyaux (core)) : l’un centré autour des aires orbito-frontales, un autre autour de l’hippocampe, le troisième autour du diencephale, le dernier autour du tronc cérébral (à gauche). Le rôle du système linguistique est également souligné dans ce modèle. (d’après [Koelsch et al., 2015]).

giques qui se trouvent ainsi délégués à des noyaux différents : l’apprentissage et la mémorisation des émotions, la satiété émotionnelle, la complexité cognitive, les émotions subjectives, le degré de conscience, etc. Ils montrent enfin comment cette perception émotionnelle nourrit l’expression et le langage et comment ceux-ci, en retour agissent comme régulateurs de l’émotion.

3.4.7 Le modèle d’émotions de Li-Hsiang Hsu d’après A. Damasio

Le modèle proposé dans [Hsu, 2009] est inspiré des théories d’António Damásio [Damasio, 1994]. Il se place d’un point de vue de la philosophie de l’esprit et décompose le traitement de l’œuvre d’art en deux flux complémentaires (modèle du double objet intentionnel, voir figure 3.13) et non quatre comme dans le modèle précédent. Il s’appuie sur un schéma fonctionnel du cerveau qui distingue la voie limbique (qui transmet les émotions), de la voie du néocortex (qui conduit à la pensée réflexive). La distinction limbique/néocortex est discutée (et critiquée) dans [LeDoux, 2000].

Dans l’approche de Hsu, un flux proprioceptif vient créer/modifier l’état émotionnel de l’observateur. Il agit de façon réflexive avec des temps de réponse très courts. Le second flux construit notre

réponse cognitive au stimulus en nourrissant les zones périphériques du néocortex. Suivant les interprétations de Damasio, les deux flux ne se combinent pas mais se juxtaposent : l'élaboration d'une pensée formulée selon la voie cognitive se superpose à l'évolution de notre perception. L'image mentale que nous formons par interprétation de l'image, association à notre contenu mémoriel, déduction de notre raisonnement, ... est sous l'influence de notre état corporel, par l'intermédiaire de transmetteurs neuro-chimiques issus de la voie limbique.

Li-Hsiang Hsu exprime ainsi la construction de notre ressenti face à une telle œuvre : *“les émotions basées sur l'image du corps ne sont pas le parasite de la raison, mais les conditions de la conscience. Elles jouent un rôle actif dans la cognition, dans la mémorisation et dans la prise de décision. En préparant l'organisme à l'action et en réduisant la complexité des problèmes rencontrés, elles constituent le soubassement permettant à la raison de fonctionner.”* ([Hsu, 2009] p. 112).

Ainsi, suivant Damasio, l'action d'observer une œuvre d'art comporte deux objectifs distincts : d'une part la conscience d'un objet intentionnel extéroceptif, visible, observable par l'action des sens, d'autre part la conscience d'un objet intentionnel intéroceptif, implicite, interne à l'observateur, observable sous la forme de sentiment de soi. Ce second objet reflète l'observateur tel que le “réveille” l'objet observé ; c'est une expression de l'observateur “éclairé” par l'objet. Ce second objet est un apport original de Damasio⁹ à la philosophie du Beau. Ce second objet serait demeuré négligé des travaux antérieurs qui se sont concentrés essentiellement sur les relations de l'objet à la conscience extéroceptive.

Ainsi, on lit dans Damasio : *“Les images sensibles de ce que vous percevez extérieurement, et les images correspondantes que vous vous rappelez, occupent la majeure partie du champ de votre esprit, mais pas sa totalité. Outre ces images, il y a aussi cette autre présence qui vous signifie vous, en tant qu'observateur des choses en images, propriétaire des choses en images, acteur potentiel des choses en image. Il y a une présence de vous, dans une relation particulière à un certain objet.”*, [Damasio, 1999], rapporté dans [Hsu, 2009].

3.4.8 D'autres modèles

D'autres modèles existent, plus ou moins formulés, plus ou moins élaborés. Nous verrons par exemple, en section 4.5.2 un modèle perceptif de textures, le modèle de S. Thumfart *et al.*, qui a permis de rattacher, par des mesures de psychophysologie, la perception des textures à trois niveaux : un niveau affectif, un niveau de jugement et un niveau émotionnel. Un tel modèle sera utilisé pour aider à l'évaluation esthétique.

On trouve dans [Vartanian and Skov, 2014, Kirsch et al., 2016] des schémas plus complets qui tentent de figurer le réseau des activations lors d'une exposition à une œuvre d'art, mettant en relation les trois grandes fonctionnalités : cognition, émotion, contrôle sensori-moteur. Ils sont encore très spéculatifs.

Par ailleurs, J.P. Changeux propose un schéma assez complet de la progression des signaux depuis l'entrée dans le système visuel jusqu'à l'accès à la conscience *“au-delà des cortex inféro-temporal et pariétal”* ([Changeux, 2016], pages 166-173). Ce schéma représente bien les aiguillages qui orientent les diverses composantes du signal : couleur, contours, fond/forme, visages ... vers des zones spécialisées, puis la mobilisation des aires en charge de la mémoire, du signifié, des émotions ... , et couvre bien les fonctions d'observation consciente. Il est moins précis sur l'accès à *“l'espace de travail neuronal conscient qui s'appuie sur un réseau de neurones, dont les axones à longue distance réunissent les régions préfrontale, pariéto-temporale et cingulaire [...] l'ignition fait intervenir une rétroaction descendante en boucle [...] Cette ignition serait la signature objective de l'efficacité esthétique de l'œuvre d'art”* ([Changeux, 2016], pages 169). L'objet même que nous nous fixons en début de ce cha-

9. Sous certains aspects, l'approche de Damasio peut être interprétée en termes d'*affordance*, c'est-à-dire d'une perception fonctionnelle immédiate et non médiatée, proposée dans la “théorie écologique de la perception” de J.J. Gibson [Gibson, 1986]. Gibson faisait également référence à une double voie de perception, l'une proprioceptive, l'autre extéroceptive.

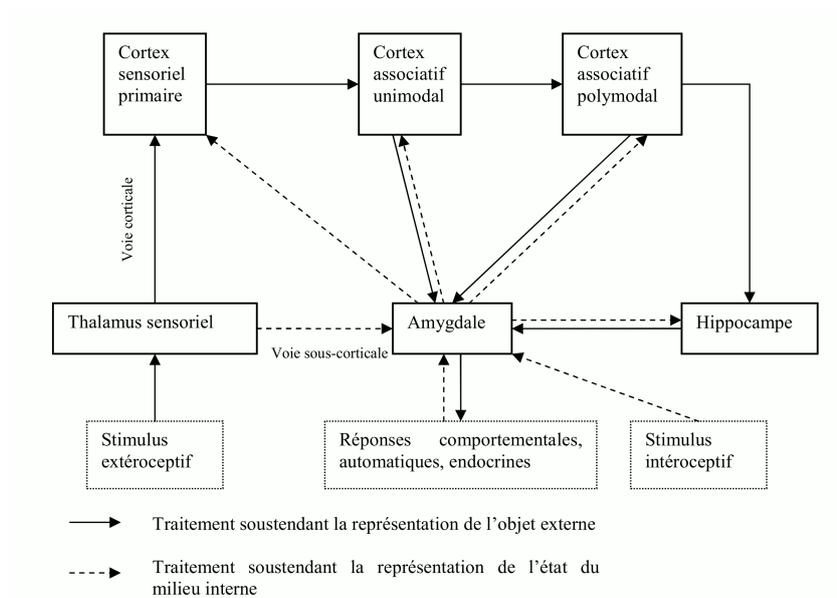


Figure 3.13 – *Modèle de double objet intentionnel de Li-Hsiang Hsu. C'est le stimulus extéroceptif (par l'intermédiaire des voies visuelles) qui constitue l'entrée du signal qui distribuera les informations vers une représentation du monde extérieur d'une part (la perception physique de l'objet) et vers une représentation proprioceptive (modification de l'état de l'observateur) qui exprimera les émotions ressenties à la vue de l'objet. La voie sous-corticale, vers le système limbique (hippocampe, amygdale, gyrus singulaire, hypothalamus), est courte, tandis que l'autre, vers le néo-cortex (vers les zones occipitales des aires visuelles d'une part, vers les zones préfrontales et orbitofrontales d'autre part) est longue. C'est la voie cognitive qui serait associée à une évaluation. (d'après [Hsu, 2009]).*

pitre est bien effectivement isolé là, mais dans une forme beaucoup trop conceptuelle et hypothétique pour convaincre définitivement. Il est indispensable que l'on progresse dans nos capacités d'examen des activités cérébrales pour confirmer ces réflexions. Cela ne pourra se faire sans des avancées importantes dans la sensibilité et la rapidité de détection des instruments de mesure (EEG, MEG) et l'on peut imaginer que l'avènement d'outils plus performants ouvrira alors de nouvelles interrogations qui pourront conduire à d'autres interprétations.

3.5 L'approche neuro-esthétique critiquée

L'importance considérable des travaux consacrés aux bases neurales de l'esthétique, et plus spécifiquement à l'approche neuro-esthétique, ainsi que le soutien considérable qu'elle a reçu dans le monde scientifique [Changeux, 2008, Changeux, 2016, Boulez et al., 2014, Le Bihan, 2012] ne doit pas cacher des critiques importantes qui lui sont adressées. Ces critiques s'adressent en général à toutes les approches de l'imagerie cérébrale [Uttal, 2002, Dupont, 2015] et visent à la fois son objet et ses méthodes. Pour la neuro-esthétique, on les trouve bien rassemblées dans [Vidal, 2011, Vidal, 2012], ainsi que des éléments de réponse dans [Pearce et al., 2016].

Critiques sur les méthodes du neuro-esthétisme : sur la méthode, on reproche aux neuro-esthètes, non sans pertinence, de mettre en évidence des "corrélations" (telle zone est active lorsque tel stimulus apparaît) et d'en déduire souvent hâtivement des modèles de fonctionnement sous-tendant des causalités. Or les fonctions cérébrales sont distribuées dans le cortex et, simultanément, diverses aires sont sollicitées à l'occasion de divers traitements. L'analyse a posteriori de cartes d'activité cérébrale, ignorant la diversité des mécanismes sous-jacents et cherchant l'explication la plus concise

entre objet et appréciation esthétique, est tentée de donner des raccourcis d'interprétation comme : « *la beauté est une qualité des objets qui est corrélée avec l'activité de cortex préfrontal médian par l'intermédiaire des sens* » [Ishizu and Zeki, 2011].

Il est clair que l'évolution des techniques d'observation, et en particulier l'émergence des méthodes qui permettent de mettre en évidence les échanges entre les aires cérébrales, sont à même de repousser cette critique à terme proche.

Une autre critique reproche à la neuro-esthétique de travailler dans l'urgence. De nombreuses études définiraient mal l'objectif de leurs expériences et confondraient ou mélangeraient des notions différentes : beau, art, intérêt, innovation ou singularité ... Les expériences ignorent souvent des dimensions déterminantes : culture, éducation, valorisation sociale, etc. Ces dimensions sont depuis de nombreuses années les composantes identifiées par les sciences sociales, l'écologie, la psychologie. Si les neurologues ne les prennent pas assez en compte, ce n'est point tant qu'ils les ignorent, mais soit ils les considèrent - sans démonstration - comme intégrées à leur démarche, soit ils les jugent trop complexes pour être convenablement traitées et préfèrent alors les exclure totalement du champ de leurs travaux.

Critiques de l'objet du neuro-esthétisme : les critiques concernant l'objet des travaux de neuro-esthétique sont plus fondamentales.

Des critiques très radicales sont exprimées dans [Brown and Dissanayake, 2009], prenant en compte l'évolution sociale liée au contexte multi-culturel contemporain. On oppose alors l'interprétation du terme *esthétique*, issue du siècle des Lumières et attachée à la réponse émotionnelle résultat de la contemplation des œuvres d'art (celle aussi que poursuivent les neuro-esthètes) et le sens historique associé à tout système de valeur lié à l'appréciation de la beauté, sens que l'on retrouve en particulier dans l'œuvre de Darwin où il s'applique à des gents animales.

Plus fondamentalement, on ne peut écarter les critiques portant sur les prémisses philosophiques des démarches de neuro-esthétique qui se fondent sur les bases que les subjectivistes expérimentalistes du XVIII^e siècle avaient posées : "il n'est pas d'inné, donc toute conscience issue de la perception est le résultat d'une activation biologique produite par l'observateur". A cela, Henri Bergson répond : "l'intelligence est inapte à nous donner la réalité des choses car elle les mortifie et ne nous livre qu'un schéma vide de ces dernières", [Bergson, 2009] et Maurice Merleau-Ponty surenchérit : "*Il est inévitable que dans son effort général d'objectivation la science en vienne à se représenter l'organisme humain comme un système physique en présence de stimuli définis eux-mêmes par leurs propriétés physico-chimiques, cherche à reconstruire sur cette base la perception effective et à fermer le cycle de la connaissance scientifique en découvrant les lois selon lesquelles se produit la connaissance elle-même, en fondant une science objective de la subjectivité. Mais il est inévitable aussi que cette tentative échoue.*", [Merleau-Ponty, 1945]. En effet, comme l'expriment Philippe Huneman et Estelle Kulich "Elle (la philosophie) ne sera pas celle d'une raison universelle, mais mettra en évidence un "champ transcendantal", c'est à dire un espace où il y a une interdépendance entre les objets et le sujet méditant, rendant ainsi caduque la séparation sujet/objet (laquelle ne tient que par l'indépendance et le caractère abstrait du sujet en philosophie)", [Huneman and Kulich, 1997].

Chapitre 4

Que disent les manuels de photo ? Qu'en pensent les psychologues expérimentaux ?

Abandonnons le domaine de la neuro-biologie et tournons nous maintenant vers les très abondantes bibliothèques des photographes (ou des artistes peintres), pour voir quelles recommandations ils nous font pour composer une belle photo.

Ce qui frappe tout d'abord, c'est le faible nombre de conseils "opérationnels" qui sont fournis. Les ouvrages se répartissent généralement en deux familles : ceux qui essaient de faire partager une expérience esthétique et ceux qui transmettent une leçon technique.

Les premiers s'intéressent généralement à l'objet (ou au sujet) de la photographie : ils abordent des questions de forme, de lumière, de composition, de couleur, ... à la façon des traités d'esthétique en peinture. Comme eux, ils ont pour but d'éduquer le photographe, c'est-à-dire de le doter de la sensibilité qui l'amènera à la production d'une belle photo, mais celle-ci est définie en termes très génériques qui font référence à un style et bien rarement à des règles. Il est très difficile d'utiliser de tels ouvrages pour répondre à la question : "Cette photo est-elle belle ?" ; tout au plus pourrions nous répondre à la question : "Cette image répond-elle aux critères de ... ?"

Les seconds s'attachent généralement à lier les réglages de l'appareil aux effets que l'on désire obtenir : le choix d'une émulsion ou d'une sensibilité, d'un diaphragme, d'un temps d'exposition, d'une lentille ou d'un éclairage, peut-être de filtres, etc.

Si les manuels de photographie donnent peu de règles utilisables pour décider de la beauté d'une image, les traiteurs d'images, et en particulier ceux qui s'intéressent à l'image animée, ont élaboré des méthodes permettant de noter la qualité d'une image ou d'une séquence d'images (voir chapitre 6 de [Maitre, 2015]). La qualité d'image est alors un attribut technique qui regroupe des aspects particuliers du signal :

1. le contraste ou la dynamique et la diversité des couleurs,
2. la mise au point en particulier sur les contours et la finesse des détails dans les textures,
3. l'absence de bruit sur les plages uniformes et la subtilité des modelés sur les zones faiblement variables,
4. éventuellement les distorsions géométriques et colorimétriques.

Ces qualités techniques permettent de préférer un mode de transmission, de codage ou d'affichage à un autre. Elles ont conduit à des algorithmes opérationnels, par exemple : [Mittal et al., 2013, Moorthy and Bovik, 2011, Charrier et al., 2012, Chandler, 2013]. Mais en n'accordant aucun rôle aux qualités esthétiques de l'image, elles ne répondent pas à nos attentes (voir la figure 4.1). Notons

cependant que la proximité des deux notions : beauté d'une part, qualité d'autre part, a conduit certains auteurs à traiter les deux problèmes avec les mêmes outils [Talebi and Milanfar, 2017].

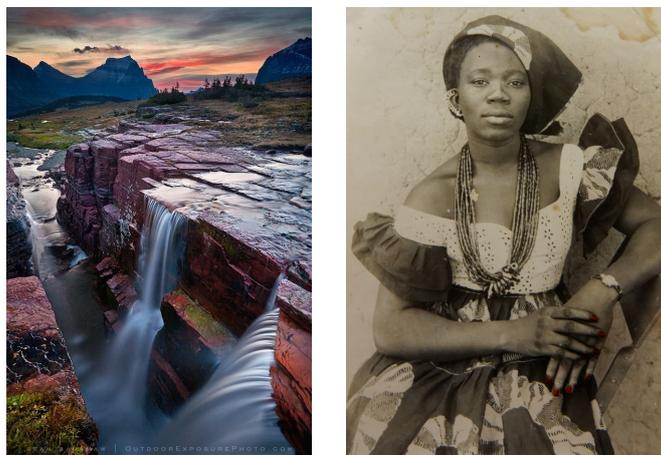


Figure 4.1 – *Qualité de la photographie et Beauté de la photographie. L'image de gauche (Triple falls in Glacier National Park, Montana de Sean Bagshaw) présente tous les attributs de qualité : résolution, mise au point, palette chromatique, nuances, textures. Elle a été distinguée dans la communauté très technophile du site Pinterest pour sa beauté ; il n'est pas sûr pourtant qu'elle rassemble les suffrages de tous les experts en ce qui concerne son esthétique. A droite, au contraire, cette photo (Sans titre de Seydou Keïta) a souffert d'un environnement très médiocre tant à la prise de vue qu'à l'archivage qui ont considérablement nui à sa qualité technique. Elle jouit cependant d'une indéniable reconnaissance quant à son esthétique.*

Notre travail dans ce chapitre consiste à collecter les critères qui sont usuellement proposés pour réaliser de belles photos, à les discuter d'un point de vue des théories disponibles et à voir comment ils ont été transposés en pratique.

Nous nous arrêterons pour commencer sur le vocabulaire utilisé pour caractériser les qualités esthétiques et constaterons que les divers termes employés se distinguent considérablement par leur sémantique. Cette constatation faite au milieu du siècle précédent semble cependant avoir eu peu d'impact sur les approches modernes. Avec les nouvelles approches exploitant les grandes masses d'information, il serait peut-être souhaitable d'en exploiter les mérites.

Nous distinguerons ensuite successivement les critères objectifs (le format, la composition, les contrastes, les propriétés spectrales, la palette chromatique ...) qui se mesurent directement dans l'image, puis ceux, que la terminologie désigne par *esthétiques*, que l'on pourrait également dénommer subjectifs parce qu'ils font intensivement appel à des observateurs pour être vérifiés. Pour chaque rubrique, nous nous efforçons de confronter trois aspects : tout d'abord les règles issues du domaine de la photographie ou de l'esthétique picturale, puis les mécanismes qui peuvent être mis en jeu pour les expliquer et que justifieraient des études de psycho-physiologie expérimentale et enfin des vérifications statistiques faites à posteriori sur de grands corpus de documents. Malheureusement, dans de nombreuses situations il n'est pas possible de recouper tous ces avis.

4.1 Le vocabulaire de l'évaluation esthétique

De longs débats [Sibley, 1959, Elton, 1954, Kivy, 1968, Iseminger, 1981] ont pris place au sein de ce qu'il est convenu d'appeler l'école d'esthétique analytique afin de caractériser le langage utilisé dans la critique esthétique. Nous avons discuté de ce point en section 2.1.

Le point de départ est la constatation que si certains qualitatifs relèvent directement de la perception par nos sens : *rouge, contrasté, sinueux*, d'autres relèvent de façon complexe de l'observation de l'œuvre : *gracieux, dynamique, triste* ... Ces termes sont dénotés *esthétiques* par ces chercheurs

qui se sont attachés à cerner la façon dont ils sont élaborés par l'observateur puisque des mots de ce type constituent souvent la trame de l'explication fournie par l'expert pour justifier son jugement.

Pour certains, il apparaîtrait que ces qualificatifs “esthétiques” dérivent directement des caractéristiques “non-esthétiques”, mais ils en dérivent de façon très complexes. A la différence d'un terme comme *rectangulaire* qui se déduit logiquement de quelques règles simples (“des angles droits”, des “côtés parallèles”), à la différence même de qualificatifs plus complexes, comme par exemple *intelligent* qui peut se concevoir de la vérification de l'une quelconque d'un grand nombre de conditions suffisantes (“avoir de hauts diplômes”, “bien jouer aux échecs”, “parler plusieurs langues”, “répondre brillamment” . . .), il ne semble pas qu'aucune batterie de tests puisse permettre de décider qu'une peinture est *banale* ou *romantique*¹. Le passage de ces propriétés non-esthétiques, partageables par tous, à ces propriétés esthétiques qui en découleraient, fait tout l'enjeu du débat philosophique en référence et la controverse est rude autour des chemins pour y parvenir. Pour [Cometti et al., 2000], les propriétés esthétiques “surviennent” (au sens de Levinson [Levinson, 1980]) sur les propriétés non-esthétiques². Pour d'autres philosophes, les éléments “esthétiques” font l'objet d'un traitement séparé de la vision ou de l'audition, par un sens qui serait distinct : le “goût”³, posant de nouveaux problèmes [Kivy, 1968].

D'autres questions philosophiques se posent : les qualités esthétiques sont-elles évaluatives ? Oui, nous le souhaitons pour reconnaître qu'une photo est plus ou moins “belle”, mais que veut dire évaluatif pour une photo “triste” ? De nombreuses qualités sont métaphoriques (“dramatique”, “romantique” . . .) comment conçoit-on le franchissement du fossé sémantique dans ce cas ? Que faut-il ajouter à la photographie pour qu'elle devienne “romantique” ? Y a t'il dans l'agencement des pixels des configurations qui seraient partagées avec la succession des notes d'un Nocturne de Frédéric Chopin ou de vers de Marceline Desbordes-Valmore ? Y a t'il “propriété esthétique” sans la culture de l'observateur, sans son éducation, sans son tempérament ? Et finalement y a t'il “propriété esthétique” sans l'observateur ? Et, comme le dit pertinemment E. Gilson⁴ : que deviennent les propriétés esthétiques d'un tableau lorsque se referme la porte sur le dernier visiteur ?

Nous entrons de plain pied dans le débat que nous allons traiter en des termes à peine nouveaux. Remplaçons “propriétés non-esthétiques” par “mesure par ordinateur” et nous avons le même questionnement que nous simplifions tout de même un petit peu car nous nous contentons de rechercher une “valeur esthétique” et non toute la richesse des “propriétés esthétiques” que proposent les critiques.

Utilisant un autre vocabulaire et constatant que l'œuvre se projette sur la toile selon des lignes parfaitement objectives, donc *objectivistes* au sens de la philosophie de la perception, est-il possible d'en déduire par un enchaînement logique (conduit par l'ordinateur dans notre propos), une mesure de la qualité esthétique de l'œuvre ? La démarche que nous entreprenons ci-dessous a pour but de faire l'inventaire de tous les indices “objectifs” qui conduisent à des propositions esthétiques. Nous verrons dans les chapitres suivants s'il est possible ensuite d'exploiter ces indices, de les brasser et de les raffiner pour en extraire l'“évaluation esthétique” grâce à l'ordinateur.

1. “*Quoiqu'en regardant l'image on puisse dire, et à juste titre, qu'elle est délicate ou sereine ou calme ou maladive ou insipide, aucune description en termes non-esthétiques permettent de dire que ces termes ou tout autre terme esthétique, doit lui être indéniablement appliqué.*” [Sibley, 1959].

2. “*Les propriétés esthétiques surviennent sur les propriétés non esthétiques : pas de différence esthétique sans différence non esthétique. En gros, les propriétés esthétiques seraient ainsi survenantes en tant qu'émergentes : elles ne seraient pas nécessairement réductibles à des propriétés non esthétiques, physiques et perceptibles, mais elles en dépendraient et co-varieraient avec elles, c'est à dire qu'elles sont rattachées et ancrées sur ces dernières.*” [Cometti et al., 2000].

3. Ce goût (*taste* en anglais) serait à distinguer de celui qui nous permet de décider des saveurs et serait commun à tous les sens.

4. Rapporté dans [Cometti et al., 2000].

4.2 Le format ou rapport d'aspect

Les photographies sont produites selon des formats⁵ très variables (voir par exemple le chapitre 3 de [Maitre, 2015]). Le format carré est celui qui tire le meilleur parti de l'objectif photographique. Le format 2/3 a été posé en standard de fait pour des raisons commerciales du temps du film en bobine, le format 4/3 s'est imposé pour des raisons identiques lors de l'avènement des écrans d'ordinateur, considéré avant tout comme un outil de travail à l'époque de son développement. Le format 16/9 a accompagné l'essor des salles de cinéma afin de satisfaire le confort d'une observation à distance. Chaque fabricant propose, de plus, un format particulier lié à la géométrie de ses capteurs : 2/3, 4/3, 5/3, 16/9 . . . , tandis que les fabricants de papier photographique disposent de leurs propres formats, issus souvent des normes de l'imprimerie. Si l'on ajoute que la photographie peut se présenter en format “*portrait*” ou en format “*paysage*” selon le besoin, ce que ne peut faire ni le cinéma ni la vidéo, on voit qu'on dispose d'un très vaste champ du possible.

Il n'y a pas vraiment de recommandations esthétiques sur ce point et les ouvrages prônent une grande souplesse. Hormis ceux qui recommandent de suivre la règle du nombre d'or, les autres stipulent plutôt d'adapter le format à la scène représentée sans entrer dans une prescription précise⁶. Ainsi les marines sont très allongées, de même que les scènes de bataille et les plans de villes. Au contraire les natures mortes sont plus compactes, ainsi que les scènes d'intérieur ou de rues. Dans l'autre sens les portraits en pied gagnent à être très hauts tandis que les bustes seuls sont plus ramassés.

4.2.1 Que disent les statistiques ?

Plusieurs travaux se sont attachés à mesurer la répartition des divers formats parmi les œuvres (picturales surtout) réputées belles. Le rapport le plus probable des dimensions sur un grand nombre d'œuvres dépend du sujet, conformément aux recommandations d'Arnheim. Statistiquement, les tableaux verticaux sont préférentiellement dans le rapport 5/4, tandis que les horizontaux sont plutôt dans le rapport 4/3 [Arnheim, 1983]. Des résultats semblables ont été retrouvés dans des collections de photographies d'art, avec pour dimensions privilégiées les rapports 1/1, 2/3 et 3/4 [Arbellini, 2017], rapports typiques des formats de films.

4.2.2 Le nombre d'or

Il vaut $\Phi = \frac{1+\sqrt{5}}{2} \simeq 1,681$ et se trouve au cœur de nombreuses propriétés mathématiques. Il a fait l'objet de très nombreux travaux, d'une part pour tenter de justifier ses qualités esthétiques et donc son emploi depuis l'Antiquité, d'autre part pour le retrouver dans les constructions de la nature [Ghyka, 1931]. Des résultats un peu hâtifs de G. Fechner semblaient en avoir établi l'intérêt d'un point de vue perceptuel et, à la suite des publications de A. Zeising, puis de M. Ghyka, Φ s'est vu promu au titre de “divine proportion” et pour cela abondamment utilisé par de nombreux artistes au cours des deux derniers siècles (Seurat, Le Corbusier, Duchamp, Pissaro, Picabia, Léger, Dali).

Des travaux postérieurs ont largement montré l'arbitraire de beaucoup des vertus esthétiques attribuées à Φ [Markowsky, 1992, Neveux, 1995, Livio, 2008]. Nous retiendrons en particulier ce qui concerne l'aspect général d'une photographie : le nombre d'or n'est pas le rapport préféré pour le grand côté et le petit côté d'une surface rectangulaire quelconque (cette proposition est celle qu'a tenté de prouver Fechner, mais son résultat est entaché d'un biais méthodologique [Markowsky, 1992]) et le rapport le plus probable de ces dimensions sur un grand nombre d'œuvres considérées comme belles n'est pas égal à Φ .

5. Le mot “format” est ici pris au sens géométrique de rapport largeur/hauteur (rapport d'aspect ou *aspect ratio*). Ce terme a un deuxième sens qui indique le codage adopté pour représenter l'image (JPEG ou TIFF par exemple).

6. Ce qu'Arnheim exprime par : “*The more visual affinity there is between frame and picture, the more they can influence each other*” [Arnheim, 1983].

4.3 La composition

La composition est le choix de la disposition des divers motifs d'intérêt au sein de la photographie. Son rôle est d'offrir une ligne de lecture des éléments composant la scène [Bouleau, 2014, Crettez, 2017]. L'impact de la composition d'une photographie sur l'observateur peut être analysé au moyen de techniques d'oculométrie qui enregistrent le parcours de l'œil examinant l'image, l'enchaînement de ses visites et la durée de chaque visite [Yarbus, 1967] (voir la figure 4.2). Quoique F. Molnar ait pu distinguer "bonnes" et "mauvaises" compositions par l'examen de ces parcours (une bonne composition conduit à peu de transitions d'un site à l'autre et à de plus longues fixations [Molnar, 1977]), il demeure difficile d'aller beaucoup plus loin dans le jugement esthétique d'une œuvre par cette technique. On aimerait pouvoir coupler les examens oculométriques avec des enregistrements neurologiques, par exemple par IRMf, de façon à attacher les activations cérébrales à tel ou tel détail. Malheureusement les constantes de temps de ces deux dispositifs sont aujourd'hui incompatibles.



Figure 4.2 – Oculométrie sur une peinture de Vladimir Repin (*La visite inattendue*), à gauche. L'un des parcours du regard obtenus par A. Yarbus, à droite. Yarbus a donné plusieurs consignes à l'observateur afin de répondre à diverses questions : par exemple Quel est l'âge des personnes ? ou Quels vêtements portent-ils ? Chaque consigne conduit à des parcours différents. Ce parcours est obtenu sans consigne et correspond vraisemblablement assez bien à celui que l'on adopterait dans un musée. Notons que ce parcours n'est pas recalé sur l'image, contrairement à ceux que l'on obtient aujourd'hui. [Yarbus, 1967].

Notons cependant que les études oculométriques ont apporté un résultat annexe intéressant notre étude en mettant en évidence une différence dans la façon dont les experts en art ou les non-experts examinent une œuvre. Elles concluent que l'appréciation d'une œuvre d'art est modifiée par l'apprentissage, et par là le jugement esthétique qui lui est porté [Molnar, 1981, Nodine et al., 1993]. Cependant [Rosenberg and Klein, 2015] avance une interprétation différente, rattachant les écarts constatés à des différences de motivation entre les observateurs.

4.3.1 La règle du tiers (*Rule Of Third*, ROT)

C'est en fait la seule règle de photographie qui soit véritablement universelle ; on la trouve dans tous les manuels de photographie, même si ses origines remontent à la peinture du *XVIII^e* siècle [Amirshahi et al., 2014]. Elle prône que les zones d'intérêt de l'image doivent se trouver de préférence sur les deux horizontales ou sur les deux verticales qui partagent en trois les deux dimensions de la photographie (voir les images de la figure 4.3).

Plusieurs études ont été consacrées à la vérification de la règle du tiers. La plus complète propose, dans [Amirshahi et al., 2014], de tester la validité de cette règle sur une base d'images judicieusement choisies (les images "belles" sont sélectionnées parmi les mieux notées du site Photo.net), soit par un groupe d'observateurs, soit à l'aide d'un algorithme qui utilise comme zone d'intérêt une zone de



Figure 4.3 – La règle du tiers indique comment l'on doit disposer les zones d'attention au sein de l'image. Divisant longueur et largeur en trois, c'est sur les 4 lignes ainsi tracées qu'il faut placer les objets d'intérêt.

forte saillance. Les conclusions de cette étude, présentées sur la figure 4.4, font clairement apparaître que la règle du tiers n'est majoritairement pas respectée par les "belles" photos, ni plus ni moins que par les photos ordinaires.

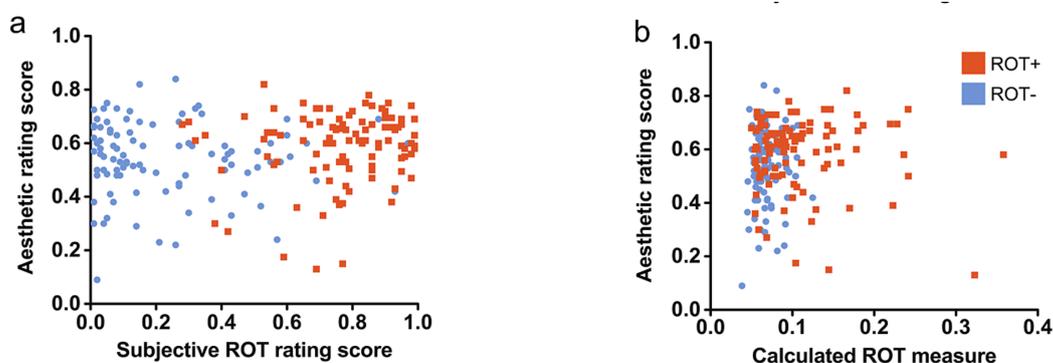
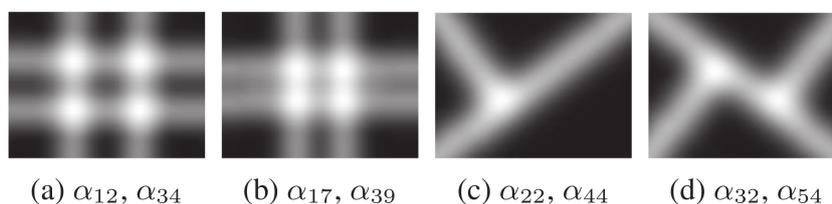


Figure 4.4 – Dans les résultats de l'expérience rapportée dans [Amirshahi et al., 2014], on observe la distribution des appréciations esthétiques de photographies, soit "belles" (notées ROT+), soit "ordinaires" (notées ROT-), en fonction de leur évaluation au regard de la règle du tiers. A gauche les résultats sont obtenus pour une ROT estimée par un observateur humain, à droite par un algorithme s'appuyant sur un calcul de saillance à l'intérieur de l'image tel que proposé dans [Mai et al., 2011].

Notons enfin que l'on trouve dans [Obrador et al., 2010] l'usage de proportions fondées sur le nombre d'or pour placer les points d'intérêt en place du tiers et également des triangles d'or pour placer des droites (voir la figure 4.5). Cet article montrerait que l'usage de ces règles de composition serait efficace pour sélectionner les "belles" images.



(a) α_{12}, α_{34}

(b) α_{17}, α_{39}

(c) α_{22}, α_{44}

(d) α_{32}, α_{54}

Figure 4.5 – Les masques permettant de détecter la composition de l'image selon [Obrador et al., 2010]. De gauche à droite : le masque de la règle du tiers, puis le masque de la moyenne d'or, puis un masque de l'un des triangles d'or, enfin un masque combinant deux triangles d'or. Chaque masque est obtenu par convolution d'une ligne ou plusieurs lignes avec un noyau gaussien. Il y en a 55 en tout, désignés ici par α_n .

4.3.2 Le centre de l'image

Mettre le point d'intérêt au centre de l'image est en opposition avec la règle précédente. C'est cependant l'une des règles proposées par le psychologue D. Ross en 1907. R. Arnheim a développé cette proposition [Arnheim, 1954, Arnheim, 1983]. Il distingue le centre géométrique du cadre et le centre de masse de l'image. Lorsque ces deux points se confondent, la construction apparaît équilibrée. Sinon il s'exerce une force de recentrage potentiellement nuisible à l'esthétique. Il suggère également que les axes horizontaux et verticaux, éventuellement diagonaux, ainsi que les coins constituent d'autres lieux de rappel de l'équilibre concurremment du centre. Une étude [McManus et al., 2011] s'est attachée à vérifier statistiquement cette propriété, sans vraiment convaincre⁷. D'autres études ont proposé des formules différentes pour mesurer l'équilibre apparent d'une composition [Hubner and Fillinger, 2016]. Elles pèchent par des protocoles de vérification utilisant des images très simples (formes géométriques noires sur fond blanc) et leurs résultats sont difficilement généralisables aux photographies.

Pour assouplir et généraliser un peu le rôle du centre de l'image, il a été proposé de définir la "focalité" comme l'existence dans l'image d'un point unique vers lequel converge le regard [Ulrich, 1983]. Une forte focalité est appréciée dans les images. Ce point est souvent en relation avec la présence de lignes fuyantes (voir ci-dessous).

4.3.3 Quelles autres règles de composition ?

Beaucoup d'autres éléments sont aussi présentés dans les manuels de photographie comme importants pour l'évaluation esthétique de l'image, mais à notre connaissance, ces propriétés n'ont pas fait l'objet d'études systématiques.

La symétrie On a vu que ce terme recouvre, à l'origine de l'esthétique grecque, des propriétés d'harmonie beaucoup plus étendues que dans son acception moderne⁸. C'est dans cette acception moderne que nous l'utilisons ici et plus particulièrement au sens de symétrie horizontale (donc par rapport à un axe vertical, l'axe le plus familier à un observateur doté également de cette symétrie). La symétrie horizontale est impérative dans les traités d'architecture classique et tout écart à son respect est jugé disgracieux. La symétrie horizontale contribue au "poids du centre" d'Arnheim sans lui être nécessaire, mais elle s'oppose à la règle du tiers. Quoique souvent estimée dans la composition pour exprimer un équilibre (en particulier dans la peinture primitive religieuse) elle ne trouve pas en peinture de soutien systématique, ni de recommandation formelle. Dans une approche plus moderne, et pour des applications de marketing, les rôles de la symétrie et de l'asymétrie pour convoier des sentiments de respectabilité ou au contraire de dynamisme dans l'affichage de logos ont été systématiquement étudiés sur des échantillons d'observateurs [Luffarelli et al., 2019]. Les conclusions exprimées en termes d'excitation perceptive (*arousal*) peuvent très probablement être étendues au domaine de la photo.

L'espace négatif On désigne parfois sous les termes "positifs" et "négatifs", les portions du champ de la photographie utilisées pour représenter les sujets principaux (ceux que l'on photographie) d'une part, et le fond ou l'environnement d'autre part. Quoique l'espace négatif ne soit pas primordial dans beaucoup de photos, il est recommandé de lui attribuer une place importante, mettant en valeur l'objet principal, et une place d'autant plus grande que l'espace positif est en mouvement ; il sera alors plus ample dans la direction du mouvement (figure 4.6 au centre).

Cette règle très empirique a des conséquences importantes pour définir le format de l'image (sujet abordé plus haut) et amène à choisir un format global faisant écho à l'aspect des objets d'intérêt (un

7. L'étude adopte une définition trop simple de la masse à partir des niveaux de gris et ne tient pas compte du biais statistique introduit par une image de taille bornée. De plus elle teste plutôt la qualité de l'observateur que celle du document. Elle conclut cependant que la taille et la couleur des stimulus sont de faible importance pour la sensation d'équilibre comparées à leur position, contrairement à une proposition d'Arnheim [McManus et al., 1985].

8. Voir paragraphe 2.1.

format allongé pour un lac ou une chaîne de montagne et plutôt carré pour un visage).

Les objets répétés et périodiques. Au contraire du cas précédent, on recommande souvent que ces espaces positifs ne soient accompagnés d’aucun espace négatif par un cadrage très serré, limité au seul espace positif, lorsque l’objet est régulier et que cette régularité est l’objet de l’image (c’est ce qui est illustré sur la figure 4.6 à gauche).

Cadrage des corps humains. Il est recommandé de ne couper les corps humains par un cadrage serré qu’en trois positions : à mi-poitrine, à la taille ou à mi-cuisse. Les autres césures sont réputées disgracieuses. Notons cependant que cette recommandation, très imprécise, laisse place à des choix très personnels.

Traitement des fuyantes. Lorsque l’image comporte des parallèles fuyantes dans l’espace objet, leur point de convergence dans l’image, virtuel ou réel est un point d’attraction de l’observation. Il doit être placé judicieusement soit s’il est le sujet principal de la photo, sur une ligne de tiers (voir la section 4.3.1), soit s’il appartient à l’espace négatif de façon à aider l’observation de l’espace positif (figure 4.6 à droite).



Figure 4.6 – *A gauche, sur un objet très régulier, le cadrage serré s’impose tandis que sur une scène en mouvement (au centre), l’espace négatif doit être ample. A droite, le point de convergence des fuyantes est le point fort de l’image.*

Usage des diagonales. Les diagonales constituent des lignes de force sur lesquelles on recommande souvent de s’appuyer soit pour introduire une profondeur dans la scène, soit pour accompagner un mouvement.

4.4 Mise au point et spectre de densité de puissance

[Mise au point] La mise au point est un élément très important dans l’évaluation de la **qualité technique** d’une photographie ([Maitre, 2015], chap 6). Sa relation à la beauté de l’image est complexe :

1. un défaut de mise au point est immédiatement perçu, et généralement de façon très négative, surtout s’il affecte l’objet principal ; pour de nombreuses scènes (extérieurs, nature), on attend une même mise au point sur tous les plans de la scène.
2. des photographes peuvent néanmoins utiliser le flou de mise au point pour exprimer des ambiances particulières : irréalité, distance, oubli, mystère⁹ ;
3. le traitement des arrière-plans s’accompagne souvent d’un flou important de façon à mettre en évidence l’objet d’intérêt. C’est une règle particulièrement importante pour les portraits, les

9. Notons l’emploi systématique du flou dans le cinéma et Noir et Blanc pour adoucir les visages de femmes.

photos de fleurs ou d'animaux. On apprécie alors souvent que la profondeur de champ soit très faible (figure 4.7, à gauche).

La recherche d'une très bonne résolution a souvent été un objectif des photographes¹⁰. L'avènement des images numériques avec des capteurs qui ne satisfont pas aux conditions d'échantillonnage de Shannon¹¹ a introduit le défaut opposé : le repliement de spectre (*aliasing*). Des dispositifs ont été mis en place pour réduire ce défaut et les boîtes à outils proposent des logiciels pour agir sur le spectre de densité de puissance du signal (modification de l'acutance¹² des photos). Les détails fins interviennent dans le piqué¹³ de la photo et contribuent, avec les textures, à la richesse de l'image. Les textures sont une partie importante de l'image, quoiqu'on les remarque peu. Ce sont des variations des valeurs de l'image qui "remplissent" les surfaces entre les contours et permettent de distinguer la nature des objets : herbe, peau, tissu, crépi ... L'impact des textures et des détails fins apparaît dans la décroissance du spectre de densité de puissance de la photographie. Des études ont montré que pour respecter une bonne hiérarchie des échelles, il convenait que la décroissance du spectre de densité de puissance suive une loi en $1/f^2$ (où f exprime la fréquence spatiale) [Alvarez et al., 1999, Gousseau and Roueff, 2007]. Le système visuel humain semble particulièrement adapté à ce type de signal [Brady and Field, 2000]. Des études ont montré que c'était bien le type de propriétés que l'on retrouve dans les images esthétiquement plaisantes [Koch et al., 2010]. De plus il a été montré [Schweinhart and Essock, 2013] que l'on appréciait que cette décroissance en $1/f^2$ soit régulière dans toutes les directions de l'espace, tandis que des écarts à cette règle conduisaient à des images "inconfortables" [Fernandez and Wilkins, 2008].



Figure 4.7 – Quelques images qui s'écartent des règles générales de la photographie. A gauche, La mise au point n'est faite que sur la fleur, le reste de l'image est flou. Au centre, image en low keys n'utilisant que des teintes sombres. A droite image en high keys n'utilisant que des fortes luminosités.

4.5 Texture et histogramme

4.5.1 Niveaux de gris

Il est souvent recommandé d'utiliser largement la gamme de luminosité disponible en évitant les zones saturées (claires ou sombres), ainsi que des plages non utilisées de l'histogramme [Kardan et al., 2017].

10. On rappelle en particulier l'éphémère mouvement d'esthétique photographique dénommé $f/64$ qui prônait l'usage exclusif de cette ouverture très petite et très contraignante, mais qui garantissait une excellente mise au point dans tout le champ.

11. Le théorème de Shannon fixe la distance qui doit séparer les échantillons d'un signal lors de sa numérisation pour ne pas introduire de défaut dans le signal ainsi créé. Le théorème de Shannon permet d'avoir un signal discret équivalent au signal continu original. Les conditions de validité de ce théorème portent sur l'extension du spectre de densité de puissance du signal continu. Les capteurs CMOS et CCD, constitués d'une mosaïque de photo-récepteurs, ne peuvent pas satisfaire aux conditions d'échantillonnage de Shannon (voir [Maitre, 2015], chapitre 2).

12. L'acutance est une mesure de la qualité d'une image qui prend en compte à la fois la fonction de transfert de l'appareil photo et la sensibilité du système visuel humain dans les conditions d'observation (voir le chapitre 6 de [Maitre, 2015]).

13. Le piqué est un terme de technique photographique qui désigne un bon rendu des détails fins accompagné d'une bonne résolution de l'image (voir [Maitre, 2015], page 220). L'optimisation de l'acutance est une façon de maximiser le piqué d'une photo donnée.

D'ailleurs, la plupart des caméras disposent par défaut de cette fonction d'étirement du contraste. Néanmoins de fameux contre-exemples offrent des dynamiques très faibles, l'ensemble de la photo se faisant par une modulation subtile des niveaux de gris (*low-keys* pour les images très sombres, *high-keys* pour les images très claires (figure 4.7)). Ainsi, il n'y a pas de recommandation impérative sur l'allure générale de la répartition des niveaux de gris (ou de la luminosité dans le cas de photos couleur). L'expérience montre que les histogrammes les plus variés peuvent accompagner de belles photos. Il n'y a donc pas de recommandation pour la valeur moyenne et pour l'écart-type de cette distribution de niveaux de gris. Surprenamment, des études ont montré que ces distributions vérifiaient des propriétés au troisième ordre (la moyenne μ et la variance σ étant les deux premiers). Si p est la distribution des niveaux de gris i , le biais γ est défini par :

$$\gamma = \frac{1}{N} \sum_{i=1}^N \frac{(p(i) - \mu)^3}{\sigma^3} \quad (4.1)$$

où N est le nombre de niveaux de gris de l'image.

On trouve dans [Attewell and Baddeley, 2007] une étude montrant que l'on préfère les images à biais négatif plutôt que celles à biais nul (comme dans les distributions gaussiennes) ou, à fortiori, à biais positif (figure 4.8). Les conclusions de [Motoyoshi et al., 2007] sont un peu différentes mais confirment l'importance du biais dans la qualité perçue. Cette sensibilité au moment d'ordre 3 serait expliquée par des mécanismes particuliers d'adaptation des cellules visuelles [Laughlin, 1987, Brady and Field, 2000, Motoyoshi et al., 2007].



Figure 4.8 – Le rôle du biais (moment d'ordre 3) dans l'apparence d'une image. Les photos du haut ont un biais négatif tandis que celles du bas, avec une moyenne et une variance identiques, ont un biais positif [Attewell and Baddeley, 2007].

4.5.2 Textures

Le rôle des textures dans l'attractivité esthétique des images a fait l'objet d'études assez complètes où se mêlent modélisation statistique et mesures psycho-visuelles [Liu et al., 2018]. Leur objectif est d'attacher des propriétés sémantiques esthétiques de haut niveau à des mesures de bas niveau des propriétés statistiques de l'image.

Le choix est fait de ne travailler que sur des images de texture pure (issues de la base de données SynTex). La représentation qui en est faite utilise des vecteurs de primitives (matrices de cooccurrence, mesures d'énergie et d'entropie, décomposition en ondelettes, ...). La centaine de primitives ainsi sélectionnée est réduite à une dizaine par une étape de sélection statistique.

Parallèlement parmi 20 couples d'antonymes sémantiques sensés décrire les sensations évoquées par

ces textures, 8 sont retenus comme les plus significatifs à l'issue d'une enquête auprès d'évaluateurs : *chaud/froid, sombre/clair, rugueux/doux, etc.*

Les mesures sont alors associées aux divers antonymes par une formulation non-linéaire pour qualifier, à travers un modèle perceptif complexe [Thumfart et al., 2011], issu d'un apprentissage, trois niveaux d'interprétation (voir la figure 4.9) :

- un niveau affectif ou perceptuel,
- une couche de jugement d'apparence,
- une couche émotionnelle.

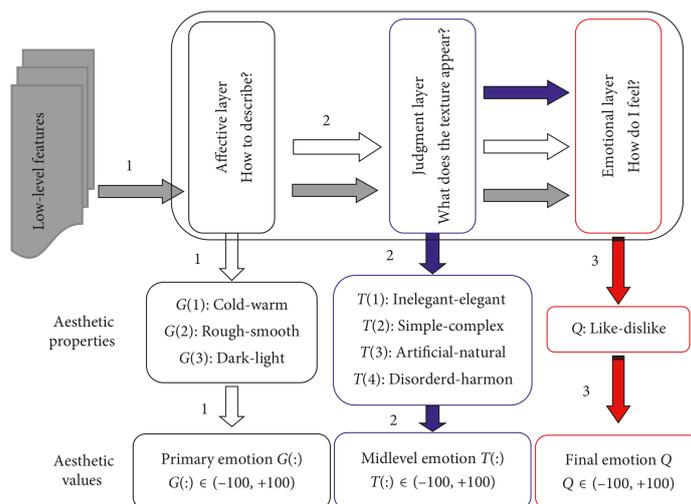


Figure 4.9 – Le modèle de perception de texture proposé par S. Thumfart et al. et utilisé par J. Liu et al. pour évaluer subjectivement l'esthétique d'une texture, repose sur trois niveaux d'interprétation : le niveau affectif, le niveau de jugement et le niveau émotionnel. Chaque niveau se caractérise par quelques couples antonymes dominants. Trois valeurs sont déduites des propriétés de bas-niveau de l'image pour chiffrer ces éléments du jugement esthétique [Thumfart et al., 2011, Liu et al., 2018].

Ces travaux procèdent d'une bonne compréhension des divers facteurs de l'évaluation esthétique. Leur limitation principale réside dans le schéma adopté pour résoudre les inconnues du problème, schéma qui fait abondamment appel aux tests psycho-physiologiques (donc mettant en jeu des observateurs humains). Ceux-ci ne peuvent alors explorer qu'une toute petite fraction du possible et il est très difficile de généraliser les résultats ainsi obtenus.

4.6 La couleur

La couleur est bien sûr un attribut très important de la photographie. Si l'on ne saurait restreindre les "belles" photos aux photos en couleur, il ne serait pas plus concevable qu'une photo en couleur soit jugée "belle" sans que l'on porte un avis positif sur la palette des couleurs qu'elle utilise. C'est d'ailleurs beaucoup sur leur couleur que portent les descriptions des œuvres d'art tant elles celle-ci est importante¹⁴. L'esthétique des couleurs mérite ainsi un traitement particulier.

Mais il semble souhaitable de revenir avant de passer à ces études sur un très court retour historique qui permettra de mieux comprendre pourquoi ce problème a été au cœur de discussions si profondes.

14. Illustrons ce point à l'aide d'une citation de Van Gogh tirée de l'un de ses dernières lettres à son frère, où il décrit le portrait qu'il avait fait de Mademoiselle Gachet en ces termes (et en ces termes seulement) : "La robe est rose, le mur dans le fond vert avec un point orangé, le tapis rouge avec un point vert, le piano violet foncé, cela a 1 mètre de haut sur 50 de large." in [van Gogh, 1992]

4.6.1 Du concept de couleur

Il pourrait sembler au scientifique que toute la connaissance est disponible sur la perception colorée à partir de 3 découvertes notoires :

1. la décomposition de la lumière blanche en un spectre de longueurs d'onde par Newton, en 1660,
2. la trivariance de la perception chromatique humaine exprimée par Grassmann, en 1853,
3. la linéarité des propriétés entre grandeurs radiométriques et photométriques, par Abney, en 1886 qui a permis le développement des travaux de la CIE¹⁵, à partir des années 1920.

Ces résultats sont cependant demeurés ignorés ou, lorsqu'ils étaient connus, ont été contestés, en particulier par le milieu artistique, jusqu'au début du XX^e siècle. Cette contestation doit être expliquée. La théorie chromatique telle qu'elle nous est connue aujourd'hui repose sur un socle physique très solide et sur une bonne connaissance de la physiologie, qui rendent bien compte de la perception dans un contexte expérimental statique, simple et contrôlé : larges plages colorées bien définies, éclairage stable et maîtrisé, généralement neutre, observateurs "normaux", reposés et attentifs, etc. Elle laisse cependant de côté de nombreux aspects plus complexes. Ainsi elle explique mal la perception éprouvée par l'observateur en présence d'un contexte chromatiquement très riche (avec des effets d'influence et de masquage), en présence de sources lumineuses variables (expérience de constance de couleur appelant des techniques de "balance de blanc"), de scènes mobiles, de flashes, etc. Elle ne rend pas compte des phénomènes d'adaptation et des visions chromatiques perçues après éblouissement, ou dans l'obscurité, ou sous l'effet de stupéfiants, dans les souvenirs ou en rêve. Or les artistes ont été très sensibles à ces phénomènes et ont tenté d'en rendre compte. Certains auteurs (Goethe, Schopenhauer) ont ainsi proposé des formulations sans réel fondement physique, mais très fortement imprégnées de ces expériences vécues. D'autres, souvent des peintres, développaient des approches pragmatiques construites sur leur vécu et la disponibilité de pigments. En gros, au verso d'une théorie des couleurs développée sur des bases de physique et de physiologie, se sont développées des modèles développés sur des bases subjectives et sensorielles.

Ainsi, pendant vingt siècles se sont développées des approches avec pour fondement l'harmonie platonicienne et des bases pythagoriciennes très incertaines où, par exemple, l'octave allant du noir au blanc se décomposait en intervalles simples, de tierce, ou de quarte. Ce sont les approches qui ont inspiré Goethe et Schopenhauer qui rejetaient les expériences de décomposition de la lumière de Newton. Pour ceux qui donnaient foi à l'approche newtonnienne, de nombreuses représentations se sont maintenues concurremment (Runge, Bezold, Chevreul, Itten) proposant différentes "roues chromatiques" où étaient placées en opposition des couleurs complémentaires, figure 4.10. Une grande confusion portait sur le nombre des primaires et sur leur choix ; elle s'est maintenue, en particulier dans le milieu artistique, tant que l'on n'a pas distingué synthèse additive (telle que réalisée par Newton à l'aide d'un prisme) et synthèse soustractive (obtenue par mélange de pigments par les peintres).

4.6.2 Les préférences des couleurs isolées

Examinons maintenant ce que l'on connaît du rôle de la couleur dans l'appréciation esthétique des images. Lorsque l'on cherche à caractériser l'intérêt que nous portons aux couleurs, il nous faut distinguer deux grands aspects bien distincts. Le premier concerne l'appréciation que l'on a pour les couleurs prises isolément (ce que nous examinons ici), le second concerne l'harmonie colorée, c'est-à-dire la façon dont est perçu un champ composé de plusieurs teintes juxtaposées (cela fera l'objet de la section suivante).

Il y a peu à dire en fait des règles générales de préférences en matière de couleurs. L'adage "*des goûts et des couleurs, on ne dispute pas*" laisse à penser qu'il n'y a ni règle ni rationalité dans le choix

15. CIE = Commission internationale de l'éclairage, c'est elle qui a conduit les études permettant la normalisation de la représentation chromatique.

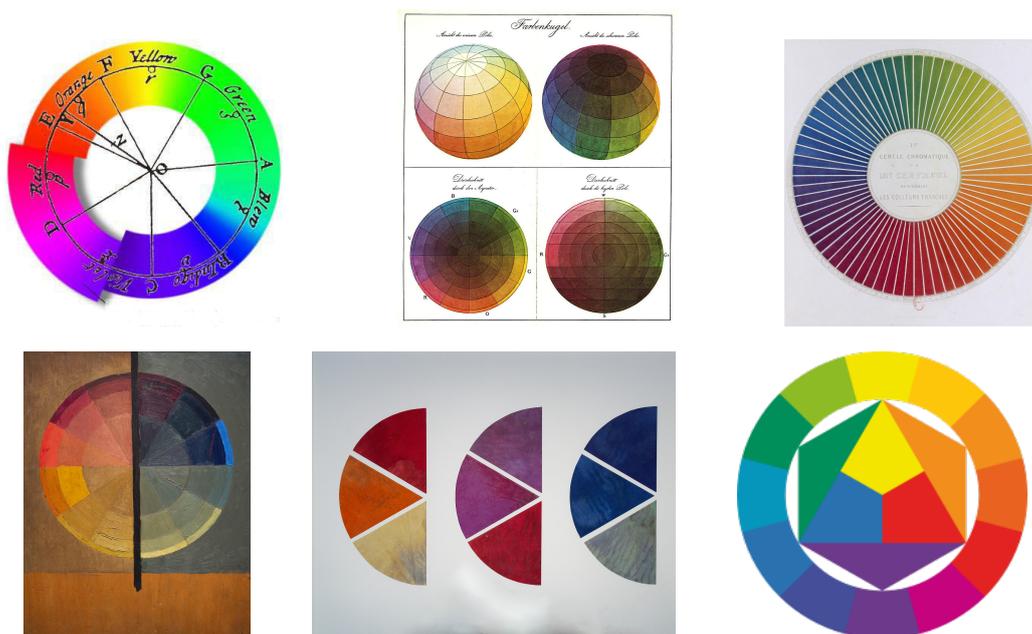


Figure 4.10 – Ligne du haut : à gauche, la première roue des couleurs telle que vue par I. Newton, au centre, la sphère des couleurs de P.O. Runge (source Wikipedia), à droite, un cercle chromatique de E. Chevreul (source : Gallica/BNF). Ligne du bas : à gauche le cercle de couleur de P. Sérusier (musée d’Orsay), au centre, 9 éléments de cercles chromatiques de W. Kandinsky (musée d’Orsay), à droite le cercle de couleurs de J. Itten (source Wikipedia). On constate que si l’ordre des couleurs est bien respecté, les couleurs complémentaires peuvent varier selon les auteurs.

des couleurs. De très nombreuses études ont cependant été conduites pour essayer de déterminer certaines tendances générales et pour les justifier¹⁶.

Une revue très complète est proposée dans [Hurlberg and Ling, 2012] des divers travaux menés depuis un siècle sur les préférences en matière de couleurs. Il apparaît très généralement que les couleurs *bleu* et *bleu-vert* sont préférées tandis que les teintes *jaune* et *jaune-vert* sont les moins appréciées [McManus et al., 1981, Ou et al., 2004]. Cet ordre de préférence est assez globalement indépendant des caractéristiques des observateurs (origine, sexe, âge, éducation), quoique de subtiles dépendances aient pu être parfois détectées (voir par exemple [Hurlberg and Ling, 2007, Taylor et al., 2013]). Il est à noter que ces préférences s’appliquent à des populations importantes mais souffrent de nombreux écarts individuels.

Des tentatives pour proposer des modèles mathématiques de prédiction de l’appréciation d’une couleur ont été faites, mais elles ne sont pas vraiment concluantes :

- en s’appuyant sur nos connaissances du système visuel ; exploitant la représentation LMS, [Vienot and Le Rohellec, 2012], [Ling and Hurlberg, 2007] utilise un modèle à base de quatre variables : *l’intensité*, *la saturation*, *le contraste vert-rouge L-M* et *le contraste jaune-bleu S-(L+M)*, les deux derniers facteurs expliquant 70 % des préférences ;
- à partir de critères de *psychologie affective* [Machajdik and Hanbury, 2010]. Les trois grandeurs *plaisir*, *excitation* et *dominance* s’expriment à partir de la luminance Y et de la saturation s

16. Nous ne rapportons ici que les études expérimentales qui confrontent des propositions à des vérifications sur le terrain. La littérature artistique fourmille d’affirmations qui reflètent les convictions de leur auteur mais ne sont pas soutenues par une validation reproductible. Elles associent souvent couleur et sentiment dans une approche très subjective. Illustrons ce que nous n’avons pas collecté à partir d’une citation, celle d’E. Delacroix : “Chacun sait que le jaune, l’orange et le rouge donnent et représentent des idées de joie, de richesse.” De telles formulations abondent, souvent pertinentes, mais inutilisables à moins de tracer le lien entre ces *attributs esthétiques* (cf. section 4.1) et les mesures de couleur.

par :

$$\begin{aligned} plaisir &= 0,69Y + 0,22s \\ excitation &= 0,31Y + 0,60s \\ dominance &= 0,76Y + 0,32s \end{aligned} \quad (4.2)$$

- en choisissant des pondérations émotionnelles des couleurs *activité colorée, poids et chaleur* [Ou et al., 2004]. Les émotions sont au nombre de 10, chacune dotée d'une polarité positive/négative, parmi lesquelles : *lourd/léger, moderne/classique, masculin/féminin, dur/doux, etc.*
- à partir d'arguments écologiques [Palmer, 2010] (voir la figure 4.11), s'appuyant sur 4 catégories : *couleurs saturées, légères, sombres et voilées.*

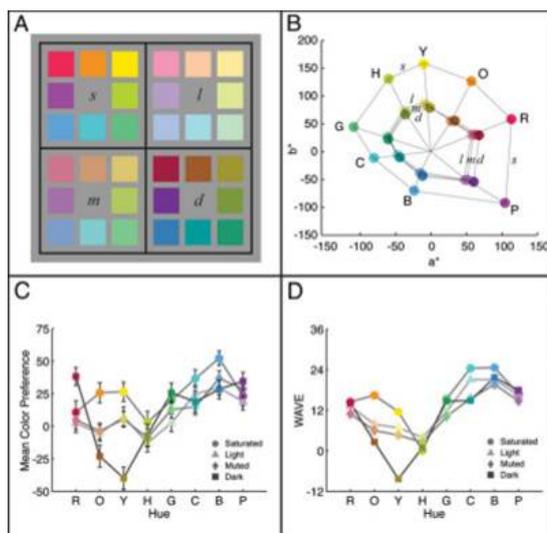


Figure 4.11 – Mise en évidence des préférences chromatiques. En A, les diverses couleurs utilisées lors de l'étude [Palmer, 2010]. On juge indépendamment des petites plages colorées, bien séparées. En B leur position dans le plan a^*b^* de l'espace Lab. En C le résultat de l'appréciation par les 48 observateurs (notée de -5 à +50), en D l'évaluation à l'aide du modèle WAVE (weighted affective valence estimate) proposé dans [Palmer, 2010], en assez bon accord avec l'expérience.

4.6.3 Les préférences concernant les palettes colorées

Si la littérature hésite à s'engager sur les recommandations concernant les couleurs prises isolément, elle est plus riche sur les propositions concernant l'assemblage de plusieurs couleurs dans une composition, ce que l'on appelle généralement la palette.

E. Chevreul est considéré à juste titre comme celui qui a le premier très largement exploré la richesse des harmonies chromatiques nécessaires, en l'occurrence, à la réalisation des tapisseries¹⁷. Néanmoins, son interprétation à partir des "complémentaires" est aujourd'hui réfutée. Une autre approche globale de la palette est celle d'Itten qui prône la recherche d'un équilibre. Cet équilibre peut être atteint avec un champ gris, ou en combinant des couleurs complémentaires. L'harmonie idéale est atteinte lorsque le bon équilibre est trouvé [Itten, 1973].

C'est à partir des travaux de A. Munsell, en 1905 (et à l'aide de son atlas rationnel des couleurs - figure 4.12), puis grâce aux travaux de normalisation de la CIE en 1931 (en particulier grâce à la mise en place de l'espace chromatique CIELab ainsi que des métriques associées), et enfin à ceux de

17. Chevreul a dépassé les travaux antérieurs (ceux de Goethe par exemple) qui ne considéraient que la teinte comme composante de l'harmonie, en distinguant les harmonies de teinte et les harmonies de contraste, qui font appel explicitement à la saturation et l'intensité en plus de la teinte [Chevreul, 1864]. En cela il a préparé les travaux de Moon et Spencer.

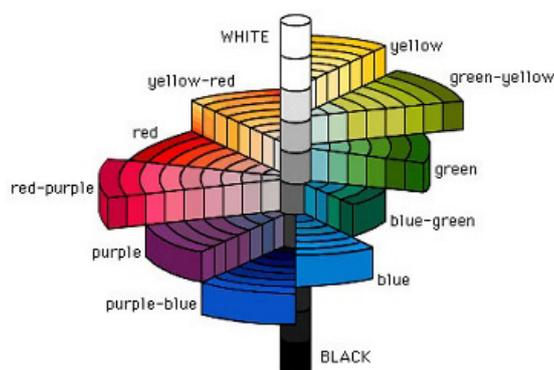


Figure 4.12 – *Le nuancier de Munsell (simplifié) qui sépare les diverses plages colorées selon des degrés de différence perceptuelle expérimentalement contrôlée.*

D.B. Judd, en 1943 (faisant le lien entre les intervalles expérimentaux de Munsell et la métrique de CIELab), qu’ont pu être menées des études systématiques et reproductibles d’harmonie chromatique.

L’harmonie selon Moon et Spencer :

Ce sont les travaux de P. Moon et D.E. Spencer qui constituent la référence scientifique la plus fréquemment invoquée lorsque l’on cherche les fondements mathématiques de l’harmonie colorée [Moon and Spencer, 1944c]. Notons cependant tout de suite que cette théorie donne les outils géométriques pour manipuler les couleurs ainsi que les règles pour évaluer la qualité esthétique d’une palette, mais qu’elle ne justifie les recommandations qu’elle fait que par l’accumulation des écrits antérieurs d’autres auteurs (Chevreul, Oswald, Bezold, Brücke, Field). Les auteurs n’invoquent ni arguments physiologiques, ni résultats de psychologie expérimentale. Pour les auteurs, les règles qu’ils proposent rendent compte d’une pratique universellement partagée dans les milieux artistiques. Elles mettent sous forme mathématique un savoir universel qu’ils considèrent avéré. Que disent les règles de l’harmonie chromatique de Moon et Spencer ? Elles recommandent de placer toute couleur dans un espace ω chromatique tel que le nuancier de Munsell, c’est-à-dire défini en coordonnées cylindriques par la luminosité, la teinte et la saturation (figure 4.13 à gauche). Les dix teintes de base de Munsell sont réparties de façon uniforme et définissent autant de secteurs du cylindre. Dans cette représentation ω , l’association de couleurs est harmonieuse si :

1. les relations entre deux couleurs quelconques ne sont jamais ambiguës,
2. les points chromatiques représentant ces couleurs sont disposés selon une distribution géométrique simple dans ω

Il y a ambiguïté de deux couleurs si leur teinte appartient à deux secteurs de teintes voisines (figure 4.13 à droite). Selon la loi 1, on autorise des teintes à être soit proches soit en contraste chromatique. Selon la loi 2, on peut faire varier deux ou trois variables à condition de laisser les points de couleur dans des plans ou sur des droites de l’espace ω . Il existe donc une infinité d’harmonies, les auteurs se gardant de les hiérarchiser. De plus, les auteurs proposent une classification à deux classes (harmonieuse ou dis-harmonieuse), reconnaissant cependant que la frontière est probablement floue entre les deux et qu’il est probablement possible de trouver des gradations dans l’harmonie.

Moon et Spencer proposent des valeurs numériques précises pour définir les zones permises et interdites (celles précisées sur les figures 4.13).

De ces postulats, Moon et Spencer tirent une classification méthodique des harmonies qu’ils explorent systématiquement. Trois classes principales sont distinguées selon le nombre de variables que l’on s’autorise à laisser libre. La première classe est ainsi obtenue en laissant se déplacer les points de couleur selon les 3 variables luminance, teinte ou saturation. Lorsque 2 variables sont libres, les points

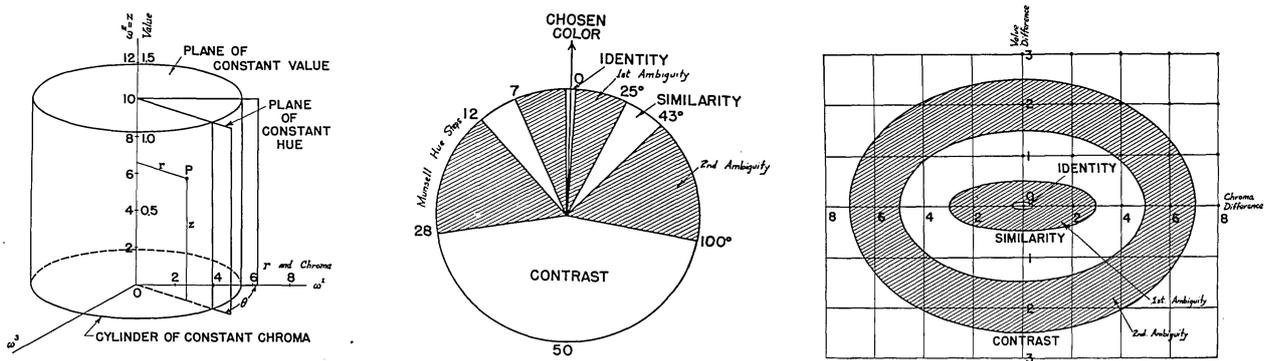


Figure 4.13 – Les travaux de Moon et Spencer : à gauche, l'espace de représentation d'une palette de couleurs. Pour les 2 dessins de droite, les zones hachurées représentent les zones interdites tandis que les zones blanches correspondent aux couleurs qui sont en relation harmonieuse. Au centre à luminance constante ; à droite à teinte constante, d'après [Moon and Spencer, 1944c].

de couleurs peuvent décrire des formes plus complexes et plus encore dans le cas à trois variables.

TABLE IV. Classification of harmonies.

Class I. One variable	
1. Variable z	
A. Achromatic	(a) Two color-points (b) Three points (c) More than three points
B. Chromatic	(a) Two points (b) Three points (c) More than three points
2. Variable r	(a) Two points (b) Three points (c) More than three points
3. Variable θ	(a) Two points (b) Three points (c) More than three points
Class II. Two variables	
1. In a plane of constant θ	A. Points on a straight line B. Triangles C. Rectangles D. Points on a circle
2. In a plane of constant z	(a) Two points (b) Three points on isosceles triangle (c) Five points on two triangles (d) n points on a circle with center on neutral axis (e) n points on a circle with center at a chromatic point
3. In a cylinder of constant r	(a) Two points (b) n points on an ellipse with center on neutral axis
Class III. Three variables	
1.	With reference to planes of constant θ
2.	With reference to planes of constant z
3.	With reference to cylinders of constant r
4.	With reference to tilted planes

TABLE V. Class I. Harmonies in one variable.

1A. Variable z , achromatic harmonies	
(a)	Two color-points
(1)	Harmony of analogy. Difference in z is between 4 and 12 ($\frac{1}{2}$ to $1\frac{1}{2}$ Munsell value steps). Example: $N 7, N 8$.
(2)	Harmony of contrast. Difference in z is between 20 and 80 ($2\frac{1}{2}$ to 10 Munsell value steps). Examples: $N 3, N 6; N 2, N 6; N 1, N 9$.
(b)	Three points
(1)	Small equal steps. Difference in z is between 4 and 12. Example: $N 4, N 5, N 6$.
(2)	Large equal steps. Difference in z is between 20 and 80. Example: $N 2, N 5, N 8$.
(3)	Large and small steps. Examples: $N 2, N 3, N 7; N 3, N 7, N 8$.
(c)	More than three points An obvious extension of the classification for three colors.

TABLE VI. Class I. Harmonies in one variable.

1B. Variable z , chromatic harmonies	
(a)	Two points
(1)	Harmony of analogy Examples: $R 6/8, 5/8; B 6/6, 7/6$.
(2)	Harmony of contrast Examples: $P 6/8, 3/8; RP 8/6, 3/6$.
(b)	Three points
(1)	Small steps Examples: $GY 7/8, 6/8, 5/8; Y 8/6, 7/6, 6/6$.
(2)	Large steps Example: $R 8/4, 5/4, 2/4$.
(3)	Large and small steps Examples: $BG 7/4, 6/4, 3/4; YR 8/4, 4/4, 3/4$.
(c)	More than three color-points

Figure 4.14 – Les divers types d'harmonie chromatique selon Moon et Spencer. A gauche l'ensemble des harmonies, à droite, celles à une seule variable libre, donc de type 1. D'après [Moon and Spencer, 1944c].

Notons qu'ils se sont plus intéressés à un examen systématique de toutes les configurations compatibles avec les deux postulats qu'ils ont établis en début d'étude qu'à une vérification de leurs qualités esthétiques. Ils ont cependant mené des expériences avec la majorité des harmonies qu'ils proposent¹⁸ et leur reconnaissent généralement d'intéressantes qualités, quoiqu'ils distinguent des configurations plus plaisantes, plus éclatantes ou plus subtiles.

18. Ces expériences étaient réalisées à l'aide d'assemblages de papiers colorés, suivant le nuancier de Munsell et disponibles dans le commerce.

Ces travaux ont été complétés d'études prenant en compte la taille des zones, ainsi que leur distance à un point d'accomodation chromatique [Moon and Spencer, 1944b]. Ces travaux peuvent se résumer ainsi :

1. une balance plaisante de n zones colorées est obtenue lorsque les moments colorés¹⁹ de ces zones sont égaux dans l'espace ω ;
2. un autre équilibre plaisant est obtenu si ces moments colorés sont dans des rapports simples.

Comme on le voit, l'approche de Moon et Spencer suit une démarche "objectiviste" de l'esthétique et laisse très peu de place à l'observateur dans l'affirmation de ses préférences. D'ailleurs, les auteurs ont affirmé cette approche en traduisant dans le formalisme objectiviste de Birkhoff (tel que nous l'avons présenté en section 5.2) leur expression de l'harmonie chromatique [Moon and Spencer, 1944a]. La formulation qui en résulte, qui prend en compte des coefficients de pondération heuristiquement déterminés, est très confuse et difficilement exploitable.

L'harmonie de Matsuda :

Y. Matsuda a proposé en 1995 une extension de l'approche de Moon et Spencer. Se plaçant dans l'espace CIE HSV²⁰, cette approche reconnaît, comme chez Moon et Spencer, des configurations harmonieuses (voir figure 4.15). Mais celles-ci sont distribuées différemment et peuvent rendre compte d'harmonies à plus de deux couleurs. La mesure de l'harmonie globale se fait par sommation de toutes les configurations compatibles avec les masques donnés, chaque configuration étant chargée d'un poids égal au produit des saturations, donnant ainsi un rôle plus important aux teintes vives. Notons qu'il n'y a aucune démonstration de la légitimité de cette approche dans les travaux de Matsuda, mais il s'appuie sur une longue pratique expérimentale.

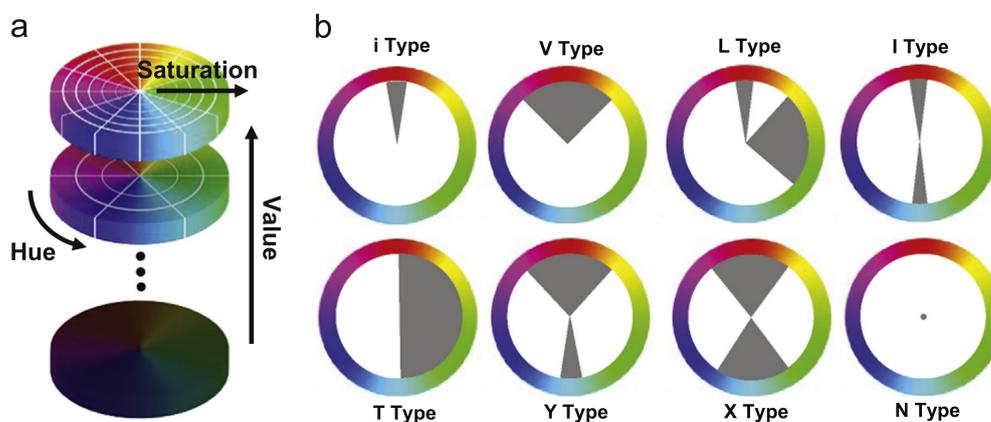


Figure 4.15 – Les configurations retenues par Matsuda. Le cercle chromatique peut librement tourner autour de son centre. Les configurations associent la teinte et la saturation. D'après une figure de [Lu et al., 2015a].

Des travaux plus récents ont cherché à déterminer les règles d'une harmonie universelle à l'aide de techniques d'apprentissage et de bases de données du web [Lu et al., 2014a, Lu et al., 2016]. Nous les examinerons plus loin (section 7.9).

19. Le moment coloré est le produit de la surface de la zone par sa distance au point d'adaptation chromatique (souvent sur l'axe des teintes neutres, mais potentiellement déplacé lorsque la scène a une couleur dominante), distance mesurée à l'aide des métriques de Munsell (exprimées en distances jnd (just noticeable difference)).

20. L'espace HSV (Hue, Saturation, Value) est un espace chromatique beaucoup utilisé dans le monde de l'informatique, retenu par la Commission Internationale de l'Eclairage (CIE).

Des harmonies qui ne sont pas universelles :

Depuis cinquante ans, de nombreux travaux s'inscrivent en faux face à l'hypothèse de couleurs universellement appréciées. De nombreux auteurs soulignent l'importance de la personnalité, de la culture, de l'éducation (voir par exemple [Mehrabian, 1977, Smith, 2008]).

Dans cette voie, Schloss et Palmer [Schloss and Palmer, 2011] ont revisité la notion d'harmonie chromatique en tempérant l'approche objectiviste. S'appuyant sur une analogie avec la musique²¹, ils ont pris en compte trois critères de qualité lors de l'appréciation d'une association de deux couleurs :

1. la *préférence* de l'observateur pour l'association des deux couleurs,
2. la perception de l'*harmonie* au sein de cette association,
3. l'appréciation de l'*accord* de ces deux couleurs avec les autres teintes considérées comme un fond.

La distinction entre *préférence d'une paire* et *harmonie d'une paire* leur semble nécessaire comme elle l'est pour distinguer des sons *harmonieux* de sons *agréables*. C'est sur cette distinction que peuvent se placer des différences inter-individus (attachées à l'éducation ou à la personnalité). Néanmoins, leurs conclusions ne laissent pas une grande place à cette distinction qui apparaît en pratique trop subtile pour se refléter nettement dans leurs expériences.

Poussant plus loin encore le refus de l'objectivisme, Z. O'Connor, dans l'intention de personnaliser complètement la mesure de l'harmonie [O'Connor, 2010] a retenu plusieurs facteurs particuliers qui interviennent dans la perception de la couleur : les différences individuelles (genre, âge, personnalité . . .), l'expérience culturelle, le contexte, les effets perceptuels et le temps (influence de la mode). Pour elle, ces facteurs s'additionnent et affectent multiplicativement une harmonie universelle (comme celle qu'aurait permis de calculer Moon et Spencer ou Matsuda) selon une formule du type :

$$\mathcal{H}_{O'Connor} = [\Phi_{individuel} + \Phi_{culturel} + \Phi_{contexte} + \Phi_{mode} + \dots] \mathcal{H}_{Matsuda} \quad (4.3)$$

Bien sûr une telle formule nécessite que de nombreux paramètres soient connus pour pouvoir être utilisée opérationnellement, ce qui est bien rarement le cas. On n'est donc pas surpris de peu rencontrer ces formules dans les travaux conduits sur d'importantes populations.

4.7 Que dit la psycho-sociologie ?

Nous avons examiné très en détail les propriétés des photographies telles que les manipulent les traiteurs d'image. Ces grandeurs sont directement issues de l'image et se prêtent bien à une interprétation objectiviste de l'esthétique puisqu'elles donnent à évaluer des grandeurs également disponibles pour tous.

D'autres études ont adopté des chemins différents, en questionnant les observateurs sur des aspects moins facilement mesurables de l'image et qui s'insèrent mieux dans une approche subjectiviste. Ces travaux ont pour méthodologie des enquêtes généralement conduites dans le cadre des départements de psychologie, de sociologie ou d'histoire de l'art. Ils offrent des résultats qui éclairent d'un jour très différent notre rapport au beau et ces résultats sont généralement très difficiles à inclure dans des approches automatisées car ils font référence à des concepts très mal modélisés de façon opérationnelle pour un informaticien. Nous ne pouvons cependant les ignorer.

4.7.1 Mesurer la beauté

Réponses à des enquêtes, les résultats obtenus reposent sur l'avis de l'observateur. Il est important alors que l'on comprenne bien ce que sera la réponse de l'observateur et, ceci fixé, que l'on trouve

21. L'analogie musicale proposée est de comparer des musiques de Mozart et de Stravinsky. Si l'on peut préférer Stravinsky à Mozart (critères individuels), on convient généralement de trouver Mozart plus harmonieux.

des mesures permettant de transporter les réponses d'observateurs différents dans un cadre où l'on pourra les comparer.

La dimension du Beau

Des travaux assez anciens se sont posé la question de l'espace dans lequel évolue le Beau. Est-ce une grandeur scalaire ? A-t-elle plusieurs dimensions ? Cet aspect a été abordé à l'aide des outils de la psychologie expérimentale, déjà anciens. Ils ont cherché à déterminer une échelle globale de mesure de l'appréciation esthétique et à la "dimension" du jugement esthétique [Dewar, 1937, Eysenck, 1939]. Ils s'appuient sur l'ordonnement d'une collection d'œuvres d'art par des expérimentateurs plutôt novices et par l'examen croisé de leurs résultats. Les œuvres choisies sont des reproductions de peintures, de vases, de fleurs, etc. qui doivent être ordonnées, à l'intérieur d'une même thématique, en fonction de leurs seules qualités esthétiques. Après une analyse statistique des résultats (plus de 300 observateurs sont employés) qui visent à éliminer l'observateur et à n'examiner que les relations entre œuvres, les auteurs concluent : "*L'analyse des facteurs suggère qu'un seul facteur général du goût artistique est responsable des classements artistiques.*" [Dewar, 1937]. Ce facteur est assimilé donc à la "Beauté" ²² et il lui est attribué une dimension 1. La seconde étude nuance ce résultat en montrant qu'un second facteur intervient si l'on introduit des documents abstraits dans le test (dessins, couvertures de livres, broderies). Ce second facteur disperse les résultats selon un axe allant du formel au représentatif [Eysenck, 1939] et les auteurs concluent que la dimension est alors 2.

A notre connaissance ces travaux n'ont été utilisés que pour concevoir des tests d'aptitude aux études artistiques mais n'ont pas été employés pour des études de mesure des qualités esthétiques.

Une échelle du Beau

Récemment, des travaux ont souligné la difficulté de travailler en esthétique sans disposer d'une mesure opérationnelle de ce qu'est le beau. En effet, la plupart des études ignorent ce problème en se reposant sur l'avis supposé sincère des expérimentateurs et sur un sens commun supposé universel.

Les travaux récents qui ont tenté d'aborder ce problème s'appuient généralement sur une étape intermédiaire, la recherche d'une mesure du plaisir hédonique, sensée refléter l'émotion éprouvée par l'observateur et donc mesurer directement le degré de beauté. Dans le domaine des arts visuels, ces travaux sont particulièrement balbutiants et leurs résultats sont loin de faire l'unanimité [Augustin et al., 2011, Hayn-Leichsenring et al., 2017, Blijlevens et al., 2017, Augustin et al., 2018] ; Ils ont plus de succès dans les champs du design ou de la conception des interfaces homme-machine, domaine où l'on sait mieux mesurer d'autres aspects liés à l'utilité fonctionnelle.

4.7.2 Images de la nature ou scènes artificielles ?

Dans le registre désigné parfois sous le nom d'esthétique environnementale [Ulrich, 1983], quelques études de psychologie expérimentale se sont attachées à rechercher des critères d'appréciation esthétique des photographies combinant une multitude de facteurs soit de bas niveau soit de haut niveau. L'idée directrice est de déterminer ce qui nous permet de décider qu'une image représente une scène naturelle (jardins, parcs, montagnes) ou un paysage en partie construit par l'homme (rues, bâtiments). Ils s'appuient sur des travaux déjà anciens [Kaplan et al., 1972], qui ont montré que l'aspect naturel est un critère important du jugement esthétique car il est bénéfique pour l'attention et la mémoire,

22. Il est intéressant de noter ici combien cette notion est fidèle à l'idée objectiviste d'un beau unique et universel. Au contraire, elle s'oppose frontalement à l'interprétation subjectiviste qu'exprime par exemple Xénophon, sophiste de l'école de Socrate, qui affirme : "*Un beau lutteur est différent d'un beau coureur, un bouclier qui est beau pour la défense est complètement différent d'un javelot, beau pour la vitesse et le lancé.*" (Xénophon, Commentaires, III,8,4, rapporté par [Tatarkiewicz, 1970]).

et ainsi pour stimuler les mécanismes impliqués dans l’appréciation esthétique²³

Ainsi, dans [Kardan et al., 2017], on compare le rôle des primitives de bas niveau dans la reconnaissance des scènes naturelles anthropisées. A partir de primitives issues de l’étalement de l’histogramme, de la palette des couleurs, de la distribution des contours ou de l’entropie, les auteurs montrent que l’on peut raisonnablement prédire si la scène est naturelle ou anthropisée et en déduire le jugement esthétique d’un observateur. On vérifie au passage que, à complexité équivalente, les scènes de nature sont préférées aux paysages artificiels.

Les trois primitives les plus significatives sont :

- l’absence de segments droits dans les contours, les ombres et les silhouettes,
- une dominante chromatique plutôt portée par l’antagonisme jaune-vert que par l’antagonisme bleu-rouge,
- une grande dynamique dans la saturation chromatique.

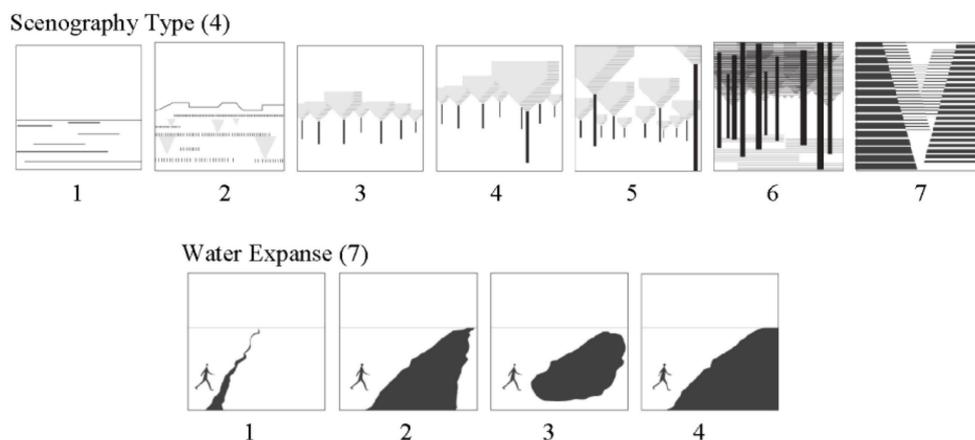


Figure 4.16 – Deux exemples des primitives de haut niveau proposées par [Hunter and Askarinejad, 2015] et utilisées dans [Ibarra et al., 2017] pour mesurer les propriétés “naturelles” ou “artificielles” d’une scène. La scénographie décrit la répartition d’ensemble des diverses masses de l’image, tandis que la mesure nommée Water Expanse s’intéresse à la disposition d’une masse d’eau. Les autres mesures également significative pour mesurer l’aspect naturel d’où découlerait la préférence esthétique sont : la distribution des bâtiments, les surfaces construites pour permettre le mouvement, la géométrie de la ligne d’horizon et l’ondulation maximale de la ligne d’horizon (qui correspond à sa variation totale).

Les auteurs de [Kardan et al., 2017] laissent une remarque en suspens. Lorsque l’on choisit des photos où les primitives ci-dessus ne sont pas déterminantes, on parvient cependant à conclure entre images naturelles et anthropisées, mais au prix d’un temps de réaction plus long qui laisse penser que d’autres mécanismes que le système visuel sont actifs.

Dans [Ibarra et al., 2017], l’étude précédente est complétée par la prise en compte non seulement de primitives de bas niveau, mais aussi de primitives de haut niveau. Celles-ci sont choisies parmi celles proposées par [Hunter and Askarinejad, 2015] et traduisent des aspects d’une grande variété thématique (végétation, bâti, voies de circulation) et d’une grande richesse sémantique (obstacle à la pénétration, franchissement, organisation ...). Ces primitives sont au nombre de 62; 2 sont

23. Les travaux de [Kaplan et al., 1972] ont été revisités dans [Kotabe et al., 2017] qui cherche à expliquer notre préférence pour les scènes naturelles. H.P. Kotabe et ses collègues aboutissent à une conclusion paradoxale : les scènes naturelles sont effectivement préférées, malgré leurs propriétés “désordonnées”, ce qui est contradictoire avec notre préférence vérifiée assez régulièrement pour l’ordre [Berlyne, 1971, Ulrich, 1983]. Ils testent trois hypothèses concurrentes pour résoudre ce paradoxe :

1. la préférence pour les scènes naturelles est plus forte que l’aversion pour le désordre,
2. le désordre n’intervient pas dans l’évaluation de scènes naturelles (masquage par la sémantique),
3. le désordre est apprécié dans les scènes naturelles.

Ils concluent que la première hypothèse est la plus probable.

présentées sur la figure 4.16. Les auteurs montrent que six d'entre elles permettent, mieux encore que les primitives de bas niveau, de prédire le jugement esthétique d'un observateur. La séparation entre naturel et anthropisé se fait à près de 70 % sur ces critères plutôt que sur les indices de bas niveau, tandis la prédiction esthétique les utilise à plus de 50 %.

Ces études sont conduites sur quelques centaines de photos et soumises au jugement d'une cinquantaine d'évaluateurs expérimentés. Les traitements de bas niveaux sont assurés par un logiciel de traitement d'images tandis que les mesures de haut niveau sont le résultat de la décision d'un expert et sont mesurées numériquement par l'intermédiaire d'une tablette graphique.

4.7.3 Le rôle du contexte

Ces travaux trouvent leur place au sein d'une interprétation historique de l'esthétique (voir section ??). Ils visent à mesurer l'influence sur l'évaluation d'une œuvre, des informations qui peuvent accompagner son examen.

Les travaux de V. Swami [Swami, 2013] se sont attachés à déterminer l'influence du type de connaissances attachées à une œuvre sur deux propriétés particulières de l'observateur : d'une part sa capacité à comprendre l'œuvre, d'autre part le jugement esthétique qu'il porte sur celle-ci. Les expériences conçues procèdent en collectant les opinions de quatre cohortes d'expérimentateurs confrontés à la visualisation de peintures abstraites, statistiquement équivalentes, mais dotées d'informations préalables différentes. Les réponses attendues permettent de tester tout d'abord le niveau de compréhension de l'œuvre, puis de mesurer son évaluation esthétique.

L'expérience est constituée de trois tests différents :

- Tout d'abord, devant un lot de peintures de Max Ernst, les cohortes sont dotées, (I_1) soit d'aucune information, (I_2) soit du titre de l'œuvre²⁴, (I_3) soit du titre et d'une présentation de l'artiste mentionnant quelques traits de sa méthode artistique (mais sans référence à l'œuvre présentée), (I_4) soit, en plus des informations précédentes, d'une description de l'œuvre et d'une note sur l'intention de l'artiste pour la produire.

Les conclusions de cette étude montrent que, si la capacité à comprendre l'œuvre progresse régulièrement de I_1 à I_4 , la valeur esthétique attribuée à l'œuvre est identique pour I_1 , I_2 et I_3 et ne progresse très sensiblement que pour la cohorte I_4 . Cette expérience montre que la valeur esthétique est dépendante d'une connaissance sur la manière et l'intention de cette œuvre.

- Dans cette seconde étude, on cherche à déterminer l'influence de l'information additionnelle (de type I_4 ci-dessus) dans le cas d'œuvres abstraites ou représentationnelles (des peintures de Picasso issues soit de sa période cubiste dénotées P_c (donc abstraites) soit de sa période Bleue, dénotées P_b). Pour cela on réduit les tests aux expériences I_1 et I_4 ci-dessus, faites sur les informations P_c ou P_b .

Les résultats montrent que l'information contextuelle I_4 est très importante pour améliorer la compréhension de P_c , mais peu utile pour la compréhension de P_b . Cette information permet également de faire croître très sensiblement le jugement esthétique de P_c et plus modestement celui de P_b .

- Dans une troisième expérience, l'auteur souhaite tester le rôle d'une information non-pertinente dans l'interprétation et le jugement de l'œuvre. Pour cela il reprend l'expérience 1 d'évaluation d'œuvres de M. Ernst ci-dessus, avec deux cohortes recevant soit I_1 soit I_4 , puis deux autres cohortes qui reçoivent l'information P_c et P_b ci-dessus (exposant la technique de Picasso), plus enfin une cohorte utilisant l'information P_b mais avec un autre nom que Picasso comme auteur. Les résultats montrent que comme dans la première expérience, l'information exacte contribue très largement aussi bien à la compréhension qu'à l'appréciation. Au contraire, les fausses informations ne contribuent ni à l'une ni à l'autre : il n'y a pas d'évolution sensible des notes attribuées.

24. D'autres études ont été faites sur le rôle du seul titre dans l'appréciation esthétique [Leder et al., 2006].

On déduit de ces travaux que l'appréciation esthétique dépend de la connaissance que l'on a sur l'intention et la manière de l'auteur lors de sa création. Ces résultats se transportent probablement dans le domaine de la photographie, mais il faut prévoir que leur rôle est important. Probablement faudrait-il mesurer de façon identique l'apport d'information sur le contexte historico-géographique de la prise de vue, mais il n'est pas clair que l'accroissement de l'intérêt porté à l'image ne serait pas dû principalement à un effet de *relevance* (voir le paragraphe 1.1).

4.7.4 La perception globale : unité et prototypalité

On doit aux écoles de *design* des études plus formelles sur des éléments qui influencent le jugement esthétique d'une scène. Ces travaux ont été orientés vers la compréhension des avis portés sur les œuvres créées par l'homme : mobilier, architecture d'intérieur, haute-couture, mais rien ne s'oppose à ce que leurs résultats soit importés dans le champ des beaux-arts et de la photographie.

Ils ont permis de dégager des critères spécifiques et nous en retiendrons deux qui nous semblent particulièrement pertinents : l'unité et la prototypicalité.

Unité : L'unité est relation qui relie les diverses parties de la scène, en particulier en mettant en relation d'harmonie des caractéristiques visuelles locales saillantes. Une photo exprimera des qualités d'unité si les diverses parties expriment les mêmes propriétés perceptives : contraste, finesse des détails, mise au point, palette de couleurs, etc. Les travaux [Bell et al., 1991] et [Veryzer Jr and Hutchinson, 1998] apportent des éclairages sur le rôle de l'unité dans l'appréciation d'une scène. Ces études suggèrent que les objets esthétiquement attractifs présentent une forte unité, tandis que les objets à faible unité reçoivent des avis esthétiques plus faibles.

Prototypicalité : La prototypicalité est la propriété d'un objet qui le rattache à une catégorie ou aux valeurs moyennes des objets de cette catégorie. Certains domaines de la photos s'exposent plus à cette propriété : le nu, les portraits, les photos d'animaux, les couchers de soleil. Les études conduites sur la prototypicalité dans le domaine du design aboutissent à une conclusion assez unanime que plus de prototypicalité conduit à plus d'esthétique [Barsalou, 1985], ce qui est cependant assez paradoxal. Certes, en s'approchant de la forme type on réunit très probablement le plus grand nombre de traits appréciés dans chaque expression de la forme. Certes on a l'impression que l'on retrouve dans le prototype les "canons" de la beauté aristotélicienne. Certes on s'attend avec des formes proches du prototype à offrir des lignes familières à l'observateur et donc à lui faciliter le travail. Mais cependant on écarte par là-même des éléments justement au cœur de l'œuvre artistique et qui en sont usuellement inséparables : l'innovation et la création. Ces résultats sont cependant confirmés par de nombreuses expériences [Whitfield and Slatter, 1979, Veryzer Jr and Hutchinson, 1998] quoiqu'aucune n'ait été conduite dans le domaine de l'image ou de la photographie où la notion de prototype mériterait d'être développée²⁵ ?

25. Une exception doit être faite cependant dans le domaine de l'esthétique des portraits où des travaux se sont attachés à définir un visage moyen et où il a été montré que ce visage est particulièrement attractif [Trujillo et al., 2014]. On est cependant loin des critères de beauté tels que ceux adoptés dans cet article.

Chapitre 5

Les approches algorithmiques de calcul de la beauté

Nous allons examiner maintenant les travaux qui ont proposé de faire calculer la beauté d'une photographie (ou d'une image, ou d'une peinture) par un programme et non plus par un humain.

Ces travaux se divisent en deux grandes parties très inégales : la première portée par ce chapitre est consacrée aux méthodes algorithmiques qui proposent de confronter la photo à un calcul dont les paramètres sont connus et universels. Ces méthodes s'appuient donc beaucoup sur l'approche objectiviste. La seconde partie traite des méthodes qui s'appuient sur un apprentissage à partir d'exemples notés par un observateur humain. Cette partie sera abordée en trois chapitres. Le premier examine les bases de données qui servent à l'apprentissage ainsi que l'expertise utilisée en référence. Le suivant examine les méthodes développées de 2000 à 2014 qui utilisent des primitives extraites des images, des règles mises en œuvre des classifieurs agissant sur les primitives. Le chapitre suivant est consacré aux réseaux de neurones profonds qui ont été utilisés depuis 2015.

5.1 Les premiers essais : Ch. Henry

Les premières contributions à cette détermination automatique de la beauté sont dues à Charles Henry [Henry, 1885, Henry, 1891]. Identifiant très lucidement que la détermination exacte des modifications physico-chimiques corrélatives aux perceptions esthétiques était hors de la portée de la science de son temps, il a proposé un cadre général pour évaluer les liaisons entre excitation et sensation, quelle que soit la modalité de perception. Fondée à la fois sur l'harmonie "antique" des proportions (en particulier les 12 premières puissances positives et négatives de $3/2$ et les polygones inscrits réguliers), ainsi que sur les théories thermodynamiques de Van't Hoff et le principe de moindre action de Lagrange, cette théorie introduit un principe unificateur de *dynamogénie* qui nous semble aujourd'hui assez confus [Henry, 1889]. Il semble avoir retrouvé ainsi par les mathématiques des évaluations esthétiques des harmonies chromatiques corroborées par les travaux de Chevreul et d'Helmholtz, des formes [Henry, 1895] ainsi que des musiques telles qu'enseignées par Rameau. Les travaux d'Henry, quoiqu'ils furent très bien reçus en leur temps, n'ont pas été repris à sa mort.

5.2 L'approche mathématique de G. D. Birkhoff

Le mathématicien et probabiliste Georges D. Birkhoff a adopté une approche tout à fait différente, mais elle aussi empreinte de l'idée fondamentale que la beauté réside dans l'objet et qu'elle peut se mesurer par une combinaison judicieuse de ses attributs [Birkhoff, 1933]. Il propose pour cela une

formule de la beauté qu’il teste assez systématiquement sur des données très variées. Cette formule relie trois grandeurs :

- une grandeur, notée \mathcal{M} , qui mesure la *beauté* de l’objet, c’est-à-dire sa valeur esthétique et qui apparaît comme une récompense de l’effort qui est réclamé pour examiner l’objet¹ ;
- une grandeur, notée \mathcal{O} , qu’il appelle l’*ordre* et qui mesure les qualités d’harmonie, de régularité et de symétrie indispensables à l’expression esthétique ;
- une grandeur, notée \mathcal{C} , qu’il appelle la *complexité* et qui correspond à l’effort d’attention ci-dessus.

Il propose alors pour conjecture :

$$\mathcal{M} = \frac{\mathcal{O}}{\mathcal{C}} \quad (5.1)$$

Il passera une bonne partie des dernières années de sa vie à vérifier la validité de cette conjecture dans diverses situations : en poésie, en musique, en dessin, en sculpture. Elle exprime la beauté comme une sorte de “récompense” à un effort consenti par l’observateur pour intégrer la richesse de la scène. Notons qu’il ne dispose à cette époque d’outillage ni pour décrire des formes (cela viendra plus tard avec les outils de traitement d’images et de reconnaissance des formes), ni pour exprimer la complexité (cela viendra avec les travaux de Shannon et de Kolmogoroff). Enfin il ne peut s’appuyer sur les travaux de psychologie expérimentale de l’école de la Gestalt qui sont publiés dans des cercles assez étroits dans cette période [Wertheimer, 1922, Wertheimer, 1923] et lui sont encore inconnus. Birkhoff est donc amené à exprimer de façon très spécifique les termes de son équation pour chaque problème traité, ce qui réduit considérablement la portée de ses travaux : pour chaque figure (vase, frise), il lui faut créer un “dictionnaire” des primitives, puis recenser les propriétés qui créent un “ordre” parmi ces formes (voir figure 5.1). On trouvera également dans [Delahaye, 2015] une discussion des résultats de Birkhoff.



Figure 5.1 – Birkhoff a expérimenté la validité de son équation sur des vases chinois en porcelaine. La complexité de la forme s’exprimait à partir de la diversité des plans tangents, des courbures et des symétries dans divers plans, ainsi que par la position des points d’inflexion et des extrêmes de courbure. Le résultat de ses mesures de beauté était finalement comparé au prix affiché de ces vases dans des salles de vente, considéré comme l’expression objective de leur beauté. L’accord des deux mesures l’a conforté dans ses choix (source [Birkhoff, 1933]).

1. Cette attention doit aller jusqu’à l’empathie de l’observateur avec son objet pour que s’exprime pleinement la dimension esthétique de l’objet. Cela est directement dans la ligne des travaux du psychologue T. Lipps qui a inspiré Birkhoff [Lipps, 1903].

5.3 A la suite de G. Birkhoff

5.3.1 L'après-guerre : les *designers*

Après-guerre, sous l'impulsion des grands courants de design industriel, les formulations de Birkhoff inspirent plusieurs chercheurs qui prennent en compte à la fois les progrès faits en psychologie expérimentale et en théorie de l'information. Ce sont en particulier A. Moles en France [Moles, 1957] et M. Bense en Allemagne qui apportent sur ces points les contributions les plus originales [Bense, 1969, Bense, 2007]. L'un et l'autre sont fortement imprégnés de la toute nouvelle théorie de l'information, de sémiotique et de cybernétique qu'ils appliquent à la création artistique. Ainsi, Bense propose de transformer la formule 5.1 en utilisant l'entropie de Shannon H appliquée à la seule couleur de l'image. L'ordre \mathcal{O} s'exprime alors comme la réduction d'entropie entre la palette complète équi-répartie H_{max} et la palette effectivement utilisée pour une œuvre donnée, H_o , conduisant à une formule de beauté :

$$\mathcal{M}_B = \frac{H_{max} - H_o}{H_{max}} \quad (5.2)$$

5.3.2 Une approche dynamique : P. Machado et A. Cardoso

Plus tard, P. Machado et A. Cardoso [Machado and Cardoso, 1998] proposent une formule "à la Birkhoff" (dont il semble cependant qu'ils n'aient pas eu connaissance), où la complexité \mathcal{C} et l'ordre \mathcal{O} sont remplacés respectivement par la *complexité de traitement* PC , et la *complexité de l'image* IC . Après quelques manipulations, la formule 5.1 devient :

$$\mathcal{M}_{MC} = \frac{IC^a}{(PC(t_0)PC(t_1))^b \left[\frac{PC(t_1) - PC(t_0)}{PC(t_1)} \right]^c} \quad (5.3)$$

où les exposants a , b , et c sont arbitraires et les instants t_0 et t_1 expriment le temps d'observation. Les auteurs évaluent enfin la différence entre la complexité de l'image et la complexité de traitement par le système visuel à l'aide de la différence entre image originale et image compressée (par un codage JPEG ou un codage fractal). Ces travaux sont remarquables car ils proposent une intéressante vérification expérimentale à partir d'une base d'images de tests de capacités artistiques. Par ailleurs, ils introduisent explicitement la durée d'observation $t_1 - t_0$ comme paramètre de la beauté, une hypothèse qui sera également au cœur des travaux de J. Schmidhuber.

D'autres auteurs ont exploité plus à fond la théorie de la complexité pour en dériver des mesures de \mathcal{O} et \mathcal{C} en utilisant les travaux de Kolmogorov et de Levin sur la complexité [Koshelev et al., 1998].

5.3.3 Les travaux de J. Rigau *et al.*

Il faudra cependant attendre plusieurs années pour que ces travaux soient adaptés à une étude plus construite des œuvres peintes. J. Rigau, M. Feixas et M. Sbert ont produit le travail le plus accompli à partir de la formulation de Birkhoff, appliquée à l'évaluation des qualités esthétiques de peintures des XIX^e et XX^e siècles [Rigau et al., 2008]. Pour cela ils ont mobilisé les outils de traitement des images (segmentation par zones, classification des couleurs) et la théorie de la complexité.

Plusieurs formulations sont proposées. Tout d'abord la complexité de Kolmogorov K peut avantageusement être employée dans l'expression 5.2 :

$$\mathcal{M}_{RFB_1} = \frac{nH_{max} - K}{nH_{max}} \quad (5.4)$$

où n est le nombre de pixels de l'image. La complexité de Komogorov K étant non calculable, on l'approche en utilisant les algorithmes de compression du commerce, en l'occurrence JPEG (comme

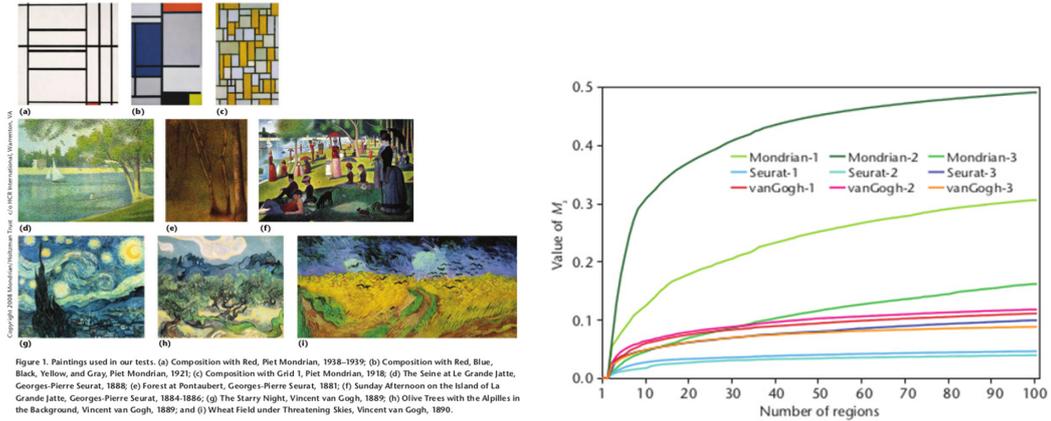


Figure 5.2 – Dans leur travail, Rigau et al. ont choisi une base expérimentale de peintures très variées qu'ils ont décomposées en plages colorées. La mesure de beauté de Birkhoff est déduite d'un calcul complexe mettant en jeu la complexité de Kolmogorov. A droite, l'évolution de la beauté de Birkhoff en fonction du nombre de plages retenues (source [Rigau et al., 2008]).

dans [Machado and Cardoso, 1998]). La formule ci-dessus est ensuite raffinée en utilisant l'entropie physique H_o de Zureck qui conduit à :

$$\mathcal{M}_{RFB_2} = \frac{nH_o - K}{nH_o} \quad (5.5)$$

Finalement les auteurs construisent un algorithme de segmentation de l'image par arbre, en ν zones, sur un critère de couleur et expriment la variation de la beauté de l'approximation $B(\nu)$ obtenue, représentée par la description $\hat{R}(\nu)$ par la formule (illustrée sur la figure 5.2 à droite) :

$$\mathcal{M}_{RFB_3}(\nu) = \frac{H[B(\nu), \hat{R}(\nu)] - H[B(\nu)|\hat{R}(\nu)]}{H[B(\nu)]} \quad (5.6)$$

où $H(x, y)$ exprime l'entropie jointe des processus x et y et $H(x|y)$ l'entropie conditionnelle de x sachant y .

Ces travaux méritent une expérimentation plus complète. Tout d'abord, il semble difficile de traiter de la même façon des styles très différents en raison du fort contexte culturel qui ne manque pas de masquer les seules évaluations esthétiques comme nous l'avons souligné en introduction. Par ailleurs, la prise en compte des propriétés spécifiques du système visuel humain semble indispensable lors de l'évaluation de la "complexité" de l'image, la métrique de l'espace chromatique et les effets de masquage en particulier.

5.4 Une approche algorithmique d'intelligence artificielle

Les travaux de J. Schmidhuber ne sont pas liés aux précédents en ce qu'ils ont recours à une formulation originale de la beauté [Schmidhuber, 1997]. Ils ne sont pas non-plus spécifiquement applicables dans un contexte objectiviste puisque leur auteur défend l'idée que l'évaluation de la beauté dépend de l'observateur et des conditions précises de l'observation. Il nous semble cependant que l'approche algorithmique proposée, ainsi que le poids accordé à la simplicité dans la mesure de la beauté les placent dans la lignée de ces travaux que nous venons de voir, plus que dans l'esprit des méthodes d'apprentissage que nous verrons plus loin.

J. Schmidhuber choisit une approche de théorie de l'information et de codage, qu'il appelle le modèle du "cerveau paresseux", où l'image, notée D est transmise à l'observateur qui s'en fait une

représentation D' par un algorithme de codage noté C . D, D' et C peuvent être décrits par des chaînes de symboles, conformément à la représentation des objets informatiques proposée par Turing. C représente, à un instant donné, toutes les connaissances que l'observateur peut avoir acquises sur D . Si $D' < D$, alors D est redondant par rapport aux connaissances de l'observateur et il est donc compressible au sens de la théorie de l'information. Pour J. Schmidhuber, il y a beauté s'il y a une forte probabilité de D' conditionnelle à D sachant C [Schmidhuber, 2007]. La beauté est directement mesurée par le gain en compression obtenu par l'observateur pour passer de D à D' . La beauté est donc d'autant plus grande que l'on dispose des règles pour interpréter le schéma visuel qui nous est proposé et en faire une description simple au regard de nos connaissances.

Ce schéma semble adapté pour décider, en présence d'un grand nombre de stimuli $D_1, D_2, D_3...$ lequel nous préférons, mais il ne semble pas à même de nous guider dans un choix d'action devant des stimuli très différents. J. Schmidhuber, propose, dans cette situation, de définir la démarche esthétique comme celle qui, conduit à un progrès dans la capacité de compression. Dans le but de modéliser cette démarche d'apprentissage, il a recours à la théorie de Kolmogorov pour caractériser la complexité du compresseur, ou plus précisément la variation de complexité du compresseur au cours du temps, par exemple dans une expérience d'apprentissage par renforcement. En l'absence de toute autre récompense extérieure, le comportement optimal consiste alors à optimiser ce qu'il appelle la récompense de "curiosité" qui s'exprime en fonction des connaissances C_t et C_{t-1} que l'on aurait si l'on devait observer D_t en utilisant la connaissance de D'_t ou non. L'implémentation de cette stratégie - très théorique - utilise abondamment les schémas de prédiction de Solomonoff et les machines virtuelles de Gödel.

En pratique, la mise en œuvre de cette approche demeure très complexe et, comme dans le cas des expériences de D. Birkhoff, repose sur une détection *ad'hoc* des meilleurs schémas de codage au cas par cas (figure 5.3). De plus J. Schmidhuber concède lui-même que de nombreux schémas de représentation qui devraient conduire à des codage très simples ne fournissent pas des résultats remarquablement heureux. Le modèle de J. Schmidhuber semble trouver une bonne confirmation dans les études faites sur les visages moyens. Plusieurs travaux ont souligné que les "visages moyens" obtenus par exemple pour les travaux de reconnaissance faciale, bénéficiaient d'un très bon jugement esthétique. Or ce sont probablement ceux que nous apprenons par défaut au cours de notre vie.

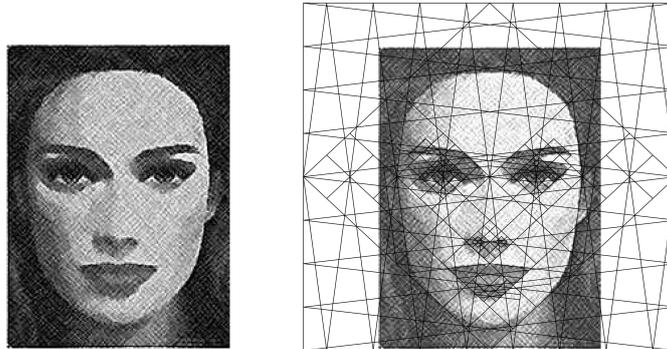


Figure 5.3 – L'image de gauche est belle car elle est simple. J. Schmidhuber propose de la déterminer à partir d'un schéma itératif très simple qui aboutit au réseau des lignes de l'image de droite. Les points les plus remarquables se trouvent représentés de façon très compacte sur ce maillage adapté (d'après [Schmidhuber, 2007]).

Chapitre 6

Le Saint-Graal de la Société du Numérique : l'intelligence artificielle

Les trois chapitres qui commencent ici seront consacrés aux méthodes dites par apprentissage automatique. Le principe de toutes les approches que nous verrons dans les deux chapitres suivants, repose sur l'hypothèse que les règles qui président à l'évaluation d'une photo peuvent être déduites de l'observation d'un grand nombre d'images qui ont été évaluées par des experts humains. Implicitement, on fait donc l'hypothèse que les règles sont universelles et intemporelles. Si l'on n'introduit pas d'autres règles, on est donc pleinement plongé dans le cadre objectiviste.

Nous verrons que deux familles successives de méthodes ont été proposées. Tout d'abord ont été développés des algorithmes de classification, assez simples, utilisant des informations sélectionnées dans l'image, le plus souvent choisies pour leur rôle dans le jugement esthétique. Puis les méthodes à base de réseau de neurones profonds les ont progressivement évincés à partir de 2015 ; ils ont fait porter tout l'effort sur la puissance du système de classification (le réseau profond) en laissant de côté l'étape de sélection de primitives. En contre-partie, ils ont réclamé de très grosses bases de données pour assurer un apprentissage de qualité.

Nous allons nous consacrer, dans le présent chapitre, à l'examen des sources d'images et des sources d'expertises communes à ces deux familles de travaux, ainsi qu'à l'ensemble des considérations qui sont partagées par ces techniques dites "par apprentissage"¹

6.1 L'Intelligence Artificielle (IA), pourquoi faire ?

Devant la difficulté des approches de psycho-physiologie ou de neuro-biologie de l'esthétique des images, de nombreux travaux ont été proposés afin d'aborder ce problème par les techniques informatiques. Ces travaux s'appuient sur la conjonction de trois progrès importants :

1. le développement de méthodes d'apprentissage (couramment dénommées *machine learning*),
2. l'accès aisé à de grandes quantités d'images mises à disposition sur le web, bases de données accompagnées d'éléments de jugement de qualité,
3. la disposition de processeurs puissants à même de traiter de très grosses quantités d'information.

1. La méthode de Schmidhuber, vue au chapitre précédent (paragraphe 5.4), se présente également comme une méthode d'IA et peut être vue comme une méthode d'apprentissage, mais l'objectif n'est pas d'apprendre des critères d'esthétique des images, mais pour chaque image, le comportement, dans le temps, des connaissances de l'observateur [Schmidhuber, 2007].

Nous séparerons les méthodes d'apprentissage intensif utilisées pour les études d'esthétique en deux familles principales :

- La première (chapitre 7) regroupe les **classifieurs fonctionnant sur des primitives extraites des images**. Particulièrement utilisées ces dernières années, les méthodes à base de machine SVM (*Support Vector Machine*), de mise en œuvre assez simple, conduisent à des temps d'apprentissage raisonnables avec des bases d'apprentissage assez réduites (fort pouvoir de généralisation). Leurs performances dépendent beaucoup de la qualité des primitives qui sont utilisées. Elles sont souvent bonnes, mais il est assez difficile d'apprendre des relations subtiles (relations hiérarchiques, exceptions, classification sémantiquement complexe, etc.).
- Depuis quelques années on leur préfère des classifieurs de type **RNP "Réseaux de Neurones Profonds"** (*deep neural network*) (chapitre 8). Les RNP sont des classifieurs "de bout en bout", c'est-à-dire qu'ils prennent en entrée l'image elle-même et donnent en sortie la classe ou la mesure que l'on attend. Les RNP sont de mise en œuvre beaucoup plus complexe car ils enchaînent des couches successives (de 5 à 15) de traitements selon des architectures savamment agencées. Ils nécessitent également de très longues étapes d'apprentissage de façon à configurer chaque couche de classifieur (apprentissage des dépendances entre couches). Pour cela, il faut disposer de très nombreuses données d'apprentissage. En contrepartie, les RNP offrent des performances remarquables, même dans les cas très difficiles et nécessitent très peu d'expertise.

6.1.1 Les résultats attendus

Tous les travaux qui s'intéressent à l'esthétique des photographies n'ont pas le même objectif. Ainsi, certains ne cherchent qu'à séparer les images médiocres des bonnes photos (c'est-à-dire de qualité utilisable par exemple en tant qu'illustration). D'autres cherchent à distinguer les photos d'amateurs des photos de professionnels en vue par exemple d'une commercialisation. Certains cherchent à classer, sur un critère esthétique, un ensemble limité de photos pour constituer un résumé, un album ou un répertoire. C'est en particulier le cas des systèmes qui se proposent en assistance du photographe, pour décider du meilleur cadrage d'un ensemble de photos. Cela peut se faire à posteriori sur une station de travail, mais c'est de plus en plus proposé comme système embarqué [Lo et al., 2013, Schwarz et al., 2016] Enfin d'autres s'attachent à attribuer une note dans un référentiel universel (de 0 à 100, ou sur 5 niveaux).

Il est rare que les systèmes d'évaluation définissent clairement l'échelle de leur notation. Les échelles les plus utilisées (par exemple celle de DPChallenge² qui est la source de la base de données AVA la plus utilisée pour notre propos) se contentent généralement de définir l'intervalle de notation, laissant à l'utilisateur la liberté de fixer la gradation. Récemment cependant, un effort a été fait pour fixer la progression d'une échelle sur 4 niveaux [Fang and Zhang, 2017], distinguant :

1. le niveau débutant,
2. le niveau amateur,
3. le niveau semi-professionnel,
4. le niveau professionnel.

Chaque niveau, ainsi que des niveaux intermédiaires, sont succinctement illustrés par les commentaires qu'ont apportés des professionnels qui ont classé des photos à ce niveau (voir la figure 6.1).

Les auteurs de [Fang and Zhang, 2017] proposent de faire correspondre au niveau débutant les 15 % d'images de plus faible score de DPChallenge, à la classe 2 les 55 % d'images suivantes, à la classe 3 les 15 % d'images suivantes et de réserver la classe 4 aux seuls 15 % d'images de score DPChallenge le plus élevé.

Si la plupart des travaux se contentent d'attribuer une note aux images, certains cherchent à "expliquer" les facteurs de leur jugement en fournissant, par exemple, des notes partielles sur divers

2. Les instructions de vote de DPChallenge se réduisent à "rate entries on a scale of 1 to 10. A score of 1 is a "bad" photo, and a score of 10 is a "good" photo. (notez la photo sur une échelle de 1 à 10. Un score de 1 est une mauvaise photo et un score de 10 une belle photo.)"

Score	Description
1.0	<p>Beginner: Point-and-shoot without consideration for composition, lighting etc. (it can still be technically good, i.e., good focus, good exposure...)</p> <ul style="list-style-type: none"> - This photo has no attention to light, composition, or moment and there is no clear intent. Looks like a photo taken by accident or like someone just held the camera up and pressed the shutter button... it's not clear why anyone would take this photo or want to look at this photo. - This score represents failure to use the camera properly and/or create anything that resembles a passable photo. - These photographs clearly are point and shoot with no intention of a composition in mind. These photographs usually lack contrast or thought and have nothing interesting in their subject matter. - No artistic merit or intention and does not follow basic rules of composition, lighting or focus.
2.0	<p>Amateur: Good photo for people without a background in photography. Nothing stands out as embarrassing. But nothing artistic either. Lighting and composition are OK. The majority of population falls between 1 and 2.</p> <ul style="list-style-type: none"> - I'm simply bored when I look at this image and it creates a sense of indifference. The image does not engage me at all and there is nothing interesting in the photo. This could be a combination of a poorly-lit image of a boring, cliched subject, and/or, an unsettled composition with poor tonal values. - This photo has a clear subject and things like composition and moment are ok but not great. You can tell that this person saw something of interest and they documented it, but they are still not thinking like a photographer because there is no attention to light, which is what photography is all about. - Pure point-and-shoot. These photos show us what is in front of the person taking the photograph. There is a passing sense for composing the elements within the frame. The angle is almost always at standing eye level. However, many of these images are acceptable to look at, and will often show interesting landscapes. The key point of classification here, is that the image does not exhibit professional photographic skill. A point-and-shoot photographer can still make a nice image if they are standing in front of beautiful landscape. - These photographs have some degree of thought put into the composition, with a strong image in mind. The image may have a clear message, but they miss the mark by either forgetting about contrast or composition. These are images that would usually fall into a camera phone photograph category. - Intention in composition, lighting or framing, but poorly executed. Possible bad editing, out of focus, fragmenting, pixelation or poor quality.
2.5	<p>Something artistic obviously presents in this photo. However, the attempt in this photo, content or editing, can not be called successful.</p> <ul style="list-style-type: none"> - The photo feels like it was taken, not made. Though the image is clear, there may be an attempt at composition without resolve, there is no point of focus, and/or the tonal range may be only in the mid-tones making it flat. The light has not been considered to best capture the subject, and/or there are distracting objects in the photo that keep it from the 3, 3.5, and 4 caliber. - You can tell in this photo, that the person is paying attention to light! They are starting to think like a photographer. Things like composition, moment and subject are an improvement but still lacking. There is attention to the way that light interacts with a subject or environment, even though the usage of light might not be very good. - This rating marks a step in the right direction beyond point-and-shoot photography. There is an effort here to create a better photo (interesting angle, composition, use of silhouette, compelling lighting, etc...) However, the image still does not fully add up to a well-made photograph. (These are the sort of images you'll see in a Photo-1 class. Good effort and intentions, but more skill needs to be applied). - These photographs have the intention of a good photograph but are missing out on many of the key elements of a professional photograph. This usually means lack of focus or composition. The photo may have a beautiful image but is cropping out a person. The photograph may have a beautiful mountainside, but the entire image is not straight. - Average photo, not good or bad. Follows rules of composition with lighting and framing, but not particularly well executed.
3.0	<p>Semi-Pro: One is on a path to become a professional photographer!</p> <ul style="list-style-type: none"> - I feel this is the critical break-point. The image is good with an effort to make/capture the shot, but the photo reveals a skill-level of one who does not have a lot of experience making great images and may employ the use of Photoshop in an attempt to enhance the photo. This is what I call over-cooking the image. It cheapens the photo and is a dead giveaway of an exploring amateur. It is almost as if the photographer is trying too hard. The photo may also be technically and aesthetically missing an element that would make it a 3.5. I would expect to find many of this caliber of imagery in a local art fair. - This photo has attention to light, a clear subject, good composition and a clear intent but more than a few factors are still lacking. Usage of light is better, but not great. the moment is a little awkward and the subject is boring. - This is a good photograph that works. The general approach has created a worthwhile landscape. This often includes at least one professional strategy that brings the image together (strong composition, depth-of-field, interesting angle, compelling lighting, etc...) - These photographs show a strong understanding of imagery and composition with a clear intention. These photographs fall short when it comes to a subject matter that defines a perfect photograph. - Above average image with clear thought, focus and framing put into it.
3.5	<ul style="list-style-type: none"> - The image is better than most, but has been done before in a more complete way. Usually the subject is amazing but the lighting could be better at a different time of day, or has any combination of great and slightly sub-great components. Its almost a 4, but I reserve 4s for only the best. - This photo has excellent use of light, a clear subject, a clear intent and almost all of the characteristics of a professionally crafted photo but there is just one factor that's off. Either the subject is boring, the moment is a little awkward, or the composition is a little messy. - This can be a tricky rating. For this, I often ask "what could have been done here that would make this photo even better, and worthy of a 4 rating?" In that sense, I use 3.5 to mark down from 4. Maybe the photographer oversaturated a perfectly good landscape, maybe they collided some elements within the frame (Ansel Adams often mentioned this). Or maybe the image just needs one more element, something that a professional would be mindful of. - These photographs have beautiful imagery but do not have the focus or the perfect composition that make a photograph truly professional. These are usually almost perfect photographs but are missing out on the techniques that make a photograph stand out as a perfect image. - Great image with purposeful depth of field and framing clearly taken by someone with photographic knowledge.
4.0	<p>Pro: photos you think deserved to be called taken by a professional.</p> <ul style="list-style-type: none"> - This photo was made, not taken. Everything in the image is working together to the sum of a great image. Without question this image was made by a skilled craftsman, one who is technically fluent, environmentally aware, has good timing and/or patience, is in command of post production and does not use cliched and overused filters, and offers a controlled composition that has a relationship with the subject. Anyone who sees this image would consider it professional. - This photo was created by someone who has studied photography and refined their craft. There is a great moment/interesting subject in great light. There is meaningful interaction between light and subject. Excellent use of composition. The moment is just right, and you can clearly see the photographer's intent. - This is a well-made professional photograph which exhibits experience, technical know-how, and above all else - a sense for the strategies which go into making strong landscape imagery. - These photographs are clearly shot by a professional with a precise composition in mind. There is a strong contrast of darks and lights. These photographs use techniques that show a strong understanding of their camera equipment. - Excellent image, reserved only for the best images with well thought out intentional and dynamic compositions, good lighting with balance in colors and tones and purposefully in or out of focus.

Figure 6.1 – *Explicitation de l'échelle à 4 niveaux de Fang et Zhang (d'après [Fang and Zhang, 2017]). Chaque niveau est illustré par un ensemble de propriétés formelles qu'un professionnel a attaché à l'observation d'une photographie particulière. Le niveau 4 n'est attribué qu'à un tout petit nombre de clichés de qualité esthétique exceptionnelle. Remarquons que les commentaires font fréquemment référence au "professionalisme" de l'auteur, un critère qui est probablement plus pris à son sens symbolique que strictement social.*

critères : *composition, couleur, contraste, etc.* ou des cartes qui indiquent les zones où les critères ont été détectés.

Enfin quelques travaux s'appuient sur les informations complémentaires que l'on peut trouver dans les bases de données sur la thématique dont relève l'image, par exemple *portrait, paysages* . . . de

façon à adapter la note en privilégiant certains critères. Lorsque la thématique n'est pas donnée, il peut arriver que l'on demande à l'algorithme de la déterminer avant de procéder à l'évaluation. Enfin dans un petit nombre d'études, on associe à l'image des descriptifs d'ambiance, liées à des émotions : *bonheur, effroi, surprise* . . . pour aider à l'évaluation. Ces émotions sont alors données par des photo-interprètes quoique, de plus en plus, des algorithmes se proposent de les estimer lorsque l'image comporte par exemple des visages en gros plan.

6.1.2 Les expertises

Les méthodes de *machine learning* ont recours à des experts lors de la phase initiale d'apprentissage qui permet de fixer les nombreux paramètres des algorithmes. Le choix de ces experts est crucial puisque c'est leur façon de décider que cherchent à imiter tous les algorithmes. Contrairement à ce qui se faisait pour la conception d'autres outils de décision automatique (comme les systèmes experts, maintenant obsolètes), les experts ne participent pas directement à la construction des méthodes de *machine learning*. Ces nouvelles méthodes ne nécessitent que la connaissance du résultat de l'expertise sur une grande variété de cas et ignorent tout de la démarche qui a conduit à cette expertise.

De grandes différences existent entre les études sur ce point. Visitons les rapidement.

Les expertises les plus reconnues sont issues du **domaine professionnel de la photographie** : des critiques d'art, des journalistes, des photographes professionnels, des jurys de concours photographiques. Ces expertises sont disponibles sur les sites spécialisés, dans les revues en ligne, mais aussi à l'occasion de compétitions souvent largement médiatisées. Si les avis ainsi exprimés sont généralement peu contestables, ils nécessitent souvent une étape d'interprétation car ils s'expriment rarement sous une forme immédiatement exploitable (on souhaiterait idéalement un avis sous forme de note dans un référentiel absolu). Cette étape est très réductrice et il n'est pas clair qu'elle ne dénature pas exagérément ces expertises. De plus, ces avis peuvent refléter des partis pris d'école ou de tendance, soit parce que l'expert en est imprégné, soit parce que le cadre de la compétition le réclame. De plus, ainsi que nous l'avons dit en section 1.1 sur les diverses raisons de regarder une photographie, il est parfois délicat de faire la part de ce qui provient, dans un commentaire élogieux, des seuls critères esthétiques. Notons en particulier que la photographie, suivant la tendance générale de l'art au *XX^e* siècle a, aujourd'hui, tendance à privilégier le "pouvoir émotionnel" sur le "plaisir esthétique" comme nous l'avons exprimé en section 1.2, et les avis des experts reflètent souvent cette tendance (voir la note 8 du paragraphe 1.1).

Moins rares que les avis des experts, mais aussi plus discutés sont les avis des **amateurs éclairés** qui sont généralement réunis sur des sites spécialisés comme DPChallenge. Leur très grand nombre permet de conduire des études statistiques sérieuses. La grande diversité des origines des "experts" permet d'assurer une meilleure couverture des opinions sur des tranches d'âge, des origines socio-professionnelles, des contextes géographiques et culturels différents. Les avis sont généralement un peu argumentés mais se traduisent aisément en score sans risque d'altérer l'avis exprimé. Lorsque les évaluations sont faites lors de "challenges", on doit cependant souligner tous les biais des votes ouverts (comportements de groupe, bourrage d'urnes, etc.) [Reagle, 2013].

Plusieurs études se sont appuyées sur de **petites cohortes d'experts** constituées généralement d'étudiants en beaux-arts ou en traitement des images, entraînés spécifiquement pour évaluer selon des protocoles bien établis des bases d'images préparées. Ces travaux ne garantissent ni la qualité des avis des experts professionnels, ni l'étendue des résultats des amateurs éclairés, mais ils assurent une homogénéité et une cohérence des résultats auxquelles les travaux précédents ne peuvent prétendre.

Enfin quelques études se fondent sur des "**non-experts**" en prenant des avis, souvent binaires (*j'aime* ou *je n'aime pas*), directement sur les réseaux sociaux généralistes ou spécialisés pour la diffusion de photos (Instagram, Flickr). Une autre tendance très en vogue est de faire appel aux internautes à l'occasion de petites compétitions lors de campagnes de *crowdsourcing*. Dans les deux cas, on essaie de déterminer un avis "moyen" parmi la population en réduisant notablement les attentes sur les critères d'esthétique ainsi dégagés.

6.2 Les bases de données

Il est très facile de trouver beaucoup d’images sur le web mais il est beaucoup plus difficile de trouver les informations associées qui sont indispensables à l’apprentissage, et en l’occurrence des informations fiables sur la qualité esthétique des photos. Les bases de données utilisées font donc souvent l’objet de compromis entre des avis très peu fiables mais très nombreux, issus du grand public, et des avis, précieux mais très rares, issus de professionnels dans la photo d’art (voir table 6.2).

Les sites professionnels donnent d’excellentes sources de photographie, de très haute qualité, mais pour la difficulté que nous avons soulignée ci-dessus d’exploiter les expertises, ils sont finalement peu utilisés directement pour constituer des bases volumineuses. Leur mérite principal est de fournir d’excellents exemples associés à des commentaires très instructifs.

Les sites d’amateurs éclairés sont la source première d’images. Ils offrent de très grandes quantités de photos, souvent de bonne qualité. Ils permettent d’éviter l’écueil de la “tentation artistique” de façon à privilégier les qualités esthétiques, ils ne sont cependant pas à l’abri de critiques. En particulier, dans un mouvement que l’on observe bien pour tous les réseaux sociaux, ils ont tendance à privilégier des comportements “de clans”. Par exemple, on ne peut manquer d’être frappé sur certains sites par l’importance des effets spéciaux (matériels ou logiciels) mis en œuvre pour “améliorer” l’image : accroissement du piqué, du contraste, filtrage du bruit, saturation des couleurs, distorsions géométriques, etc., reflétant une esthétique véritablement technophile bien loin des critères classiques des pères de la photographie (on se reportera pour exemple à la figure 4.1 à gauche). Enfin, les jugements portés ne sont pas à l’abri de critiques liés à la nature informelle de ces consultations [Reagle, 2013].

Les bases de données généralistes d’images, comme Flickr, Tumblr, Picasa, Instagram ou Image-Net, sont également une source permettant de disposer d’une très large variété d’images, mais leur qualité est parfois médiocre et les éléments associés qui permettraient de servir d’expertise ne reposent le plus souvent pas sur des critères esthétiques. Ces sources “primaires” ne sont pas toujours utilisées en l’état pour alimenter les études sur l’esthétique. Elles sont parfois filtrées pour fournir les bases de données opérationnelles de ces études. Présentons ces bases opérationnelles ici :

- **Flickr**³ : quoique source primaire d’images, Flickr a aussi été utilisée comme base de données pour des études sur l’esthétique, à partir de la notion d’attrait (*attractiveness*) qui mêle de façon un peu confuse des critères de beauté et de popularité [San Pedro and Siersdorfer, 2009, San Pedro et al., 2012]. L’attrait est mesuré par le décompte des annotations faites par les internautes : “*attractive*” ou “*unattractive*”.
- **Photo.net**⁴ : c’est un site d’amateurs éclairés et de professionnels de la photo qui regroupe probablement plus d’un million de photos de bonne qualité, classées en 38 catégories. Les photos reçoivent des notes de 1 à 7 selon les deux critères d’esthétique et d’originalité (voir figure 7.2 à gauche). L’accès aux photos peut se faire par ordre de note moyenne décroissante.
- **DPChallenge**⁵ est aussi un site d’amateurs éclairés et de professionnels. Il comporte plus de 650 000 photos généralement de bonne qualité. Les photos peuvent être retouchées et le sont parfois assez lourdement. DPChallenge fait une place très importante aux concours photos jugés par les internautes. Les notes vont de 1 à 10. Les meilleures photo reçoivent souvent plus de 200 notes. Les très nombreux concours sont identifiés par un thème qui peut servir d’identifiant (près de 1000 thèmes sont aujourd’hui disponibles dans les archives). DPChallenge est la base primaire essentielle des études sur l’esthétique ; elle a donné naissance en particulier à la base AVA, qui est la plus utilisée. Outre le grand nombre d’images de bonne qualité, on apprécie qu’elle offre un large éventail de notes, ainsi que des commentaires littéraires sur de nombreuses images. Des études se sont attachées à bien connaître les propriétés statistiques des appréciations portées sur les images de DPChallenge.

3. Flickr : url=*http://www.flickr.com/*.

4. Photo.net : url=*http://www.photo.net*.

5. DPChallenge : url=*http://www.dpchallenge.com*.

La répartition des notes est pratiquement gaussienne autour d'une moyenne un peu supérieure à 5 (voir figure 6.2 à gauche). L'écart-type des notes est lui-même assez gaussien. Les commentaires sont répartis de façon plutôt poissonnienne, avec un mode au voisinage de 15 commentaires par photo (voir la figure 6.2). Si l'on rapproche les statistiques de DPChallenge de la notation de Fang [Fang and Zhang, 2017] vue plus haut, on attribue le niveau 1 de Fang aux photos dont la note sur DPChallenge est inférieure à 4,67, le niveau 2 aux photos dont la note est comprise entre 4,67 et 5,82, au niveau 3 celles comprises entre 5,82 et 6,18, la note 4 étant réservée à celles au dessus de 6,18.

Si l'on rapproche les commentaires et les notes, on constate que la dispersion des notes pour une même photo est forte en présence de termes qui expriment soit l'originalité de la photo, soit une forte personnalisation de l'interprétation [Kim et al., 2015]. Elle reflète assez bien ce que certains auteurs appellent la "non-conventionalité" de la photo, c'est-à-dire sa capacité à sortir des canons classiques [Talebi and Milanfar, 2017].

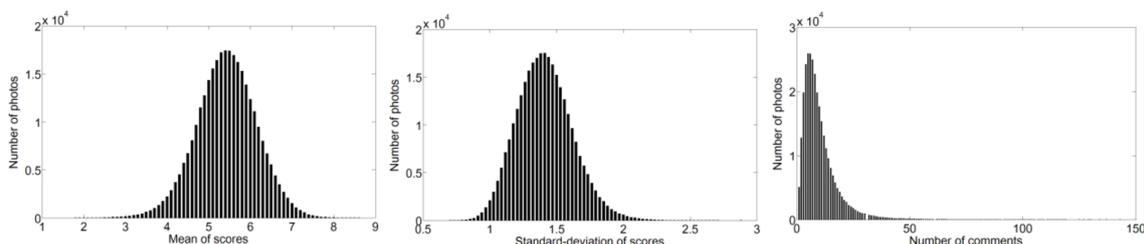


Figure 6.2 – *Dispersion des notes moyennes attribuées à chaque photo de la base de données DPChallenge (d'après [Kim et al., 2015]). A gauche, distribution sur toutes les photos des moyennes des notes attribuées à chaque photo (c'est pratiquement une gaussienne de moyenne 5,43 et d'écart-type 0,73) ; au centre distribution des variances des notes attribuées à chaque photo (c'est aussi pratiquement une gaussienne de moyenne 1,41 et de variance 0,21) ; à droite répartition du nombre de commentaires littéraires par photo ; c'est plutôt une loi poissonnienne de paramètre 15.*

- **CUHK-DB**⁶ développée par l'université de Hong-Kong, elle est issue de DPChallenge en vue de soutenir les études sur l'esthétique. Elle a été obtenue en sélectionnant des images ayant reçu une note soit appartenant aux 10 % meilleures soit aux 10 % moins bonnes. Les deux classes sont donc très peu ambiguës et les images "moyennes" ne participent pas à l'apprentissage.
- **ImageCLEF**⁷ est issue de la base de données MyFlickr issue elle-même de Flickr (ce sont donc des images ordinaires sans qualité esthétique particulière). Elle a été développée pour tester les algorithmes d'indexation et de fouille d'images. Elle comporte un million de photos, avec des annotations sur le contenu (44 tags sémantiques) ainsi que sur les émotions associées aux photos.
- **SUN**⁸ est une base de données d'images du MIT, comportant 130 000 images indexées, partagées en 900 catégories de scènes. Certains objets sont segmentés, appartenant à plus de 4 000 catégories d'objets. Ces images n'ont pas de caractéristiques particulières d'esthétique.
- **AVA**⁹ *Aesthetic Visual Analysis* est une base de données d'images spécialement développée pour les études d'esthétique. Elle est devenue progressivement la source principale des travaux utilisant des réseaux profonds. Elle est issue de DPChallenge et rassemble 250 000 images, regroupées en 60 catégories, avec une abondante indexation en scores esthétiques d'une part (plus de 200 notes en moyenne par photo), en style photographique (14), en contenu (44 tags sémantiques) ainsi qu'en émotions associées aux photos d'autre-part [Murray et al., 2012]. Elle hérite donc des propriétés statistiques de la base DPChallenge [Talebi and Milanfar, 2017], et en particulier de la correspondance des notations avec les travaux de Fang, rapportée dans le tableau 6.1.

On doit s'interroger sur la qualité des images de la base AVA. Nous l'avons dit, AVA est déduite de DPChallenge qui offre des images de très bonne qualité, à haute résolution (généralement

6. CUHK-DB : url=<http://mmlab.ie.cuhk.edu.hk/CUHKPQ/Dataset.htm>.

7. ImageCLEF : url=<http://www.imageclef.org/>.

8. SUN : url=<https://groups.csail.mit.edu/vision/SUN/>.

9. AVA url=<http://www.computervisiononline.com/dataset/1105138637>.

Qualité selon Fang	classes de Fang	Notation AVA
débutant	1	$\leq 4,67$
amateur	2	$\in]4,67 ; 5,82]$
semi-professional	3	$\in]5,82 ; 6,19]$
professional	4	$> 6,19$

Table 6.1 – Correspondance entre les classes de Fang [Fang and Zhang, 2017] et la notation entre 0 et 10 d’AVA (ou de de DPChallenge), [Murray et al., 2012].

de l’ordre de 4000 pixels par ligne). Il n’est cependant pas clair que toutes les études menées avec AVA bénéficient de cette qualité. En effet, la base disponible sur le site AVA est constituée d’images réduites à 640 pixels dans leur dimension la plus grande. Certes il est possible d’accéder à l’image source par un lien qui est fourni dans AVA, mais il n’est pas clair que cette opération de téléchargement de chaque image à haute résolution soit faite systématiquement par les études qui s’appuient sur AVA, et ce d’autant plus que DPChallenge met en place des mécanisme assez efficace de protection des images sources.

Deux types d’adaptations ont été faites avec AVA de façon à créer des ensembles d’apprentissage où les images sont nettement séparées sur des critères esthétiques. D’une part on a séparé un ensemble d’images, réputées belles, d’un ensemble d’images, réputées moins belles, par un seuil $\pm\delta$ autour de la valeur moyenne 5,5 et défini ainsi des ensembles d’apprentissages divers pour des δ de 0, 1 ou 2. D’autre part, on a réduit la base AVA aux seules 10 % d’images les plus belles et 10 % d’images les moins belles. Cette nouvelle base est appelée AVA2 et ne comporte donc plus que 50 000 images environ.

Style	Number	Style	Number
Complementary Colors	949	Duotones	1,301
High Dynamic Range	396	Image Grain	840
Light on White	1,199	Long Exposure	845
Macro	1,698	Motion Blur	609
Negative Image	959	Rule of Thirds	1,031
Shallow DOF	710	Silhouettes	1,389
Soft Focus	1,479	Vanishing Point	674

Figure 6.3 – Les 14 attributs de style utilisés dans la base de données AVA et leur nombre d’occurrences, d’après [Wang et al., 2016]

- **AADB**¹⁰ La base *Aesthetics with Attributes Database* est issue du site Flickr. Elle contient 10 000 images annotées par 5 opérateurs qui leur ont attribué une note d’esthétique (sur 5 niveaux) ainsi qu’une note pour chacun de 11 attributs descriptifs de qualité issus d’une liste fournie par des professionnels de la photo (*intérêt du contenu, éclairage, harmonie des couleurs*, etc.). La moyenne des 5 notes d’esthétique est prise comme note de référence pour l’esthétique, mais on peut aussi suivre un opérateur unique de façon à conserver éventuellement une esthétique cohérente [Kong et al., 2016].
- **BEAUTY**¹¹ est une base de 15 000 images issues de Flickr, mêlant des images de qualité très variée. L’indexation de ces images en fonction de leur qualité est confiée à la communauté des internautes : elle se fait par *crowdsourcing*. Des précautions sont prises pour éviter des dispersions trop grandes des avis, ainsi que des biais dans les résultats statistiques. Pour le premier point les seuls internautes mis à contribution sont issus d’un nombre restreint de pays à forte homogénéité culturelle et leurs résultats sont a posteriori filtrés pour écarter les comportements par trop singuliers. Pour le second point, le stock des images traitées est partagé en trois grandes familles de qualités différentes et les représentations statistiques des échantillons issus de ces familles sont notablement différentes. Enfin des critères différents sont retenus pour décrire les quatre classes : “personnes”, “nature”, “animaux” et “ville” [Schifanella et al., 2015].

10. AADB url=<http://www.ics.uci.edu/skong2/aesthetics.html>.

11. BEAUTY url=<http://www.di.unito.it/~schifane/beauty-icvsm15/>.

critères	Photo .net	DP Challenge	Image CLEF	BEAUTY	AADB	Redi	AVA	Uni Tübingen
taille ($\times 1000$)	> 1000	650	1000	15	10	100	250	380
qualité esthétique	haute	haute	faible	faible	faible	faible	haute (?)	faible
score esthétique	7	10		3	5	4	2	continu
classes sémantiques	38	\neq thèmes	44	4	non	2	44	non
étiquettes de style	non	non	non	non	11	non	14	non
annotation	non	oui	oui	non	oui	oui	oui	non
origine			Flickr	Flickr	Flickr	web	DP Chal.	Flickr

Table 6.2 – Quelques propriétés des bases de données les plus utilisées pour l'étude esthétique des images. La taille est exprimée en milliers d'images. La qualité esthétique est considérée comme "haute" si les images proviennent de professionnels ou d'amateurs éclairés, comme "faible" si elles sont issues de réseaux sociaux. Le nombre associé au score esthétique indique le nombre de niveaux du résultat de l'évaluation. Les 3 lignes suivantes indiquent quel type d'information est associé à chaque image : le nombre de classes sémantiques, l'existence de caractéristique de style, la disponibilité d'une autre annotation (par exemple les émotions associées).

- **La base de M. Redi et al.** comporte également des images (100 000) de toutes qualités, mais issues du web à partir de 8 000 requêtes en tirant aléatoirement 20 images par requête (cela évite de n'utiliser qu'une source potentiellement biaisée). La base comporte ainsi 85 % de photos et 15 % de graphiques, dessins, images de synthèse, etc. Ces images sont annotées par des professionnels selon deux indices : la *qualité esthétique* (sur 4 niveaux) et le *type* ("photo" ou "non-photo") [Redi et al., 2017]. Elle fait également l'objet d'une évaluation publique par le Turc Mécanique d'Amazon, cette évaluation se montrant plus générale, mais moins fiable.
- **La base de l'Université de Tübingen** [Schwarz et al., 2016] est constituée à partir des images de Flickr et comporte 380 000 images. L'annotation varie de 0 à 1. Elle est obtenue en prenant en compte le nombre de vues (\mathcal{V}_i) (c'est-à-dire de clicks sur le lien de l'image) et du nombre d'avis favorables (\mathcal{F}_i) (mesuré par le nombre de clicks sur la case *favor*) par la formule :

$$S(i) = \alpha \frac{\log(\mathcal{F}_i)}{\log(\mathcal{V}_i)} \quad (6.1)$$

où α permet de fixer à 1 la mesure de la meilleure image. Seules les images telles que \mathcal{F}_i est supérieur à 1 sont gardées dans le décompte. La mesure de qualité est donc celle d'observateurs néophytes mais très nombreux (en moyenne 6 800 votes par photo). La distribution des votes est pratiquement plate entre 0 et 1.

- **La base de l'Université de Iéna : JenAesthetics**¹² La base de données constituée dans le groupe de Computer Vision de l'Université de Iéna [Amirshahi et al., 2013], comme la suivante que nous examinerons ci-dessous, n'est pas une base de données de photos, mais de peintures. Elle a cependant été constituée dans le but précis de comprendre les ressorts du jugement esthétique et de confronter les résultats trouvés à l'épreuve de techniques d'apprentissage automatique. En cela elle rejoint notre propos.

Elle est constituée de 1628 photographies de haute qualité d'œuvres d'art collectées à l'occasion du *Google Art Project*, couvrant tous les styles de peinture, des Primitifs à l'art contemporain. Chaque œuvre d'art est accompagnée d'une vingtaine d'évaluations conduites par un panel de 130 observateurs, non spécialistes de l'art, interrogés sur leur appréciation de cinq critères : l'esthétique, la beauté¹³, la couleur, la composition et le contenu. L'appréciation se fait à l'aide d'un curseur continu entre 0 et 1.

- **La base de données BAM! : Behance Artistic Media dataset**¹⁴ Cette base de données a été constituée à partir du site Behance¹⁵, un portefeuille de sites spécialisés dans la présentation de professionnels dans le domaine de la production multimédia. On y trouve donc des photos, mais également des dessins, des peintures et des vidéos qui ne relèvent pas directement de

12. JenAesthetics url=<http://www.inf-cv.uni-jena.de/en/jenaesthetics>.

13. Il n'est pas précisé comment les auteurs distinguent esthétique et beauté, et les résultats statistiques montrent que ces variables sont finalement fortement corrélées.

14. BAM! Behance Artistic Media dataset url=<https://bam-dataset.org/>.

15. Behance : url=<https://behance.net/>.

notre propos. Behance comporte plus de 65 millions de documents; BAM! en a sélectionné 1,9 millions, dans les catégories suivantes : images de synthèse, bandes dessinées et dessins graphiques et peintures. A chaque image sont attribuées des étiquettes exprimant l'émotion (calme, heureux, triste, inquiet) et la catégorie de l'objet (seuls sont retenues les scènes couvertes par les termes : bicyclette, oiseau, bâtiment, voiture, chat, chien, fleur, personne et arbre). Les annotations ont été obtenues en deux étapes : tout d'abord par *crowdsourcing* d'un nombre limité d'images, puis par classification automatique itérative [Wilber et al., 2017].

- **La base FACD** (*filter aesthetic comparison dataset*¹⁶) C'est une base tirée de AVA pour comparer les qualités esthétiques des divers traitements très souvent utilisés pour "améliorer" les photos déposées sur les sites comme Instagram, [Sun et al., 2017]. Elle est composée de 1260 images originales auxquelles ont été appliqués 22 traitements issus, soit des bibliothèques d'Instagram, soit de la boîte à outil GIMP. Ces traitements procèdent essentiellement à des changements de contraste et des colorisations. Chaque image est affectée à une catégorie parmi 8 (animal, flore, paysage, architecture, nourriture, portrait, scène urbaine, et nature morte), également peuplées. La base de données comporte également 42 240 paires d'images auxquelles sont associées des jugements de préférence attribués par des observateurs humains à l'aide du Turc Mécanique d'Amazon.
- Enfin, **des sites de vente d'œuvres d'art en ligne** : de nombreux sites sont spécialisés dans la vente d'œuvres d'art en ligne; ils offrent souvent, à côté de peintures ou de sculptures, des photographies (Artspace, Artsper, ArtPhoto, KazoArt, Pixopolitan, Singulart, Ugallery ... par exemple). Ils proposent ainsi des catalogues de quelques centaines à quelques milliers de photos qui se distinguent par leur qualité artistique affirmée. Ces sites regroupent les photos par catégories (paysages, portraits, natures mortes), ils mentionnent les auteurs et annotent souvent les documents de commentaires qui peuvent être utilisés comme méta-data dans des classifications. Les ressources offertes par ces collections sont par exemple discutées dans [Messina et al., 2018]. Notons que ces sites proposent souvent des œuvres beaucoup plus engagées artistiquement que les sites d'amateurs comme DPChallenge et font référence à des critères esthétiques que l'on pourrait qualifier de "savants" en ce qu'ils mobilisent une technique photographique très avancée.

16. FACD (*filter aesthetic comparison dataset*) url=<https://wtwilsonsun.github.io/FACD/>

Chapitre 7

Les méthodes de classification à base de primitives

Les photos artistiques sont des documents de grande dimension. Plus que les autres photographies, elles réclament qu'on les traite avec toute leur résolution et toute la dimension de leurs couleurs. Elles ne se prêtent pas bien à des opérations de sous-échantillonnage ou de codage. Les méthodes présentées ici vont donc choisir d'extraire de l'image un ensemble de primitives, censé retenir les principales qualités de la photo initiale. Les méthodes que nous verrons se distinguent par le choix des primitives retenues. Une approche employée dans le domaine de l'imagerie multimédia consiste à extraire des caractéristiques connues pour bien refléter les propriétés d'une grande diversité de photos : descriptif de l'histogramme de luminosité ou des plans couleurs, distribution spatiale des principales plages colorées, moments géométriques des régions lors d'une décomposition hiérarchique, indicateurs de textures, etc. D'autres approches s'appuient sur des représentations utilisées en vision par ordinateur pour retrouver des formes et des objets : liste de points caractéristiques (SIFT, SURF, Harris, etc.). On peut également s'appuyer sur les recommandations faites dans le domaine de la photo d'art (voir section 4), pour ne conserver que les primitives qui reflètent des critères exprimés de beauté. C'est alors ce que l'on dénomme couramment dans la littérature par les primitives "artisanales" (*hand-crafted*) qui visent à refléter les recommandations des experts. Dans le tableau présenté sur la figure 7.1 on observe un mélange de ces divers types de primitives.

Néanmoins, quelles que soient les primitives choisies, le schéma qu'adoptent ces travaux se fait selon deux étapes :

1. une étape d'apprentissage :
 - (a) une source d'images est identifiée, fournissant pour chaque image une évaluation de la qualité esthétique ;
 - (b) des primitives sont extraites de chacune des images de façon à mettre en avant des propriétés discriminantes des qualités esthétiques des images ;
 - (c) un algorithme d'apprentissage supervisé est mis en place qui détermine l'association judicieuse des primitives de façon à rendre compte du classement esthétique ;
2. une étape de notation d'une photo dont les qualités esthétiques sont inconnues :
 - (a) l'image inconnue est décomposée dans les mêmes primitives que celles utilisées lors de l'apprentissage ;
 - (b) un classifieur, identique à celui qui a été entraîné, mais dont les paramètres sont maintenant fixés, fournit une estimation de la qualité de l'image.

Name	Description
Brightness AVG and STD	(1) Average and standard deviation of the brightness, using the V channel in the HSV space.
Color Variance	(1) Variance of colors in the LAB space.
Contrast	(1) Width of the middle 96% mass of the histogram of the V channel in the HSV space.
#Edges and #Edges L, R, T, B, C	(1) We split the canny map into 16×16 blocks and we compute the number of blocks containing more than 10% of edges. We also compute this number on the left, right, top, bottom and center regions of the image.
Hue Count	(1) Approximation of the number of unique hues [18].
Saturation AVG and STD	(1) Average and standard deviation of the saturation.
Sharpness	(1) Variance of the Laplacian. [20]
Distance to the Center	(2) Distance of the salient region to the center of the image.
Rule of Thirds	(2) Shortest distance of the salient region to a power point.
Salient Hue, Brightness and Saturation	(2) Average hue, brightness and saturation of the salient region.
Salient Sharpness	(2) Sharpness of the salient region.
Salient Size	(2) Size of the salient region.
Salient LOC	(2) We split the image into nine equal parts, and compute the proportion of the salient region in each part. LOC can then take nine values: Top-Left, Middle-Left, Bottom-Right...
Color Difference	(3) Difference of colors in the LAB space between the salient object and the background.
Hue, Saturation and Brightness Difference	(3) Difference of hue, saturation and brightness between the salient region and the background.
Sharpness Difference	(3) Difference of sharpness between the salient region and the background.

Figure 7.1 – Quelques primitives couramment utilisées pour mesurer la beauté d'une photographie. Certaines sont des primitives génériques du traitement de l'image (comme la luminosité ou le contraste), d'autres relèvent plus de critères esthétiques (distance au centre, règle du tiers) (d'après [Simond et al., 2015]).

7.1 Primitives multimedia : le système Acquine et les travaux de R. Datta *et al.*

Ces travaux sont parmi les premiers à avoir abordé le problème de détermination des qualités esthétiques des photos par ordinateur [Datta et al., 2006, Datta et al., 2008, Datta and Wang, 2010]. Ils ont abouti à un service en ligne d'évaluation de la qualité d'une photographie : ACQUINE¹.

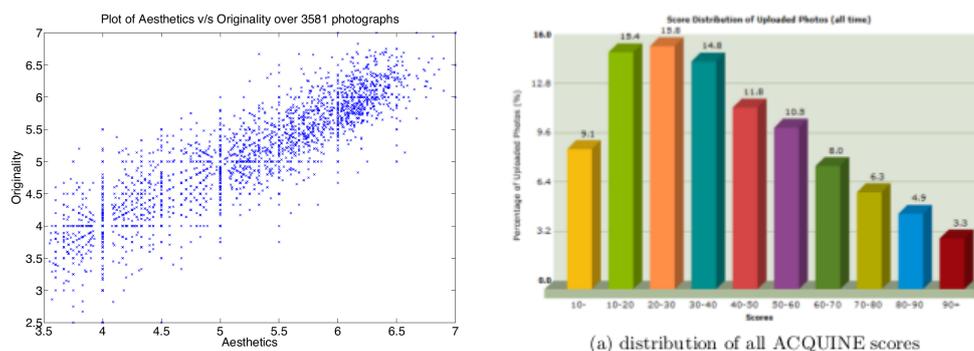


Figure 7.2 – A gauche, relation entre les notes d'esthétique et d'originalité dans la base *Photo.net* (d'après [Datta et al., 2006]). A droite distribution des scores d'Acquine sur l'ensemble des photos traitées (150 000 environ) : les très bonnes notes sont rares et les notes les plus probables sont faibles (d'après [Datta and Wang, 2010]).

R. Datta a choisi des photos issues du site *Photo.net*. Seules ont été considérées les images ayant reçu au moins deux évaluations supérieures à 6 sur le seul critère de l'esthétique. Ainsi environ 3 800 images ont été retenues.

1. Disponible à l'adresse : <http://infolab.stanford.edu/~wangz/project/imsearch/Aesthetics/MIR10/>.

Le résultat attendu de l'algorithme est en deux parties : une classification à deux classes : *low* et *high*, définies à partir d'images de notes inférieures à 4,2 d'une part et supérieures à 5,6 d'autre part, et une note individuelle de 0 à 100, croissante avec la qualité esthétique de l'image.

Les primitives extraites des images sont au nombre de 56 :

- le niveau de gris moyen,
- une mesure de coloration est déduite d'une quantification grossière de l'espace RVB en 64 cubes, puis d'un calcul de distance (dans l'espace Luv) entre cette image et une image de couleurs équi-réparties ;
- les valeurs moyennes de saturation et de teinte ;
- 3 mesures donnant l'accord à la règle du tiers dans les 3 composantes HSV ;
- 2 indicateurs de *familiarité* qui comparent l'image à un ensemble de référence et calculent la distance moyenne aux 20 ou 100 plus proches ;
- 12 termes de texture sont déduits de décompositions en ondelettes des composantes HSV ;
- la taille et le facteur d'aspect sont également retenus ;
- une décomposition par régions (classification par k -moyennes dans l'espace chromatique) ne retient que les 5 régions les plus importantes qui sont décrites par leur taille et leur point de couleur moyen. De plus 2 indicateurs expriment la distribution des teintes et leur complémentarité sur la roue des couleurs ;
- 3 indicateurs de profondeur de champ permettant de distinguer les images nettes sur fond flou ;
- enfin un facteur qui mesure la convexité et la régularité des formes.

Le classifieur utilisé est un SVM classique, entraîné par permutation croisée de l'ensemble d'apprentissage et de l'ensemble de test après sélection des primitives (30 sur 58 sont conservées). Enfin une régression linéaire sur l'expression polynomiale des facteurs retenus permet de déduire une formule approchée de l'esthétique de l'image à partir des seules valeurs des primitives.

Notons que le taux de bonne classification *low* et *high* est de l'ordre de 63 %, ce qui reste assez modeste, avec une variance de l'ordre de 0,5 sur la notation à 7 niveaux. Il est difficile de tirer des conclusions du choix des primitives par les algorithmes de sélection : arriveraient en premiers la description des régions, le niveau de gris moyen et la convexité des formes, mais ces conclusions ne sont pas vraiment convaincantes.

Le système a été mis en expérimentation en ligne sous le nom d'ACQUINE² qui a connu un certain succès chez les photographes amateurs (voir les résultats de la figure 7.2 à droite). Parmi les plus intéressants résultats, notons qu'un forum a été investi par les internautes qui ont cherché à retrouver les critères de notation du moteur d'inférence en soumettant des successions d'images de façon à avoir soit le score le plus élevé, soit le plus faible.

7.2 Contours et distance chromatique : l'approche de Y. Ke et al.

A peu près en même temps que les travaux précédents ont été proposés les travaux de Y. Ke [Ke et al., 2006] dont l'objectif est d'affecter toute photo à l'une des deux classes *Amateur* ou *Professionnel*. Notons tout de suite que d'autres critères entrent en jeu dans cette classification que les critères esthétiques, puisque certains éléments relèvent plus spécifiquement de ce que l'on appelle *la qualité* de l'image.

La base d'images utilisée est celle de *DPChallenge.com* ; seules les photos ayant reçu au moins 100 voix sont prises en compte et la moyenne des notes reçues leur est attribuée comme note unique. Les seules photos appartenant aux 10 % les mieux notées sont conservées pour former la classe P : *Professionnel* et les 10 % inférieurs pour la classe A : *Amateur*, constituant ainsi une base de données B de 6 000 photos (voir figure 7.3).

2. ACQUINE = *aesthetic quality inference engine* [Datta and Wang, 2010].

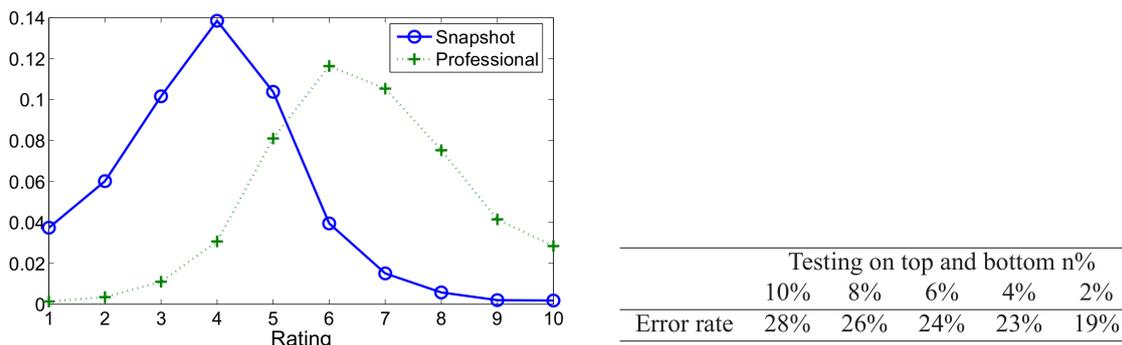


Figure 7.3 – *A gauche, les images P ou A ne sont pas totalement séparées dans les notes qu’elles reçoivent sur DPContest.com. A droite, en choisissant de plus en plus strictement (de gauche à droite) les photos sélectionnées comme P ou A dans la base de données, on améliore un peu les performances du classifieur, mais pas beaucoup, (d’après [Ke et al., 2006]).*

Pour être normalisées, les images sont ramenées à un format carré. Les primitives qui ont été utilisées relèvent du domaine du traitement de l’image :

- une carte des contours détectés par un laplacien est comparée à la carte moyenne des contours des deux répartitions P et A ; une boîte englobante de la majorité des contours est évaluée et rapportée à la taille de l’image (les “belles” images ont tendance à avoir leurs contours concentrés loin des bords) ;
- après une quantification des plans de couleurs, on recherche les n images de B dont les histogrammes sont les plus proches les uns des autres (aux k-plus-proches-voisins). Si n_P et n_A sont les images issues de P et de A, on utilise $n_P - n_A$ comme indicateur de qualité ;
- un indice de teinte est dérivé d’un décompte du nombre de couleurs dont la teinte est supérieure à 5 % de la dynamique des teintes ;
- un indice de contraste et un indice de brillance.

De ces primitives sont tirées cinq mesures de qualité qui sont combinées au sein d’un classifieur bayésien naïf, éventuellement suivi d’un classifieur Real-AdaBoost.

Les résultats obtenus par cette méthode restent modestes et les enseignements sur les primitives les plus importantes pour le jugement esthétique demeurent rares.

7.3 Objet et fond : l’approche de Y. Luo et X. Tang

Ces travaux sont rapportés dans [Luo and Tang, 2008]. Ils se mettent dans le prolongement des travaux de Y. Ke *et al.*, mais adoptent une stratégie fondée sur la distinction entre l’objet principal de la photo et son fond. La séparation s’appuie sur une hypothèse que l’on pourra mettre en question : le sujet est généralement clair et au point tandis que le fond est plus sombre et flou. L’algorithme de détection reprend donc ce principe et mesure localement diverses statistiques de l’intensité et de ses dérivées horizontales et verticales.

Les auteurs peuvent alors mettre en œuvre des critères différents sur l’objet et sur son fond, ainsi que sur le contraste qui les différencie. Ainsi le fond gagnera-t-il à être assez uniforme avec une palette chromatique réduite, tandis que l’objet bénéficiera d’une plus grande richesse harmonique mesurée dans l’espace HSV par des histogrammes marginaux. La qualité de la palette chromatique sera obtenue en comparant ces histogrammes marginaux à ceux des “belles” images de la base d’apprentissage les plus proches. Enfin des règles de disposition spatiale sont également mesurées, comme la règle du tiers comptée à partir du centre de gravité de l’objet.

La même base de données d’apprentissage issue de DPChallenge, modifiée par Y. Ke *et al.* est

utilisée et l'apprentissage se fait sur 3 000 photos de bonne qualité et 3 000 mauvaises photos. Le classement se fait soit par Bayes, soit par SVM suivi d'AdaBoost. Les résultats, sur les mêmes images, sont nettement supérieurs à ceux de Y. Ke *et al.* (figure 7.4). Les auteurs attribuent ce gain au mérite des primitives mieux adaptées à la photographie.

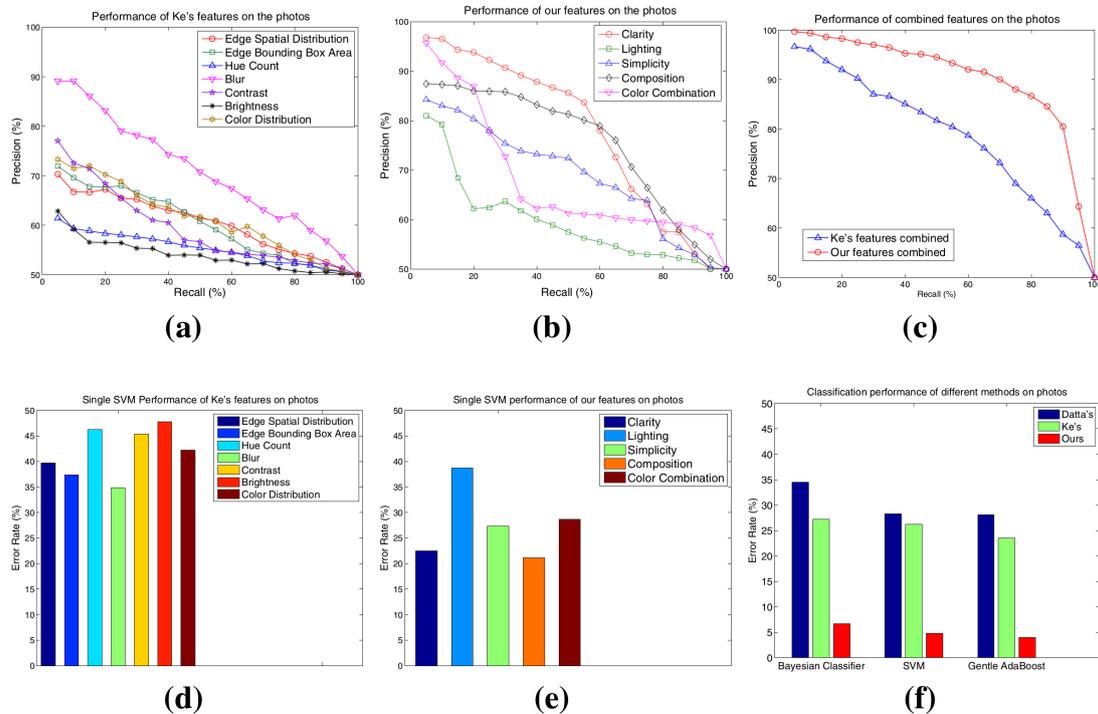


Figure 7.4 – *Comparaison des résultats (Précision en fonction du Rappel) de Y. Ke et de Y. Luo. Classification bayésienne (a) avec les primitives de Ke, (b) avec les primitives de Luo, (c) avec les primitives combinées; (d) et (e) rôle des primitives de Ke et de Luo dans la classification (pour Ke, c'est l'intensité qui est la grandeur dominante, pour Luo c'est l'éclairage); (f) performances des divers classificateurs (les primitives de Luo induit une plus faible erreur que celles de Ke ou de Datta, quel que soit le classificateur) (d'après [Luo and Tang, 2008]).*

7.4 Des primitives de haut niveau : l'approche de S. Dhar *et al.*

[Primitives de haut niveau] Ces travaux se placent également en parallèle de ceux de Y. Ke, mais les auteurs ont choisi d'utiliser des primitives de haut niveau, décrivant explicitement la composition de l'image, son contenu et l'illumination de façon interprétable par un interprète humain [Dhar et al., 2011]. La figure 7.5 indique de quelle façon ces primitives sont calculées à partir des primitives de Y. Ke. La détection de nombreuses primitives est complexe (détection des visages, des animaux, reconnaissance des types de scène - montagne, intérieur ...) et nécessitent la mise en œuvre de sous-systèmes de classification (17 SVM sont utilisés pour apprendre des tâches particulières), avec leurs propres bases d'apprentissage (par exemple un système pour chacun des 10 types d'animaux identifiés). Ils retiennent ainsi 26 attributs.

L'étude montre ensuite les bénéfices de cette lourde approche lors de deux tâches particulières : l'une (qui nous concerne) sur l'évaluation de la qualité esthétique, l'autre sur l'évaluation de l'intérêt de l'image. Donnons ici les résultats sur la seule tâche d'évaluation esthétique. Ils sont regroupés sur

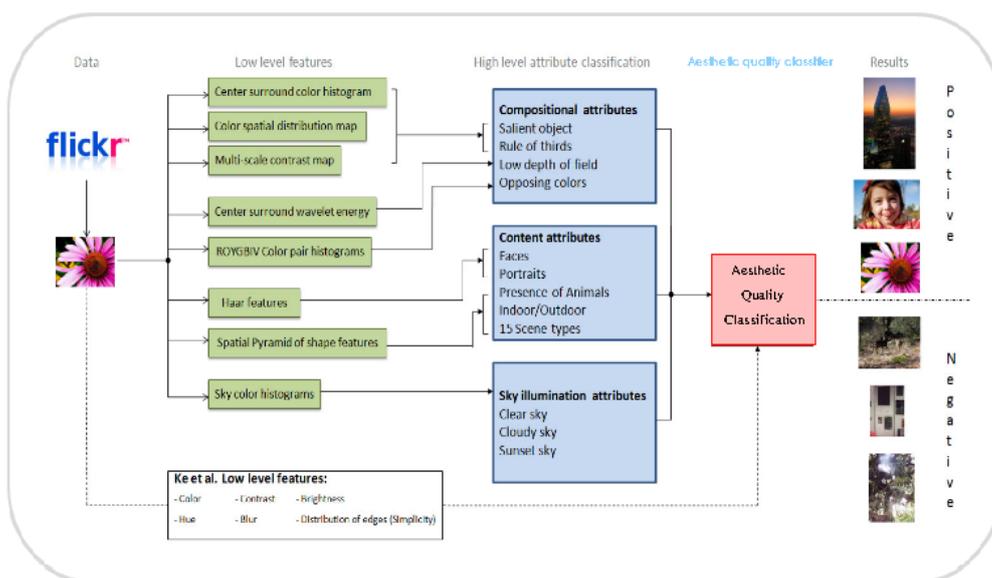


Figure 7.5 – Le schéma de calcul des primitives de haut niveau de S. Dhar et al. utilise les détections faites au bas niveau par le pré-traitement de Y. Ke (le schéma ci-dessus décrit l'évaluation de "l'intérêt" de l'image, mais un schéma identique permet d'évaluer "l'esthétique". Il ne diffère pas de celui-ci dans son architecture mais subit un entraînement différent de façon que la dernière couche de neurones fournisse un score d'esthétique et non d'intérêt (d'après [Dhar et al., 2011]).

la figure 7.6. Les auteurs préconisent l'emploi d'un classifieur SVM qui donne les meilleurs résultats sur une base 16 000 images issues de DPChallenge.

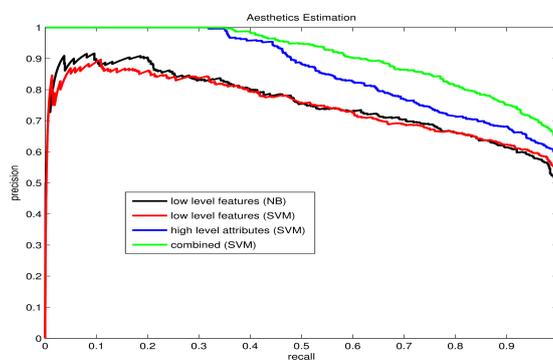


Figure 7.6 – Comparaison des résultats de l'approche de Y. Ke et de l'approche de Y. Dhar en termes de Rappel et Précision. Les courbes marquées low level ont été obtenues avec les primitives de Y. Ke. La courbe marquée high level utilise les primitives de Dahr et la courbe marquée combined utilise toutes les primitives. L'expérience avec les primitives de bas niveau a été conduite avec le classifieur recommandé par Y. Ke (Bayes naïf, noté NB), ou avec un classifieur SVM (d'après [Dhar et al., 2011]).

7.5 La bibliothèque de primitives esthétiques de H.-H. Su et al.

Dans les travaux présentés dans [Su et al., 2012], on se propose d'apprendre un catalogue de configurations spécifiques de photographies esthétiques à partir de belles photographies de DPChallenge.

Pour cela on décompose les photos selon des mailles fixes, déterminées une fois pour toutes, reflétant les principales échelles de composition des photos (figure 7.7). Sur ces maillages sont mesurées des statistiques particulières de certaines primitives (moyenne et contraste entre les mesures faites sur les sites noirs et les sites blancs de la figure 7.7). Ces primitives concernent la couleur (RVB ou HSV), la distribution des contours (détecteur HOG (*histogram of oriented gradient* [Dalal and Triggs, 2005])), une mesure de texture dénommée LBP et une mesure de salience. Le calcul des diverses configurations possibles, par exemple dans la subdivision (d) de 6×6 , est obtenu par un tri des 36 fenêtres de mesure, permettant d'obtenir très rapidement les configurations qui optimisent les contrastes et leur distance aux fenêtres de la bibliothèque.

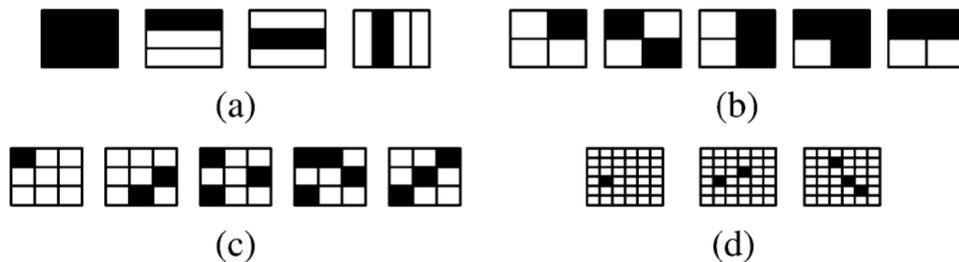


Figure 7.7 – Les fenêtres d'analyse proposées par H.-H. Su et al., qui fourniront la bibliothèque des primitives esthétiques. L'image est divisée selon les 4 configurations décrites dans les sous-figures (a) à (d). Pour chaque configuration sont calculées les propriétés particulières des zones noires par rapport aux zones blanches (d'après [Su et al., 2012]).

Cette technique se montre particulièrement efficace pour proposer un recadrage de l'image autour des zones les plus intéressantes, ce qui est l'objectif principal de ce travail.

7.6 Le système OSCAR de L. Yao *et al.*

Le système OSCAR (*on-site composition and aesthetic feedback through exemplars*) [Yao et al., 2012] développé à l'université de Pennsylvania se place dans la lignée des travaux de R. Datta *et al.* [Datta et al., 2008, Datta and Wang, 2010], mais s'en dégage par une architecture audacieuse et une approche originale. Le système n'a malheureusement pas été conduit à son terme, laissant plusieurs pans de développement annoncés mais non réalisés et un ensemble assez incomplet.

L'idée de base consiste à porter assistance à un photographe (muni d'un simple téléphone mobile) en le mettant en relation lors de sa prise de vue avec un système centralisé qui lui transmet des recommandations sur les paramètres à employer pour la photo qu'il souhaite prendre. Le schéma global du système est présenté sur la figure 7.8. Le système centralisé dispose d'une base de données d'images (issues de Photo.net), dotées d'une évaluation selon des critères esthétiques.

L'utilisateur transmet l'image qu'il souhaite prendre dans le nuage (*cloud*) où une analyse est faite de l'image sous plusieurs aspects :

- les couleurs sont extraites par classification dans l'espace HSV ;
- des descripteurs de compositions sont extraits en nombre de 5 : organisation diagonale, verticale, horizontale, centrale ou texture. Ces descripteurs s'appuient sur des détecteurs de contours suivis de filtrage par pertinence ;
- des éléments de Gestalt sont utilisés pour extraire des motifs significatifs sur des critères de parallélisme, similarité et continuité ;
- le bruit granulaire est estimé dans l'image source.

Muni de ces éléments de description, les images les plus semblables sur chacun de ces critères séparément, sont sélectionnées par le serveur qui les propose comme exemples à l'utilisateur. La sélection est faite par des processeurs en parallèle. Le moteur ACQUINE de Datta est l'un d'eux,

ainsi que le classifieur SIMPLiCity.

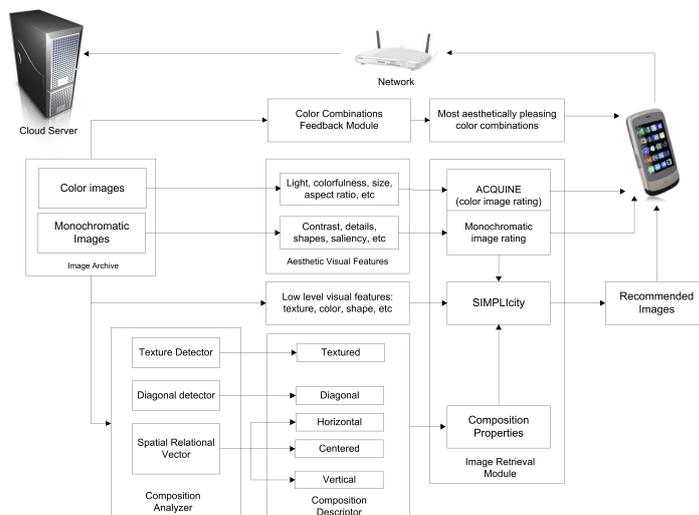


Figure 7.8 – Schéma de principe du système OSCAR, d’après [Yao et al., 2012]. Ce système fait une grande place aux règles de composition et d’harmonie chromatique pour évaluer la qualité des images de la base de données qui seront fournies en exemple pour l’éducation du photographe.

7.7 Un système embarqué : l’approche de K.Y. Lo et al.

Dans les articles [Lo et al., 2012, Lo et al., 2013], une technique de détermination de la qualité esthétique est développée, avec l’objectif de l’embarquer dans un appareil, c’est-à-dire d’assurer une complexité de calcul modeste et une faible consommation électrique. Ces travaux sont à rapprocher de ceux de L. Yao et al. (vus ci-dessus) ou ceux de W.N. Wang et al. [Wang et al., 2014] qui cherchent également à embarquer le système, d’évaluation dans un téléphone mobile. Ils ont donc recours à la puissance du *cloud* pour y transporter les plus gros calculs, ne gardant dans le téléphone que l’interface avec l’utilisateur. Les techniques employées par K.Y. Lo et al. sont assez proches de celles de l’étude de S. Dhar et al. (section 7.4). On attribue 5 notes à chaque image jugeant : la couleur, la saturation, la composition, la richesse et le contraste (voir figure 7.9 à gauche). Un apprentissage est fait sur 2 classes : l’une des images de haute qualité (la classe dénotée *P* ci-dessus), une autre sur la classe de basse qualité (la classe dénotée *A* ci-dessus). Le classifieur est un SVM.

Les primitives détectées sont :

- des descripteurs de couleur semblables à ceux utilisés dans l’étude précédente, construits sur les histogrammes fortement quantifiés (voir figure 7.9 à droite) ;
- des indicateurs de teinte, saturation et luminance obtenus à partir de la moyenne des plans HSV déduits de l’image, puis un indicateur global égal à la somme des 3 moyennes précédentes ;
- des indicateurs de contours sur les 4 canaux H, S, V et H+S+V ;
- des indicateurs de texture calculés après division de l’image en tiers horizontalement et verticalement et intégration des intensités d’une part et des gradients d’autre part dans chacune de ces bandes ;
- un indicateur de flou, un indicateur de canal noir et un décompte des zéros de l’histogramme HSV quantifié.

C’est donc au total une trentaine de primitives qui sont utilisées. La base de données d’images est celle de l’Université de Hong-Kong. Plus de 10 000 images sont utilisées. Les images sont de très petite taille (moins de 250 000 pixels). Elles comportent des photos de types “professionnel” et “amateur”, classées sous 7 rubriques différentes selon leur thématique.

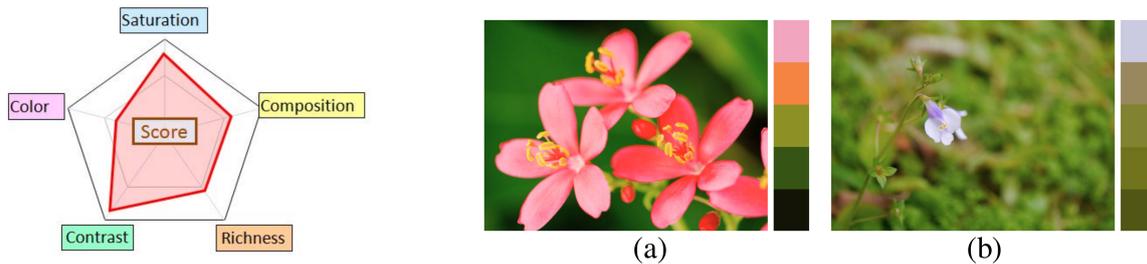


Figure 7.9 – L’approche de Lo et al.. A gauche, la note finale se déduit de 5 notes partielles dont l’importance varie selon la catégorie des images traitées. A droite, la palette est réduite à 5 teintes dominantes et permet de distinguer les images avec une harmonie agréable (en (a)) et celles moins plaisantes (en (b)), (d’après [Lo et al., 2012]).

7.8 Descripteurs génériques de la vision : l’approche de L. Marchesotti *et al.*

Ces travaux ont été présentés dans [Marchesotti et al., 2011, Marchesotti et al., 2013]. Sans chercher à coder des “règles de beauté”, les auteurs utilisent des primitives de description génériques développées pour l’indexation et la fouille d’images. Ce sont soit des sacs de mots visuels (BOVW = *Bags of visual words*) soit des vecteurs de Fisher (FV). Ils opèrent au second niveau de représentation du signal, s’appuyant sur des primitives de bas niveau (en l’occurrence des SIFT³ d’une part et des descripteurs de couleur d’autre part). Les SIFT et les descripteurs de couleur sont traités séparément. Ils sont réduits par une analyse en composantes principales (PCA) en 64 dimensions puis regroupés dans des mots obtenus en concaténant en une chaîne unique soit leur nombre d’occurrences (BOVW) soit leurs statistiques (à l’aide d’une modélisation par mélange de gaussiennes) jusqu’à l’ordre 2 (FV).

Une information sur la composition de l’image peut être donnée par représentation pyramidale mais ne semble pas donner de meilleures performances.

La classification se fait par SVM et l’apprentissage à l’aide d’un algorithme stochastique de descente de gradient. La classification est conduite séparément sur les SIFT et sur la couleur, et leurs scores sont ensuite moyennés. Comme précédemment, deux classes sont recherchées : *Low* et *High*. La base de données utilisée est tout d’abord celle de *Photo.net* un peu modifiée pour en retirer les cadres de certaines images (résultant en une sélection de 3 200 photos environ). Les auteurs ont ensuite utilisé la base CUHK (12 000 images). Ces travaux ont finalement contribué à la création de la base de données annotée AVA, spécialement dédiée à l’étude des propriétés esthétiques des images [Murray et al., 2012].

En conclusion de ces travaux, les auteurs soulignent les très bonnes performances des primitives génériques associées à une classification de second niveau (en particulier les FV). Cette approche permet de dépasser les performances données par les seules primitives “esthétiques” en confiant au classifieur FV le soin d’incorporer des propriétés de haut niveau dans la décision. Notons enfin que cette technique de recherche associant esthétique et sémantique a fait l’objet d’un brevet [Murray et al., 2013].

7.9 Harmonie chromatique : les travaux de P. Lu *et al.*

Les travaux présentés dans [Lu et al., 2014a, Lu et al., 2015a, Lu et al., 2016] ont un objectif plus réduit que les précédents puisqu’ils visent uniquement à déterminer l’esthétique de la palette d’une

3. Les SIFT (*Scale-invariant feature transform*) sont des détecteurs de points caractéristiques multi-échelle.

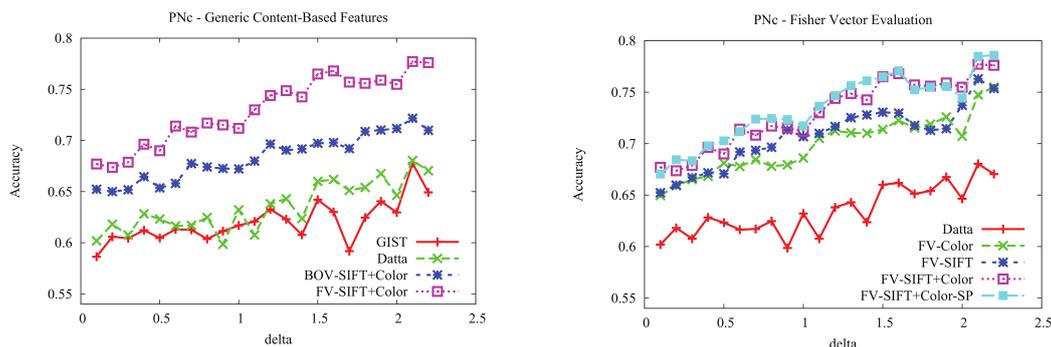


Figure 7.10 – L’approche de *L. Marchesotti et al.* À gauche, comparaison des performances de reconnaissance sur la base de données *Photo.net*, des diverses stratégies de représentation des images. La courbe notée “*Datta*” est celle décrite en section 7.1. La variable “*delta*” portée en abscisse correspond à l’écart entre les notes des images “*Low*” et “*High*” imposé lors de l’apprentissage. À droite, rôle des descripteurs *SIFT* ou de couleur dans la performance du classifieur : ils sont assez complémentaires. La courbe marquée *SP* utilise une décomposition pyramidale. Pour une plus grande complexité, elle n’offre pas des performances supérieures, d’après [*Marchesotti et al., 2011, Marchesotti et al., 2013*].

photographie.

P. Lu et al. ont choisi d’apprendre par une technique automatique les règles de l’harmonie. Ils comparent un apprentissage automatique avec les règles d’harmonie de *Moon et Spencer* et de *Matsuda* présentées dans la section 4.6.3.

Pour cela, ils s’appuient sur la base de données annotée *AVA* [*Murray et al., 2012*] (250 000 images), triée par grandes thématiques (fleurs, paysages, etc.). Huit thématiques sont considérées. Pour chaque thématique, ils distinguent deux ensembles : celui des images harmonieuses et celui des images banales. Puis, après avoir décomposé chaque image en une mosaïque de vignettes, ils utilisent une approche bayésienne (la méthode *Latent Dirichlet Allocation (LDA)*) pour déterminer les associations de vignettes les plus fréquemment attribuées aux “belles” images. Pour cela, ils emploient un régulariseur laplacien, (normalisé ou non), puis une régression par *LASSO* pour estimer le score d’esthétique de l’image [*Lu et al., 2016*].

Leurs résultats montrent des performances de reconnaissance des “belles” images supérieures à celles des performances obtenues avec les critères de *Moon et Spencer*, mais surtout avec les critères de *Matsuda* (voir la figure 4.15) qui apparaissent les plus mauvais (figure 7.11).

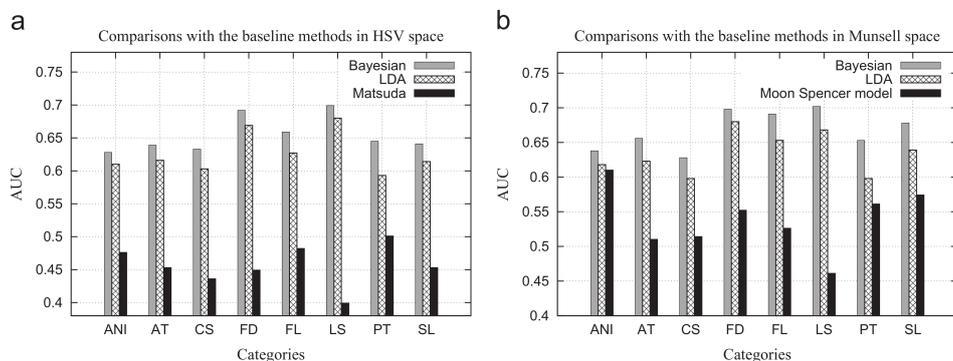


Figure 7.11 – Apprentissage de l’harmonie des couleurs sur la base *AVA*. À gauche avec les critères de *Matsuda* dans l’espace *HSV*, à droite avec les critères de *Moon et Spencer* dans l’espace de *Munsell*, d’après [*Lu et al., 2014a*].

7.10 Réseaux sociaux et *crowdsourcing* : l'approche de R. Schifanella *et al.*

Les travaux présentés dans [Schifanella et al., 2015] se distinguent des précédents sur plusieurs points. Tout d'abord, ils font appel à des images ordinaires, la base BEAUTY issue du réseau social Flickr indexée en fonction de leur qualité par la communauté des internautes et *crowdsourcing* avec de grandes précautions pour éviter des biais dans les résultats statistiques (voir 6.2). La question posée : *How beautiful is this picture ?* se réfère aux seules qualités esthétiques et non au contenu de l'image. La note est attendue parmi 5 : *inacceptable, avec des défauts, ordinaire, professionnelle, exceptionnelle*. Enfin des critères différents sont utilisés pour les diverses classes : "personnes", "nature", "animaux" et "ville".

Les primitives utilisées sont des primitives visuelles de couleur (à partir des critères issus de la psychologie affective (voir la section 4.6.2 [Machajdik and Hanbury, 2010]), de texture et de disposition spatiale (règle du tiers, distribution des gradients). La classification se fait par régression quadratique partielle (PLSR = *Partial least square regression*).

La mesure de qualité obtenue pour une image inconnue à l'aide du classifieur précédent peut être comparée à la note de popularité de l'image sur Flickr. Elle ne montre aucune relation. Elle peut également être comparée à celle obtenue par des experts (la méthode de Marchesotti, section 7.8). Elle montre alors une tendance assez nette à surnoter les photos.

7.11 Prise en compte de commentaires : les travaux de J. San Pedro *et al.*

Une approche très originale mérite ici d'être évoquée quoiqu'elle nous éloigne sensiblement des méthodes automatiques d'évaluation de la qualité à partir des propriétés iconographique. C'est celle rapportée dans [San Pedro et al., 2012]. L'idée développée consiste non pas à traiter l'image mais à examiner les commentaires qui accompagnent l'image. Les auteurs utilisent DPChallenge qui fournit à la fois les images et les commentaires associés déposés par les internautes.

Pour ce faire les auteurs identifient des mots de vocabulaire dotés de significations variées : certains désignent des propriétés de l'image : "éclairage", "composition", "couleurs", d'autres des valeurs positives ou négatives de jugement des internautes. Collectant toutes les opinions portées sur la photo, les auteurs déterminent, pour chaque attribut, le sentiment qui ressort globalement de ces avis. Cela est fait à l'aide d'un processus markovien qui estime un état caché (la qualité de la photo). L'apprentissage des dépendances entre critères et beauté est fait de façon supervisée par maximisation d'entropie selon le schéma popularisé par le système de traitement du langage naturel *OpinionMiner* [Jin et al., 2009].

Cette stratégie est employée pour réordonner les résultats d'une requête sur une base d'images, de façon à placer en tête des retours les images les mieux classées par les critères esthétiques ainsi dégagés. Une courte expérimentation montre que la sélection s'appuyant sur l'analyse des commentaires est plus pertinente qu'une sélection sur les seuls mots clefs de la requête ou qu'une sélection utilisant des critères esthétiques tirés de l'image (suivant les algorithmes proposés dans [Datta et al., 2006], section 7.1 et dans [San Pedro and Siersdorfer, 2009]).

Chapitre 8

Les systèmes à base de Réseaux de Neurons Profonds

Les techniques à base de Réseaux de Neurons Profonds sont l’objet d’une grande attention depuis quelques années chaque fois que l’on cherche à reproduire une prise de décision complexe ordinairement assurée par un expert. L’évaluation esthétique relève tout à fait de ce critère et l’on a donc vu plusieurs travaux s’engouffrer dans cette voie. La référence [Deng et al., 2017b] donne une bonne revue des travaux disponibles jusqu’aux premiers mois de 2017.

Comme nous l’avons dit les RNP proposent une approche “de bout en bout” pour laquelle l’image à évaluer est directement mise en entrée du système. La couche d’entrée dispose donc d’un neurone par pixel, ou plutôt de 3 neurones par pixel puisque l’on traite généralement des images en couleur. Les premières couches (en nombre variable, de 3 à 5), sont traditionnellement des couches de neurones convolutifs¹ qui permettent de réduire progressivement la taille des cartes traitées en reliant un neurone de la couche n à un petit nombre (3×3 , 7×7) neurones de la couche $n - 1$. On peut aussi réduire la dimension des couches par des opérations de mutualisation (*pooling*) qui remplacent le sous-échantillonnage des couches convolutives par d’autres fonctions de décimation (fonction “sup”, “médian” ou “inf”). Les couches suivantes sont souvent totalement connectées, de façon que chaque neurone de la couche n soit lié à tous les neurones de la couche $n - 1$. La dernière couche comporte autant de neurones que de classes traitées (par exemple 2 classes pour un jugement “Amateur” ou “Professionnel”, ou 5 classes pour un jugement à 5 niveaux). Les architectures utilisées sont celles disponibles en ligne : AlexNet, ResNet, VGG, Inception ou MobileNet par exemple.

Les limitations des RNP pour l’évaluation de l’esthétique des images sont étroitement liées à l’exigence d’avoir des images de bonne qualité disposant d’une finesse de détails et d’une richesse de champ compatible avec la qualité des photos professionnelles. Cela devrait se traduire en pratique par :

1. une couche d’entrée de plusieurs millions de neurones, généralement triplée pour rendre compte de la couleur ;
2. un nombre de connexions considérable à l’intérieur du réseau, conduisant à des temps de calcul extrêmement longs pour l’apprentissage des paramètres et l’obligation de disposer de bases de données annotées gigantesques pour conduire cet apprentissage.

Les systèmes opérationnels en 2017 fonctionnent surtout sur des imagerie de taille réduite (généralement inférieure à 256×256) et ne rendent compte de la finesse à haute résolution, lorsqu’ils s’en préoccupent, que par des pirouettes (voir les sections 8.1.1 et 8.1.2 par exemple). Seul le travail présenté en section 8.1.3 aborde ce problème de front. Pourtant, une petite taille fait perdre d’importants détails pour juger la qualité de l’image et l’on ne peut guère faire confiance à des résultats obtenus sur de si médiocres représentations.

1. Les réseaux convolutifs sont désignés dans la littérature par l’acronyme CNN (*convolutionnal neural network*).

Au-delà de cet aspect lié à la taille de l'image, se pose le problème de son format (ou facteur d'aspect). Les images peuvent se présenter selon des formats très différents tandis que le réseau de neurones est par nature de taille fixée (généralement carrée que nous désignerons par $N \times N$). Il est classique de répondre à ce problème en choisissant l'une des 4 pistes suivantes (illustrées sur la figure 8.1) :

1. en sélectionnant une sous-image $N \times N$ au sein de l'image (opération de recadrage, *cropping*) ;
2. en déformant l'image pour la ramener au format $N \times N$ (opération de rééchantillonnage) ;
3. en réduisant l'image de façon que sa plus grande dimension vaille N et en complétant la plus petite dimension par l'image repliée (opération de remplissage, *padding*) ;
4. en utilisant une version inversée de l'image dans l'opération de *padding* précédente, de façon à éviter les discontinuités lors du repliement (opération de remplissage après retournement, *padding/flipping*).

D'un point de vue de la conservation du signal, le remplissage après retournement préserve le mieux à la fois les propriétés géométriques et les propriétés spectrales.

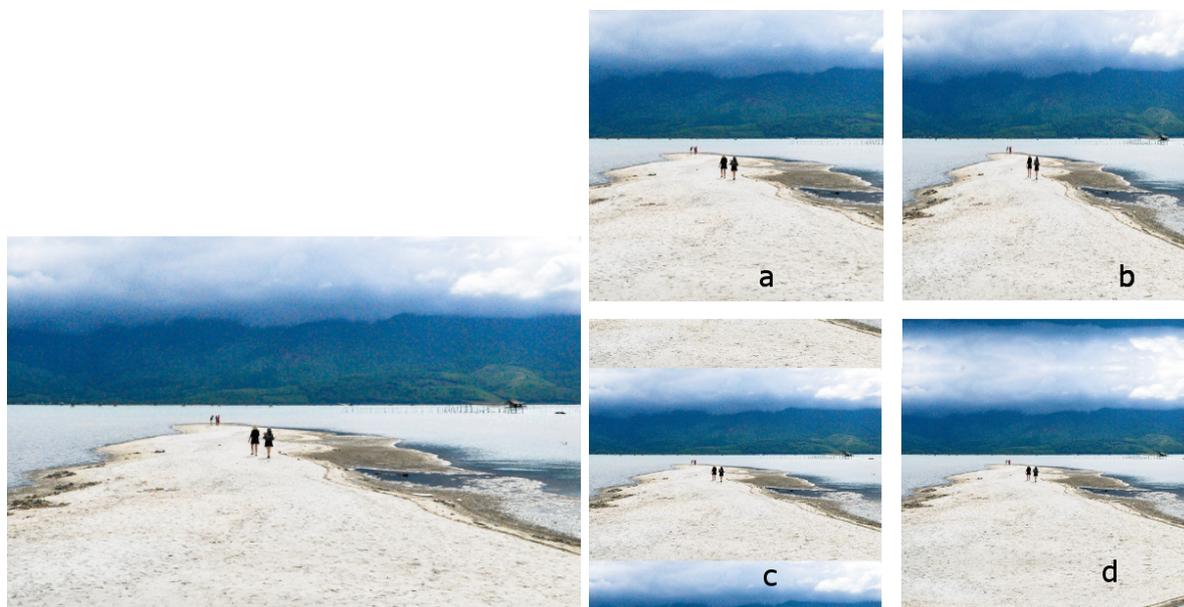


Figure 8.1 – On peut adapter une image rectangulaire (à gauche) à la couche d'entrée carrée d'un RNP, en a) par recadrage, *cropping*, en b) par rééchantillonnage affine, en c) par remplissage, *padding*, en d) par remplissage après retournement, *padding after flipping*). C'est cette dernière solution qu'il faut privilégier.

Examinons maintenant quelques uns de ces travaux.

8.1 Les RNP dédiés à l'évaluation esthétique

8.1.1 Le système RAPID de X. Lu *et al.*

Dans [Lu et al., 2014b, Lu et al., 2014c, Lu et al., 2015b] est présenté le premier système utilisant l'apprentissage profond (RNP (réseaux de neurones profonds) ou DNN (*Deep Neural Network*)) : RAPID (*R*ating *P*ictorial *a*esthetics using *D*eep *l*earning). L'usage des DNN doit permettre de supprimer l'étape de choix des primitives placées à l'entrée des systèmes de classification, en nourrissant directement le classifieur par l'image. Celle-ci est alors traitée par une succession de réseaux convolutifs qui réduisent progressivement la taille du signal avant de le traiter par des couches de réseaux totalement

δ	[Marchesotti et al., 2013]	SCNN	AVG-SCNN	DCNN	RDCNN
0	66.7 %	71.20 %	69.91 %	73.25 %	74.46 %
1	67 %	68.63 %	71.26 %	73.05 %	73.70 %

Table 8.1 – Résultats de l'apprentissage par RNP. Le paramètre δ est le même que le delta de la figure 7.10. Le réseau SCNN est à une seule colonne. AVG-SCNN est le résultat du moyennage de 2 réseaux indépendants. Le réseau DCNN est celui qui utilise 2 colonnes couplées à la dernière couche, tandis que le réseau RDCNN est celui qui est régularisé par l'information de "style", d'après [Lu et al., 2014b].

connectés. Ainsi le RNP extrait lui-même les primitives les plus pertinentes. Le travail présenté dans [Lu et al., 2014c] prend en compte de façon très originale la taille et la diversité des formats d'images.

Comme dans tous les autres systèmes qui utilisent l'architecture VGG-Net², la taille est ramenée à la dimension 256×256 , sur 3 canaux, par un zoom et l'image est normalisée en taille et en aspect en la complétant par des zéros. Cela constituera le RNP principal de traitement. Il est alors utilisé un second RNP pour traiter une vignette, elle aussi de taille $256 \times 256 \times 3$, dont la position est tirée aléatoirement, à pleine résolution, dans l'image originale. Pour rendre compte de la richesse de l'image, lors de l'apprentissage, 50 vignettes sont ainsi tirées. Chaque vignette, associée au canal principal, sera alors traitée par le double réseau et l'on retient celle qui donne le meilleur score. Ces deux RNP partagent finalement la dernière couche totalement connectée des réseaux pour fournir un score unique : 0 ou 1 (figure 8.2).

Ces travaux exploitent la base de données AVA pour l'apprentissage (voir section 6.2).

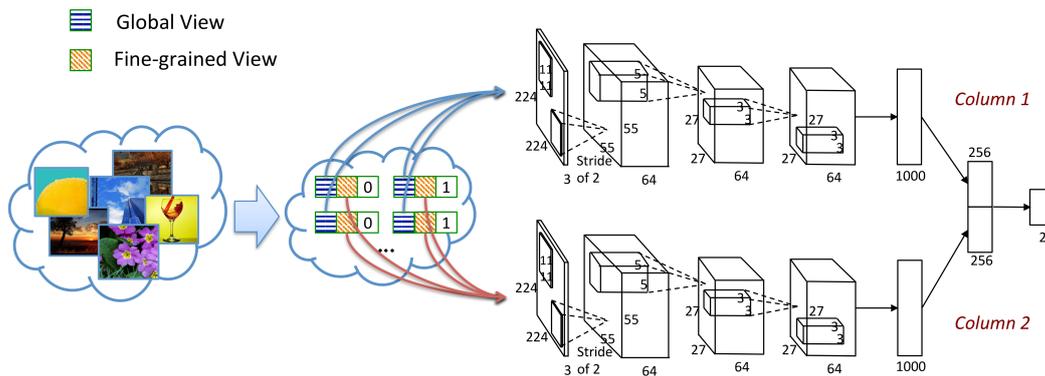


Figure 8.2 – Réseau neuro-mimétique profond à deux colonnes proposé par X. Lu et al. afin de réaliser l'apprentissage. Les deux colonnes sont identiques. L'une est nourrie par l'image réduite et normalisée, l'autre par la vignette extraite de façon aléatoire de l'image d'origine, à pleine résolution. Les 4 premières couches sont convolutives (avec des noyaux de $11 \times 11 \times 3$, $5 \times 5 \times 64$ et deux fois $3 \times 3 \times 64$). La réduction de dimension des couches totalement connectées se fait par mutualisation par maximum (max-pooling). La première a 1000 neurones, la seconde 256. Enfin la sortie est à 2 états. D'après [Lu et al., 2014c].

Les auteurs disposant d'informations de style³ sur certaines images de la base de données font alors un premier apprentissage du style (au moyen d'un RNP semblable au précédent), puis, utilisant l'information de style comme élément de régularisation, procèdent à la détermination de l'appréciation

2. VGG-Net est une architecture développée par le Visual Geometry Group de l'Université d'Oxford. Il se distingue par des convolutions limitées à des noyaux 3×3 . Il a obtenu la seconde place à la compétition ILSVRC-2014 (*ImageNet Large Scale Visual Recognition Competition*). VGG-Net est maintenant un produit de Torch, association de développeurs de logiciels d'intelligence artificielle issus de diverses sociétés des GAFA. VGG16, souvent utilisé pour les études sur l'esthétique offre 16 couches de traitement

3. Le style porte sur des propriétés particulières de l'image, annotées manuellement : présence de couleurs complémentaires, images à deux tons, images en Noir et Blanc, images HDR, photo macro, bougé, déplacement, etc.

esthétique.

Les résultats (table 8.1) permettent de valider le choix de l'architecture ainsi que le protocole d'apprentissage et l'utilisation d'un vecteur de style pour contraindre la solution. Ces résultats sont supérieurs à ceux présentés dans [Marchesotti et al., 2013] pour un coût d'implémentation cependant considérablement accru. Sur la base AVA, ils donnent un accord à plus de 73% avec la notation à 2 classes "Low" et "High".

8.1.2 L'architecture multi-voies DMA-Net de X. Lu *et al.*

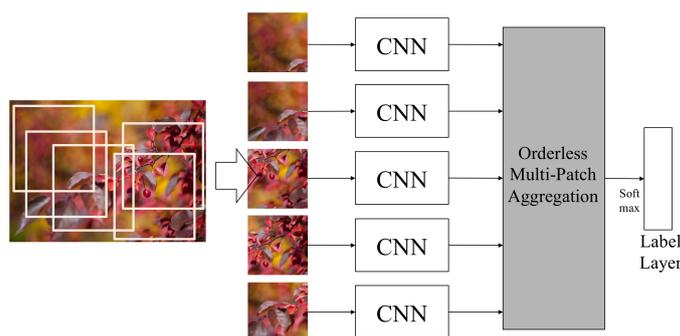


Figure 8.3 – Schéma d'apprentissage par réseau de neurones profond multi-voies. Afin de simplifier l'apprentissage, les divers CNN sont identiques. L'étape de fusion (Orderless Multi-Patch aggregation) est décrit sur la figure 8.4. Le choix de sous-images aléatoires se montre, en pratique, préférable à un positionnement prédéfini, d'après [Lu et al., 2015b].

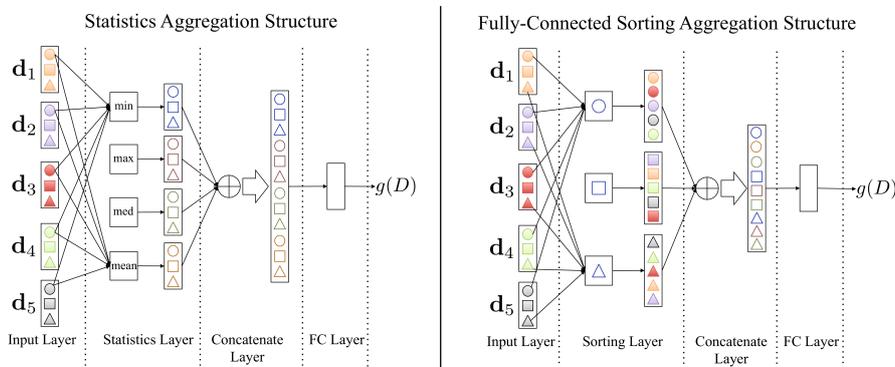


Figure 8.4 – Fusion des réseaux multi-voies (Orderless Multi-Patch aggregation) de la figure 8.3. Les réseaux neuro-mimétiques étant sensibles à l'ordre des données en entrée, il est souhaitable d'introduire une couche assurant l'indépendance du résultat à l'ordre de choix des sous-fenêtres avant l'étape de fusion. Cela peut être fait soit en fusionnant tous les neurones de même rang par un opérateur indépendant de l'ordre (max, min, médian) (à gauche), soit en triant les données de ces neurones de façon à les ordonner (à droite), d'après [Lu et al., 2015b].

L'architecture de RAPID est ensuite étendue dans [Lu et al., 2015b] en remplaçant l'architecture à deux voies par une architecture multi-voies, conformément au schéma de la figure 8.3. Elle prend le nom de DMA-Net (*Deep multipatch aggregative network*). On adopte un schéma à cinq voies, chaque voie étant une sous-image à pleine résolution tirée aléatoirement au sein de l'image à classer. Afin de simplifier les modules d'apprentissage, tous les CNN traitant une sous-image ont des paramètres identiques. Les cinq voies sont ensuite fusionnées selon deux stratégies qui donnent en pratique des résultats assez semblables : soit un opérateur statistique indépendant de l'ordre de ses

Méthode	Précision	Mesure-F	Score AUC
VGG avec recadrage	71.2 %	0.83	0.66
VGG avec chgt d'échelle	73.8 %	0.83	0.74
VGG avec remplissage	72.9 %	0.83	0.73
SPP-CNN	76.06 %	0.84	0.77
MNA-CNN	77.1 %	0.85	0.79

Table 8.2 – Comparaison des performances d'évaluation des images de la base AVA selon le type d'entrée adopté : images réduites par recadrage (*cropping*), changement d'échelle ou remplissage (*padding*) (en haut), à l'aide d'un réseau à multi-voies (SPP est le réseau original de [He et al., 2015b], MNA est celui de L. Mai et al.). Les deux mesures (Mesure-F et AUC) sont déduites de la courbe ROC de reconnaissance, d'après [Mai et al., 2016].

entrées (moyenne, max ou médian), soit une couche de neurones prenant en entrée des vecteurs triés (figure 8.4).

En 2017, le système DMA-Net présenté dans [Lu et al., 2015b] sert de référence pour la plupart des travaux conduits sur la qualité esthétique des photos par RNP.

8.1.3 S'adapter à la taille de l'image : les travaux de L. Mai *et al.*

Ces travaux se sont eux aussi attaqués au problème du format d'entrée de l'image [Mai et al., 2016]. Ils ont choisi une solution complexe mais efficace en adoptant une approche par décomposition pyramidale d'une part et par parallélisation des voies de traitement d'autre part.

L'idée de base consiste à faire précéder les couches de traitement classiques d'un RNP d'une couche de convolution et de mutualisation *pooling*. Pour s'adapter aux diverses tailles possibles d'images, les auteurs proposent de mettre ainsi en parallèle cinq RNP, avec des facteurs de décimation de 12×12 , 9×9 , 6×6 , 4×4 et 2×2 respectivement (voir la figure 8.5 à gauche)⁴. Par ailleurs, le rapport d'aspect de l'image pouvant ne pas être le même que celui (généralement carré) de la couche d'entrée du RNP, ils proposent une technique de mutualisation adaptative, dérivée, comme la précédente, des travaux de [He et al., 2015b]. L'architecture ainsi conçue est appelée MNA-CNN (*Multi-Net adaptive pooling CNN*) ; elle est réalisée à l'aide d'un réseau VGG-Net (voir note 2), pré-entraîné à l'aide du modèle Zoo de BVLC/CAFFE⁵

Ils adoptent ensuite une amélioration supplémentaire en prenant en compte le contenu sémantique de l'image. Cela est fait en ajoutant un canal supplémentaire, en pratique, celui développé par Google (Places 205-GoogLeNet)⁶, dont ils réduisent le nombre de classes à 7 (au lieu de 205) : *humains, plantes, architecture, paysage, nature-morte, animal et nuit* (figure 8.5 à droite).

L'expérimentation est conduite sur la base AVA. Les résultats ainsi obtenus sont meilleurs que ceux donnés par les méthodes classiques de recadrage *cropping*, de sous-échantillonnage, de mise à l'échelle ou de pavage (voir la table 8.2). Ils montrent aussi le bénéfice du canal de catégorisation.

8.1.4 La beauté du web : travaux de M. Redi *et al.*

Présentée dans [Redi et al., 2017], cette démarche a pour objectif de trouver des critères de beauté, non pas dans les photos de sites spécialisés, mais dans les images ordinaires issues du web qui comporte

4. Il semble que l'on puisse ainsi traiter des images dont la dimension maximale serait $12 \times 256 = 3072$ pixels dans la grande dimension, mais cela n'est pas précisé dans l'article.

5. CAFFE est une plateforme logicielle de réseau neuromimétique profond, développée par Berkeley AI Research (BAIR) et créée par Yangqing Jia au BVLC (Berkeley Vision and Learning Center).

6. GoogLeNet est un réseau à 22 couches développé par Google, 205 indique le nombre de classes de la couche de sortie. GoogLeNet a remporté la compétition ILSVRC-2014.

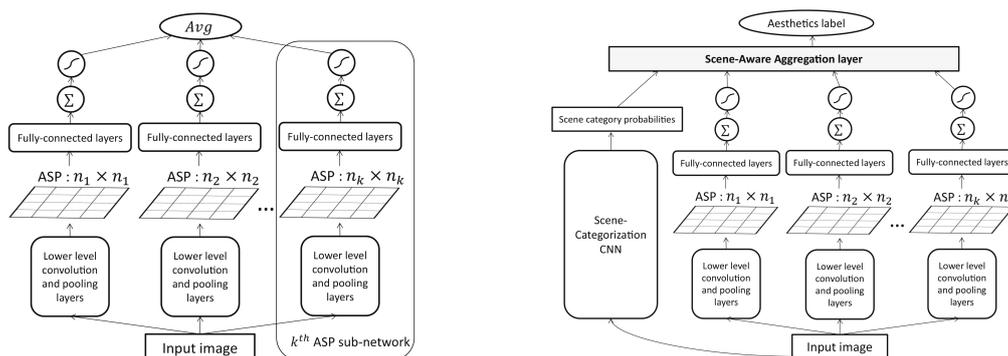


Figure 8.5 – Les deux architectures proposées par L. Mai et al.. A gauche pour traiter le problème d’adaptation au format de l’image, cinq réseaux sont mis en parallèle, précédés chacun d’une couche (convolution + mutualisation) de décimation variable. A droite un canal supplémentaire est ajouté pour prendre en compte la catégorie de l’image d’entrée (d’après [Mai et al., 2016]).

beaucoup plus d’images médiocres, voire de dessin ou de graphiques, que les sites de photo. Elle a pour cela utilisé une base spécifique (voir la section 6.2).

Les auteurs souhaitent également comparer les performances de trois chaînes de traitement différentes :

1. une méthode à base de primitives et d’un classifieur par forêts aléatoires (*random forest*). Les primitives choisies couvrent assez largement les descripteurs d’images : couleur, texture, contraste, mais aussi des primitives sur la composition : présence d’objets, symétries et unicité. Ils sont au nombre de 27 ;
2. un réseau de neurones profond à 3 sorties : “haute-qualité”, “qualité ordinaire” et “basse qualité”, qui est entraîné entièrement à partir de la base de données issue du web ;
3. un réseau de neurones profond pré-entraîné à partir d’une base de données AVA (images essentiellement “esthétiques”) selon le protocole de [Lu et al., 2015b], puis finement réglé à l’aide d’images de la base web (donc avec des images de mauvaise qualité).

Ces travaux confirment que quelques primitives permettent de distinguer avec une raisonnable fiabilité, “photos” et “non-photos” de façon à adapter, ensuite, l’évaluation de la qualité esthétique à cette information. Ils montrent que les meilleures performances sont obtenues à l’aide du réseau de neurones profond entraîné à l’aide du web (et non de la base AVA). Ils montrent aussi que les performances du classifieur sur primitives sont très proches des meilleures. Enfin, ils montrent que, si un réseau entraîné avec les seules images de bonne qualité de la base AVA donne des résultats médiocres (en particulier à cause des erreurs commises sur les images “non-photo”), il est possible de l’améliorer sensiblement par raffinement d’apprentissage (sans toutefois atteindre les performances d’un réseau entraîné sur une base web).

8.1.5 Le réseau siamois de S. Kong et al.

Ce travail (présenté dans [Kong et al., 2016]) propose d’ambitieux objectifs pour vérifier les capacités d’évaluation esthétique des RNP. Pour cela il développe tout d’abord une base de données originale (AADB, voir la section 6.2), elle aussi renseignée par *crowdsourcing* avec une évaluation de la qualité, une catégorisation et des attributs de qualité esthétique. A la différence des autres bases de données AADB conserve l’identité des évaluateurs, photo par photo.

Le schéma adopté est un RNP travaillant sur une image réduite (256×256), éventuellement repliée, sans canal de traitement de l’information à pleine résolution. Le réseau est amélioré par un montage “siamois” [Chopra et al., 2005] qui procède par traitement parallèle (avec des RNP semblables) de paires d’images qui ont été jugées de façon à peu près équivalente sur le critère d’esthétique (figure 8.6 à gauche).

Les auteurs testent tout d'abord le rôle des diverses informations dans les performances du système et pour cela ils développent une architecture capable de suppléer l'absence d'information, sur la catégorie par exemple, ou sur les attributs (figure 8.6). Ils montrent que toutes ces informations sont utiles mais ne sont pas indispensables. Les conclusions de l'expérimentation montrent de plus

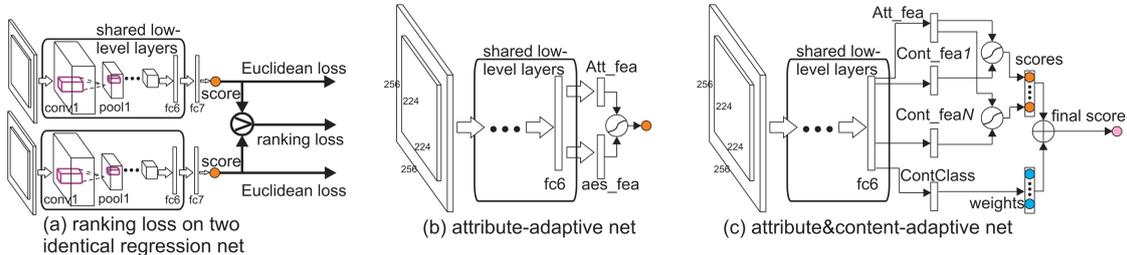


Figure 8.6 – Les divers éléments d'architecture utilisés dans [Kong et al., 2016]. A gauche, le réseau “siamois” qui traite deux images en parallèle avec les mêmes architectures. Au centre, un réseau entraîné de façon à donner simultanément une évaluation de l'esthétique (en bas) et une détection de primitives (en haut), ces deux informations étant combinées pour fournir l'évaluation finale. A droite, une architecture où l'on détecte à la fois les primitives et le contenu.

l'importance de l'observateur ayant annoté l'image dans le classement, ce que les auteurs mettent sur le compte d'une “même esthétique”, mais qui peut aussi être mis au compte de la fiabilité réduite du *crowdsourcing*. Elles montrent aussi la très faible généralisation des réseaux ainsi entraînés car les performances d'apprentissage et de reconnaissance croisées sur les bases AADB et AVA se révèlent très médiocres. Enfin les auteurs souhaitaient proposer une note “continue” entre les photos laides et les belles photos, par opposition aux approches binaires, mais aucun résultat n'est présenté sur ce point.

Une configuration assez semblable à celle des réseaux siamois, une configuration en réseaux “adverses”⁷ est utilisée dans [Deng et al., 2017a] pour améliorer une image initialement médiocre par un changement de contraste (séparément sur les canaux *Lab*) et par un recadrage. Un réseau adverse est un réseau génératif qui modifie des paramètres de l'image d'entrée de l'une des deux voies d'un réseau siamois de façon à obtenir des sorties identiques à une couche d'évaluation. Pour cela on met en opposition une image réputée belle et l'image médiocre et l'on fait évoluer les paramètres de la seconde jusqu'à rendre les sorties indiscernables.

8.1.6 Utiliser l'information sémantique : les travaux de Kao et al.

Présentés dans [Kao et al., 2017], ces travaux se proposent de démontrer que l'évaluation esthétique de la photo bénéficie d'une connaissance du contenu sémantique. Ils proposent une architecture par RNP qui accomplit simultanément les deux tâches de détermination du contenu sémantique et de l'évaluation des qualités esthétiques. Ces travaux sont conduits à l'aide du réseau décrit en figure 8.7. C'est un schéma classique de MT-CNN (*Multi task convolutional neural net*) qui est suivi d'une étape d'apprentissage des dépendances de la note esthétique en fonction de la classification sémantique.

Les images sont ici encore de taille modeste (256×256). Elles sont issues des bases AVA qui servent à l'apprentissage; les résultats d'apprentissage ont été transportés avec succès à des images issues de Photo.net.

Une image x_n ($n \in 1, \dots, N$) est décrite par un attribut d'évaluation esthétique y_n qui peut prendre C valeurs et des attributs sémantiques z_n dont les valeurs z_n^m ($m = 1, \dots, M$) sont soit 0 soit 1 selon que l'image appartient aux images de sémantique m ou non. La base de données (X, Y, Z) est donc

7. Les réseaux adverses sont désignés par AN : “*Adversarial network*”. Lorsqu'ils sont utilisés pour créer de nouvelles images, on dit qu'ils sont “génératifs” et on les désigne par GAN, [Isola et al., 2017].

δ	MT-CNN 1	MTRL-CNN
0	76.15 %	79.08 %
1	75.90 %	77.71 %

Table 8.3 – Résultats comparatifs du traitement avec le couplage entre sémantique et esthétique et sans ce couplage. MT-CNN correspond au premier étage de CNN, sans le couplage (celui de la figure 8.7). MTRL-CNN est le réseau complet avec couplage de Kao et al.. Le paramètre δ est le même que le delta en abscisse de la figure 7.10, d’après [Kao et al., 2017].

constituée de N triplets $\{x_n, y_n, z_n^m\}$. La première étape confiée au MT-CNN se fait à l’aide d’une régression logistique multinomiale. Elle est suivie par une étape de prise en compte de la dépendance explicite entre sémantique et esthétique par l’intermédiaire de la matrice de covariance Ω :

$$\Omega = \frac{(W^T W)^{1/2}}{\text{Tr}[(W^T W)^{1/2}]} \quad (8.1)$$

où W est le vecteur des paramètres de la couche finale (figure 8.7). L’optimisation se fait alors en alternant rétro-propagation du gradient à Ω constant pour calculer les W et recalcul de Ω à partir de ces valeurs de W . Cette méthode prend le nom de MTRL-CNN (*Multi Task Relationship Learning Convolutional Neural Net*).

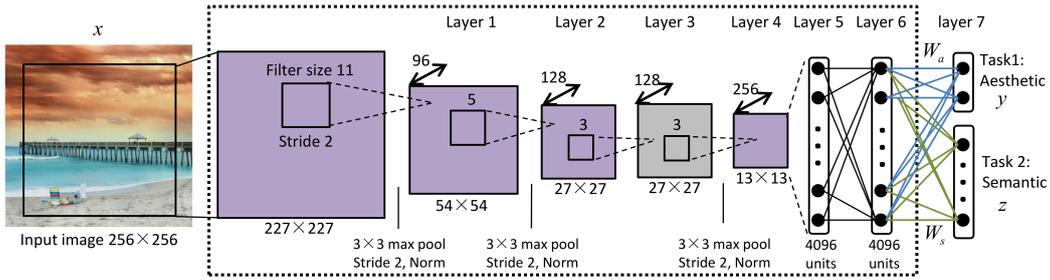


Figure 8.7 – L’architecture utilisée dans la première étape des travaux de Kao et al. On reconnaît quatre couches de réseau convolutif avec décimation, suivies de trois couches totalement connectées, la dernière étant constituée de deux réseaux spécialisés, l’un pour attribuer une note esthétique, l’autre pour reconnaître le contenu sémantique (d’après [Kao et al., 2017]).

L’expérimentation est conduite sur la base AVA dont sont tirées 185 000 images avec au moins un identifiant sémantique d’appartenance à l’une des 29 classes sémantiques disponibles ($M = 29$). L’entraînement se fait sur 165 000 images et la vérification sur 20 000. Deux valeurs d’esthétique seulement sont utilisées ($C = 2$). Les images AVA étant notées sur 10, l’apprentissage utilise un terme δ pour séparer autour de la valeur 5 les belles ($> 5 + \delta$) et les vilaines images ($< 5 - \delta$). δ vaudra 0 ou 1.

Les résultats obtenus dans ces conditions montrent un taux de bonne classification dans l’une des deux classes esthétiques supérieur à 75 % (voir table 8.3). Le bénéfice de l’information sémantique est toujours vérifié, mais modéré (de l’ordre de 2 à 3 %).

8.2 Variantes autour de l’architecture de base des RNP

De très nombreux travaux sont consacrés aux techniques d’évaluation de la qualité esthétique par RNP. Si les architectures de base sont celles que nous avons examinées ci-dessus, une tendance pourrait apparaître dans les derniers travaux à remplacer les couches totalement connectées par des couches purement convolutives. Par ailleurs, un effort est fait pour proposer des interprétations des

toute dernières couches pour justifier la classification faite [Jin et al., 2016a, Kairanbay et al., 2017, Murray and Gordo, 2017] en présentant des cartes d'activation superposées aux images. Ces résultats sont encore loin d'être convaincants.

D'autres approches visent à simplifier les architectures [Kairanbay et al., 2017] ou à rendre plus simple l'apprentissage [Srivastava and Kant, 2018] (en l'occurrence en proposant un apprentissage *from scratch*). Nous allons voir ici plus en détail quelques-unes de ces tendances.

8.2.1 Une approche qui distingue la nature de la photo : les travaux de Y. Kao *et al.*

La méthode développée par Y. Kao, K. Huang et S. Maybank [Kao et al., 2016], propose d'associer à la photographie sa classe : "Scène", "Objet" ou "Texture" puis, en fonction de la catégorie trouvée, de l'orienter vers 3 réseaux d'architectures un peu différentes.

Ainsi, pour la classe "Scène" un seul réseau classique est utilisé, doté de 5 couches convolutives, suivies de 2 couches totalement connectées. Pour la classe "Objet", le traitement est conduit en parallèle sur deux réseaux, l'un traitant l'image globale, l'autre focalisé sur la zone de plus grande saillance. Les deux canaux passent par une couche totalement connectée, puis sont regroupés avant de traverser une nouvelle couche totalement connectée. Enfin, dans le cas d'une image de "Texture", on sélectionne 16 sous-images par une fenêtre glissante, chaque fenêtre passe par un réseau un peu raccourci puis les 16 résultats sont ensuite moyennés.

Le réseau fournit deux résultats. Tout d'abord par un classement sur 1 bit, une appartenance ou non à la classe des "belles images", puis, par une régression utilisant la somme des carrés des sorties de la toute-dernière couche totalement connectée, une note continue d'évaluation de la qualité esthétique de la photo.

Le réseau ainsi développé est appelé A&C CNN : (*Aesthetic and Category*). Ses résultats sont assez bons, sans vraiment justifier l'accroissement de la complexité du à la séparation en trois catégories.

8.2.2 Comparer les photos entre-elles : la proposition de K. Schwarz *et al.*

Nous avons cité la base de données de l'Université de Tübingen, obtenue en recueillant les avis favorables des visiteurs du site Flickr. Elle a été conçue pour mener une démarche originale de détermination de l'esthétique des images à partir de l'opinion - non spécialiste - d'un grand nombre d'internautes [Schwarz et al., 2016]. Rappelons pour cela que les images se sont vu attribuer une note de qualité, continue, entre 0 et 1 (voir l'équation 6.1 et la figure 8.8). Les auteurs entreprennent alors un apprentissage automatique, à partir d'un réseau profond (architecture ResNet-50⁸ et logiciel TensorFlow⁹) qui utilise les qualités esthétiques non pas pour classer les images ("Belle" ou "Pas belle") mais pour les ordonner par comparaison 2 à 2 (réseaux siamois [Chopra et al., 2005], comme dans l'étude précédente), ou 3 à 3 par les réseaux triplets [Hoffer and Ailon, 2015]. Cet objectif (classer les images par rang d'esthétique) est, de l'avis des auteurs, plus proche de la façon de fonctionner du jugement humain.

Les qualités de ce *ranking* sont vérifiées par recours à un vote massif sur internet pour contrôler que deux images placées dans une relation d'ordre particulière par le RNP se retrouvent bien ordonnées de la même façon par les internautes. Les auteurs montrent qu'ils peuvent cependant déduire de cet ordonnancement une classification aux performances très semblables à celles des autres techniques.

Les auteurs utilisent finalement leur logiciel pour 3 applications différentes :

- pour ordonner des photos dans un album,

8. ResNet est le réseau de neurones convolutif de l'équipe Microsoft qui a gagné la compétition ILSRVC 2015 [He et al., 2015a]). ResNet-50 est l'une des versions de ce réseau.

9. TensorFlow est une librairie de logiciels libres de calcul utilisant les graphes de flots de données. Il a été développé par les équipes du projet Google Brain au sein du centre de recherche en Intelligence Artificielle de Google.

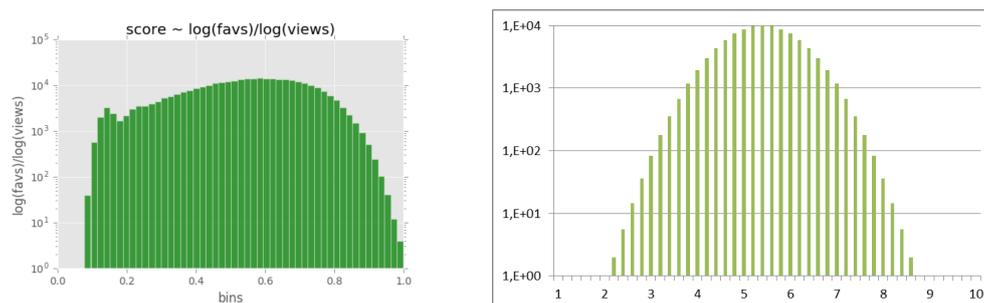


Figure 8.8 – A gauche, répartition des 350 000 notes (entre 0 et 1) de la base de données de Tübingen (d’après [Schwarz et al., 2016]). Les notes sont calculées à partir de la formule (6.1) qui tire profit des avis des internautes. On note que la distribution est presque plate, distribution très différente de celle des images de la base AVA, pratiquement gaussienne, représentée ici à droite également en échelle logarithmique, obtenue à partir de la figure 6.2).

- pour extraire la meilleure photo d’une séquence vidéo,
- pour sélectionner la meilleure sous-image dans une scène large.

8.2.3 A mi chemin entre classification et RNP : la méthode de Z. Wang et al.

Ces travaux [Wang et al., 2016] essaient de coupler l’approche par primitives sélectionnées et par RNP. Faisant fonctionner en parallèle des voies de sélection de primitives, ils se placent sous l’analogie du cerveau humain pour justifier la démarche adoptée et son nom (*Brain-inspired-CNN*), et plus particulièrement dans la lignée du modèle cérébral proposé par A. Chatterjee [Chatterjee, 2003]. Ils divisent ainsi le traitement en deux niveaux :

- dans un premier niveau, des détecteurs sont menés en parallèle sur des primitives élémentaires : la couleur (directement mesurée par ses composantes HSV), puis chacune des 14 informations de style photographique renseignée dans la base AVA (voir la figure 6.3), chaque information étant prise en charge par un RNP composé de 4 couches convolutives et 2 couches totalement connectées) ;
- dans un second niveau, les sorties des détecteurs précédents servent d’entrée à un réseau à 128 entrées, 4 couches convolutives et 2 couches totalement connectées (voir figure 8.9).

La sortie finale consiste en deux valeurs : la moyenne et la variance de la gaussienne qui reproduit la courbe de distribution des avis des membres du jury AVA (voir la figure 6.2)¹⁰. Ces valeurs sont obtenues en remplaçant la fonction de sortie *softmax* par le calcul d’une divergence de Kullback-Leibler.

Résultat intéressant de cette étude, alors que l’on s’attendrait à ce que l’apprentissage de la distribution des notes permette un classement selon une échelle continue, les auteurs insistent sur le rôle particulier de cette distribution pour améliorer la qualité du classement binaire : ”belle image “ versus ”vilaine image“.

En annexe de ces travaux, les auteurs ont conduit une étude pour voir quelles transformations d’images préservent la note d’évaluation donnée par le réseau. Leur objectif était de multiplier la base d’apprentissage en démultipliant le nombre d’images annotées par des transformations respectueuses de l’évaluation. Pour cela ils accordent une note (*LP* dans le tableau 8.4) qui décroît de 1 à 0 lorsque la distorsion réduit l’évaluation. On en déduit que seuls la réflexion, le changement d’échelle et le bruit faible assurent une bonne conservation de la note. Par suite, une campagne de tests montre que l’augmentation de données d’apprentissage à l’aide de ses 3 transformations permet de gagner 2 ou 3 % dans les performances.

10. Les auteurs notent cependant que le choix d’une distribution gaussienne est mal adapté aux photos de très bonne ou de très mauvaise qualités dont les notes ont une distribution très dissymétrique.

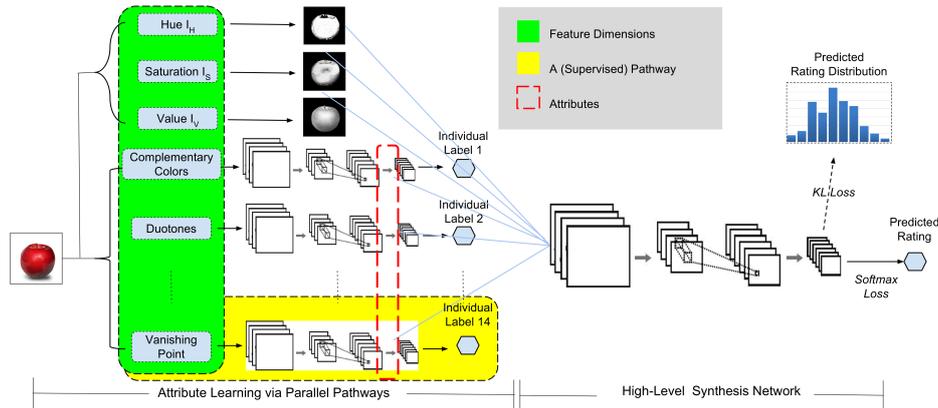


Figure 8.9 – Architecture du système BDN de Z. Wang et al., inspiré du cerveau humain (d’après [Wang et al., 2016]). En entrée, il met en œuvre un ensemble de détecteurs en parallèle, sensés détecter des primitives indépendamment les unes des autres. Hormis les 3 premiers qui mesurent les primitives chromatiques HSV, les autres canaux font appels à des RNP. Les sorties de cet étage sont elles-mêmes envoyées sur un autre RNP, dit de haut-niveau. Le réseau prédit également la distribution des notes (similaire à celle que donnerait AVA) à l’aide d’une fonction de coût de Kullback-Leibler.

transformation	Description	Note LP
réflexion	symétrie verticale	0,99
changement d’échelle aléatoire	choix d’un facteur unique dans $[0,9, 1, 1]$	0,94
écrasement/étirement	choix de 2 facteurs dans $[0,9, 1, 1]$	0,55
bruit faible	bruit gaussien $\in N(0, 5)$	0,87
bruit fort	bruit gaussien $\in N(0, 30)$	0,63
modification de la couleur	distorsion indépendante des canaux RVB	0,10
rotation	transformation affine aléatoire	0,26

Table 8.4 – Effet de diverses transformations sur le maintien de la note d’appréciation d’une photographie. Le facteur LP vaut 1 lorsque la note demeure inchangée. L’évaluation ne résiste qu’à la réflexion et au changement d’échelle, ainsi, dans une moindre mesure au bruit de faible valeur [Wang et al., 2016].

Le système proposé a été implémenté à l’aide du package Cuda-Convnet ¹¹

8.2.4 Porter attention à la construction de l’image : le système A-Lamp-CNN

Ces travaux se placent dans la lignée du système DMA-Net (voir section 8.1.2) qui utilise plusieurs voies en parallèle. Il a été présenté dans [Ma et al., 2017], son nom (A-Lamp CNN, *Adaptive layout-aware multi-patch*) exprime ses propriétés essentielles : d’une part il procède à une estimation de la composition (*layout*), d’autre part il mène en parallèle une mesure de l’esthétique sur plusieurs zones examinées (*multi-patch*) (voir la figure 8.10). L’analyse de la composition s’inspire de nombreux travaux consacrés à la distribution de zones d’intérêt dans une image (par exemple [Liu et al., 2010] ou [Obrador et al., 2010]). Elle utilise un réseau de détection adaptative de zones d’intérêt dont les résultats alimenteront également l’autre voie de traitement. Chaque zone d’intérêt, définie par sa boîte englobante, est un nœud du graphe de représentation, auquel sont affectés des attributs locaux et des attributs globaux. Les attributs locaux décrivent la position relative d’une zone d’intérêt par

¹¹. Cuda-Convnet est un package développé par l’Université de Toronto à partir de 2011 permettant de simuler tout type de RNP. Il est écrit en C++ et parallélisé en Cuda [Krizhevsky et al., 2012].

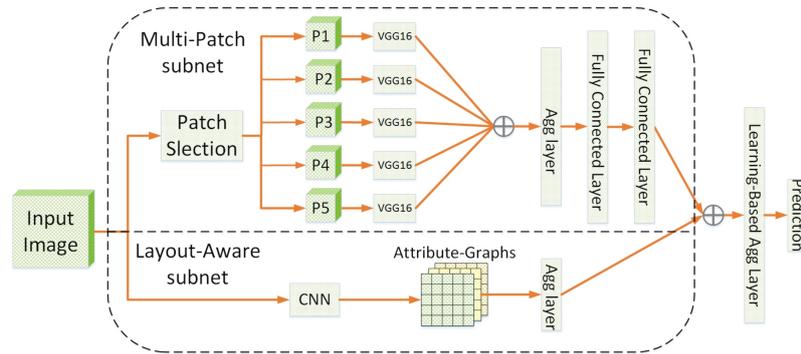


Figure 8.10 – Architecture du système A-Lamp mettant en évidence, dans la partie supérieure, le sous-réseau multi-patch qui traite en parallèle des sous-images à pleine résolution et dans la partie inférieure, le réseau traitant la composition (d’après [Ma et al., 2017]). Un réseau de détection adaptative de zones d’intérêt (pas représenté sur cette figure) est placé immédiatement après la lecture de l’image d’entrée et alimente les deux parties du traitement.

rapport à chacune des autres (distance, orientation du lien, recouvrement), tandis que les attributs globaux représentent la position de chaque nœud dans la scène.

La voie multi-patch est identique au DMA-Net de [Lu et al., 2014a] ou de [Lu et al., 2015a], mais au lieu d’utiliser des sous-images en position fixe ou en position aléatoire, la voie de traitement parallèle en pleine résolution traite les sous-images sélectionnées par le réseau de détection adaptative.

Les auteurs démontrent que cette voie multi-patch utilisant des zones pré-sélectionnées est préférable aux versions à sous-images fixes et aléatoires. Ils montrent aussi que le réseau complet, avec ses deux voies, donne de meilleurs résultats de classement binaire (“belle” ou “pas-belle”) que les autres architectures proposées antérieurement.

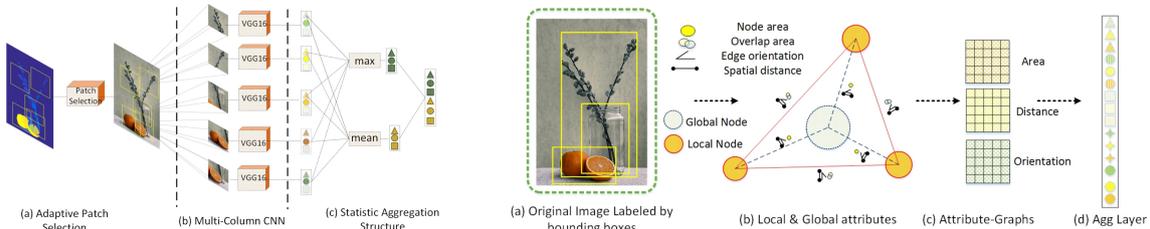


Figure 8.11 – A gauche le sous-réseau multi-colonne traite en parallèle, et en pleine résolution, les zones d’intérêt détectées par le réseau de détection adaptative de régions. A droite, le réseau traitant la composition à l’aide du graphe d’attributs (d’après [Ma et al., 2017]).

Ces travaux utilisent l’architecture Multi-Patch VGG16 (voir note 2) pré-entraîné sur ImageNet et le logiciel CAFFE (voir note 5). La détection préalable des zones d’intérêt suit les travaux de [Zhang et al., 2016].

8.2.5 Extraire une image “dramatique” d’un panorama : le système *Creatism*

Le système *Creatism* a été présenté dans [Fang and Zhang, 2017]. Il propose une chaîne de traitement complète pour extraire d’un panorama à 360 degrés, la plus belle partie du paysage, puis lui appliquer une succession de traitements pour en accroître l’attrait : renforcement des contrastes, application d’opérateurs HDR d’extension de dynamique, de filtres de vignettage pour assombrir les

bords, distorsion des gammes de gris et de couleur, puis application d'un filtre dit "dramatique" pour obscurcir les nuages et renforcer les contrastes.

La méthode repose sur l'hypothèse que la fonction de qualité esthétique se décompose en un produit de fonctions portant sur des aspects précis (contraste, couleur, etc.) pouvant être optimisées séquentiellement.

La première étape consiste à rechercher la sous-image la plus attrayante. Pour cela on divise le panorama en 6 images de 90 degrés de champ, avec donc un recouvrement des champs pour éviter de couper une zone d'intérêt potentiel. On procède ensuite par tirage éleatoire d'une fenêtre dans chaque image et on fait évoluer cette fenêtre en ajoutant ou retranchant des bandes sur les bordures, puis en rééchantillonnant la fenêtre de façon qu'elle demeure carrée. On la réévalue à travers un RNP et on itère le processus par descente de gradient.

Lorsque la meilleure fenêtre est isolée, on procède successivement aux opérations d'amélioration, en faisant évoluer les paramètres de chaque amélioration (l'étirement du contraste, la distorsion des couleurs, ...), en utilisant la boîte à outils Snapseed¹² et le RNP dans une configuration générative adverse (GAN) (voir section 8.1.5).

Le système Creatism a été testé par des professionnels qui ont confirmé sa capacité à trouver de belles sous-images et à les modifier de façon esthétique. Un exemple est présenté sur la figure 8.12.



Figure 8.12 – Le système Creatism, partant du panorama à 360 degrés à droite, recherche la sous-scène la plus esthétique, puis lui applique une succession de traitements pour en modifier le contraste, la couleur, la dynamique, etc. de façon à lui donner un caractère "dramatique". Toutes ces opérations sont conduites par un RNP, qui est en configuration générative (RNP-GAN) pour les modifications des paramètres d'apparence. Le résultat est l'image à gauche. (d'après [Fang and Zhang, 2017]).

8.2.6 Exploiter la distribution des évaluations

H. Talebi et P. Milanfar se sont penchés sur l'évaluation de l'esthétique d'une part et de la qualité des photos d'autre part. Nous ne nous intéressons ici qu'aux travaux concernant l'esthétique. Ils ont développé un système appelé NIMA (*Neural image assessment*) qui attribue à chaque photo une note entre 1 et 10, plus précise que la simple affectation à une classe "belle" ou "pas belle" [Talebi and Milanfar, 2017].

Pour cela, ils exploitent la totalité de la base AVA, sans en écarter les photos de qualité moyenne comme la plupart des autres auteurs. De plus ils exploitent une information disponible dans cette base mais qui n'a pas beaucoup été utilisée par les autres travaux¹³ : la distribution des notes affectées à chaque photo (en moyenne 200 notes, voir la figure 6.2 à gauche). Lors de l'étape d'apprentissage

12. Snapseed est une boîte à outils de traitement d'image créée en 2011, initialement développée pour Windows, maintenant disponible sous Android et IOS. Snapseed est la propriété de Google : <https://snapseed.fr.softonic.com>.

13. Cette information a été prise en compte dans les travaux de B. Jin, M.V. Ortiz-Segovia et S. Süsstrunk [Jin et al., 2016a] qui attachent à la variance des notes affectées à une photo une notion de difficulté de jugement esthétique (plus la variance est forte, plus les avis des juges sont dispersés). Dans cet article, les auteurs proposent, de plus, de pondérer les échantillons par un poids inversement proportionnel à la fréquence de la note attribuée, de façon à éviter un sur-apprentissage des photos de qualité moyenne, ce que la plupart des auteurs traitent en excluant, lors de l'apprentissage, les photos dont les notes sont comprises entre 4 et 6 (poids binaires).

des poids du réseau, c'est vers l'estimation de cette distribution qu'ils cherchent à converger, à l'aide d'une optimisation par une méthode de transport optimal (*Earth Mover's Distance*). La note unique finalement retenue pour une photo est, comme dans les autres études, la moyenne des notes données par cette distribution, comprise donc entre 1 et 10. Les résultats obtenus montrent un accord assez satisfaisant avec la note des juges humains (une erreur moyenne de 5 % et max de 12 % sur les 27 exemples présentés) quoique la corrélation moyenne sur la base AVA soit médiocre (de l'ordre de 0,6). Repassant à une classification en deux classes par seuillage, ils donnent des performances de plus de 80 % de bonnes réponses, proches de l'état de l'art.

La bibliothèque logicielle utilisée est TensorFlow (voir note 9) et le RNP est pré-entraîné sur la base ImageNet de façon à accélérer la phase d'apprentissage. Trois architectures différentes sont comparées : VGG16 (voir note 2), Inception-v2¹⁴ et Mobile-Net¹⁵. Elles donnent des résultats assez semblables, les meilleurs étant cependant obtenus avec Inception-v2.

Une approche très semblable est proposée par [Murray and Gordo, 2017]. Elle permet aussi de prédire :

- la classification en deux classes par seuillage à 5,
- la note esthétique comme moyenne de la distribution des scores,
- la distribution des scores (répartie sur 10 niveaux).

La méthode d'optimisation adoptée se distingue par un critère exploitant une norme de Huber sur les distributions (quadratique en dessous d'un seuil, linéaire au dessus).

8.2.7 Formuler une appréciation littérale

Dans le système appelé NAIR *Neural Aesthetic Image Retriever* [Wang et al., 2018], une tentative est faite d'associer dans un même système deux objectifs :

- d'une part séparer les images en haute et basse qualité comme la plupart des autres systèmes,
- d'autre part associer un commentaire littéral explicitant les raisons de la décision.

Ce second objectif ressemble à celui que nous avons décrit au paragraphe 7.11 et mis en œuvre à l'aide de primitives et de classifieur dans [San Pedro et al., 2012].

Pour la première partie, on utilise la base de données AVA et un réseau classique (modèle Inception-v3 et logiciel TensorFlow). La seconde partie tire profit des nombreux commentaires associés à la base AVA et regroupés dans un corpus appelé AVA-Comments, qui couvre plus de 20 000 photos. A l'aide de ce corpus, on entraîne un réseau capable de créer un commentaire littéral inspiré par la description de l'image, suivant le schéma proposé dans [Vinyals et al., 2015]. La tâche est délicate car les avis sont rédigés en langage naturel, ils comportent des termes non normalisés et des valences positives ou négatives issues soit du vocabulaire, soit de la syntaxe.

Nous notons, en examinant la figure 8.13, que ce très louable effort de fournir une justification de l'évaluation reste à améliorer, les termes employés étant peu précis et pas très pertinents.

Des travaux utilisant les mêmes données sont présentés dans [Hii et al., 2017]. Ils tirent également parti de AVA-Comments, mais avec comme objectif seulement d'améliorer la classification binaire. La méthode de représentation par mots utilise l'algorithme GloVe (*global vector*) [Pennington et al., 2014]. Chaque image est décrite par 100 mots appartenant au vocabulaire complet de 20 000 mots extrait des commentaires. Un réseau récurrent à deux couches est alors utilisé, actif sur les termes textuels. Ses résultats sont concaténés à ceux du réseau multiGAP utilisé pour la partie visuelle. L'ensemble fournit un classifieur à 2 états.

On trouve également dans [Ghosal et al., 2019] l'utilisation de ces informations de commentaires ou de thèmes issus de la base AVA, mais ici pour engendrer automatiquement des commentaires

14. Inception-v2 exploite une parallélisation des étapes de convolution et de mutualisation et remplace la couche de connexion complète par une mutualisation par maximum (*max pooling*), réduisant ainsi notablement le nombre de paramètres à apprendre [Szegedy et al., 2015].

15. Mobile-Net est un RNP dédié aux architectures embarquées de vision. Il se distingue par les étapes de convolution qui sont remplacées par des filtres séparables, particulièrement rapides [Howard et al., 2017].



Ground-truth aesthetic score: 4.2
(Low-aesthetic category)
Ground-truth comments: Would help if the picture was larger.
Prediction: Low-aesthetic category
Generated comments: The image is too small.



Ground-truth aesthetic score: 6.1
(High-aesthetic category)
Ground-truth comments: Great detail on this fine bird nicely separated from background A very good image
Prediction: High-aesthetic category
Generated comments: Great detail in the feathers.
Ground-truth aesthetic score: 5.6
(high-aesthetic category)
Ground-truth comments: Love the light but the leaves look a bit soft, the composition is also very nice.
Prediction: High-aesthetic category
Generated comments: The lighting is a bit harsh.



Ground-truth aesthetic score: 5.5
(high-aesthetic category)
Ground-truth comments: This is beautiful. The rim lighting on that plant is perfect.
Prediction: High-aesthetic category
Generated comments: The simplicity of this shot.



Figure 8.13 – Résultats d'évaluation de l'esthétique de quatre images de la base AVA par le système NAIR : appréciation de l'esthétique d'une part et expression littérale de l'évaluation d'autre part, comparées l'une et l'autre à leur "vérité" dans la base AVA (d'après [Wang et al., 2018]).

"esthétiques", en combinant une détection automatique de régions et un apprentissage de mots significatifs à l'aide de méthodes de traitement de texte.

8.2.8 Personnaliser l'évaluation

Le travail présenté dans [Park et al., 2017] par K. Park *et al.* est singulier et particulièrement intéressant car il cherche à donner une appréciation particulière, conforme aux choix d'un observateur précis.

Le principe consiste à utiliser un algorithme d'ordre et non de classification binaire. L'ordre prend en compte deux critères simultanément : une préférence générale en deux classes ("beau"/"pasbeau", comme dans la plupart des algorithmes) et des goûts particuliers obtenus à partir de réponses faites par l'observateur à un petit jeu de test. L'objectif est alors de fournir une classification qui respecte autant que possible l'avis général, mais en prenant en compte l'avis particulier.

La préférence individuelle se construit à partir, tout d'abord, du choix fait par l'observateur d'un échantillon préféré au sein d'une dizaine de paires, puis que par l'ordonnancement complet d'un ensemble de 10 images. La préférence est ensuite propagée à la base de données par des techniques de plus proches voisins.

L'implémentation logicielle est assez complexe. Tout d'abord on choisit de décrire chaque échantillon par un vecteur de dimension 4 096 obtenu à la sortie de l'avant dernière couche d'un RNP classique [Krizhevsky et al., 2012]. Puis on recherche une optimisation selon deux critères l'un exprimant l'accord à la classification globale, l'autre la régression selon les préférences individuelles. L'algorithme original, inspiré du *support vector regression* (SVR) d'une part et du *ranking support vector machine* (R-SVM) d'autre part est appelé *support vector regression* (SVR). Il est testé sur la base AVA. Les premiers résultats laissent place à de possibles progrès. Il semble en particulier nécessaire de bien comprendre comment les goûts individuels peuvent être généralisés de façon convaincante à partir d'un tout petit nombre d'apprentissages (condition nécessaire pour rendre le projet acceptable par l'utilisateur).

Chapitre 9

Analyse critique des techniques d'apprentissage automatique

9.1 Popularité des études sur l'esthétique

Nous allons tout d'abord rapprocher deux chiffres : d'une part celui des articles publiés sur le thème de l'esthétique 9.1 à gauche, d'autre part celui des publications qui associent les deux mots à priori assez antithétiques "Esthétique" et "Ordinateur" figure 9.1 à droite. Si les études sur l'esthétique restent très nombreuses (plus de 50 000 par ans depuis plus de 10 ans), les travaux mêlant esthétique et ordinateur croissent de façon beaucoup plus forte. Ils constituent plus de la moitié des travaux sur l'esthétique en 2018. Il y a donc aujourd'hui une bonne acceptation du milieu intellectuel à associer ces deux concepts dans un même raisonnement, ce qui était très rare il y a 15 ans.

Comme nous l'avons vu par ailleurs, même si quelques travaux sont encore publiés sur des approches algorithmiques (essentiellement issues des travaux de Birkhoff), depuis 15 ans ce sont les approches par apprentissage qui mobilisent la communauté.

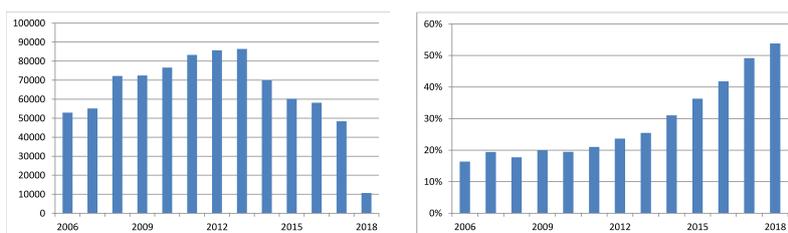


Figure 9.1 – *A gauche, recensement des références trouvées lors de la requête "Aesthetics" dans Google Scholar, par année de publication. A droite pourcentage des publications précédentes qui répondent au double critère "Aesthetics + Computer".*

9.2 Bilan des méthodes d'apprentissage

Ces systèmes font l'objet du tableau 9.1 où sont comparés certains de leurs attributs. Il montre bien comment les systèmes à base de primitives et de classificateurs, qui ont ouvert le chemin de l'esthétique par ordinateur, ont progressivement cédé la place aux techniques à base de réseaux de neurones. Notons cependant que des techniques spécifiques de classification peuvent encore parfois être employées pour des tâches particulières en aval d'un traitement par RNP comme dans [Park et al., 2017].

référence	base de données	type d'évaluation	primitives	classifieur
[Ke et al., 2006]	DPChallenge	2 classes	attributs génériques	Bayes, AdaBoost
[Luo and Tang, 2008]	DPChallenge	2 classes	primitives photo	Bayes, SVM, AdaBoost
[Datta and Wang, 2010]	Photo.net	2 classes+note	primitives photo	SVM
[Marchesotti et al., 2011]	Photo.net+CUHK	note	attributs génériques	PCA+FV+SVM
[Dhar et al., 2011]	DPChallenge	2 classes	primitives haut niveau	SVM
[Yao et al., 2012]	Photo.net	recommandations	composition,texture,color	ACQUINE+ SIMPLicity
[Lo et al., 2012]	CUHK	2 classes	primitives photo	SVM par catégorie
[San Pedro et al., 2012]	DPChallenge	note continue	opinions internautes	SV- ϵ
[Su et al., 2012]	DPChallenge	2 classes	bibliothèque esthétique	AdaBoost
[Lu et al., 2014a]	AVA	2 classes	HSV ou Munsell	LDA + Lasso
[Lu et al., 2014b]	AVA	2 classes	photo + ss-photo	RNP - SCNN + RD-CNN
[Schifanella et al., 2015]	BEAUTY	5 niveaux	primitives photo	PLSR
[Jin et al., 2016b]	AVA	2 classes	image réduite	RNP = ILGNet
[Kong et al., 2016]	AADB + AVA	2 classes	image réduite + catégorie	RNP "siamois"
[Mai et al., 2016]	AVA	2 classes	pyramide + info class.	RNP multitâche MNA
[Schwarz et al., 2016]	Tübingen	classe ment 2 à 2	image : 1 piste	RNP siamois + triplets
[Kao et al., 2016]	AVA	2 classes + distrib.	image + 3 catégories	A&C RNP
[Wang et al., 2016]	AVA - AVA2	2 classes + distrib.	histo. de notes.	Brain-inspired RNP
[Redi et al., 2017]	AVA \rightarrow Redi	3 classes	apprentissage images web	RNP
[Kao et al., 2017]	AVA	2 classes	esthétique + catégorie	RNP multitâche
[Kairanbay et al., 2017]	AVA - Comments	2 classes	images + style	RNP + couches GAP
[Ma et al., 2017]	AVA	2 classes	composition + ss-images	A-Lamp, multitâche
[Murray and Gordo, 2017]	AVA	note continue	histo. de notes.	RNP= ResNet
[Park et al., 2017]	AVA + interact.	choix personnalisé	préférences personnelles	RNP +SVR
[Talebi and Milanfar, 2017]	AVA	note continue	histo. de notes.+catégorie	RNP multitâche
[Srivastava and Kant, 2018]	AVA	2 classes	image réduite+couleur LAB	RNP = ILGNet
[Wang et al., 2018]	AVA - Comments	explications	AVA-Reviews	RNP = Inception

Table 9.1 – Les divers systèmes d'apprentissage automatique de qualité de photographie dans l'ordre de leur apparition. On recouvre ici sous le terme "primitives photo" les primitives qui sont spécifiquement déduites de règles photographiques : règle du tiers, mise au point sur l'objet et flou du fond, importance de la construction, etc. Les primitives de "haut niveau" sont calculées dans l'image, tandis que les "catégories" sont issues des données, souvent comme métadonnées, parfois calculées par le RNP. Les études utilisant des RNP sont présentées sur un fond grisé.

9.2.1 Quelles architectures ?

Sans surprise, nous constatons que les travaux présentés se contentent souvent d'exploiter des architectures classiques ayant fait leurs preuves par ailleurs dans les compétitions internationales de reconnaissance des formes. Les architectures ainsi que les logiciels sont alors les meilleurs à l'instant t et la génération qui triomphera à la prochaine compétition les remplacera. Pour les études sur l'esthétique, une attention particulière est portée sur le format des images et concerne aussi bien la résolution que le rapport d'aspect, mais là aussi les solutions adoptées sont celles retenues pour d'autres problèmes. Les auteurs ne semblent pas rencontrer de problème à traiter des images de taille très modeste (la base AVA universellement utilisée pour l'apprentissage ne dispose que de réductions des images originales de DPChallenge de taille 640×480 .)

Certains développements des RNP sont suivis avec un grand intérêt par la communauté de détermination de l'esthétique par RNP : les réseaux multitâches et les réseaux siamois surtout, mais aussi les réseaux récurrents. On met ainsi en parallèle des tâches très différentes comme celles qui traitent l'image elle-même d'une part et celles qui traitent les métadonnées ou qui cherchent à les déduire de l'image : catégorie, style, contexte, etc.

Au contraire de la tendance observée en 2018 dans les applications les plus générales des RNP qui privilégient une multiplication de couches convolutives, l'utilisation conjointe de couches convolutives et de couches totalement connectées demeure la règle dans les études sur l'esthétique et les architectures adoptées sont quasiment toutes "en entonnoir" plutôt qu'en "yoyo" comme c'est la tendance dans d'autres domaines.

L'apprentissage du réseau reste un point délicat. Compte tenu de la grande profondeur des réseaux actuellement employés, les bases de données dédiées à l'esthétique apparaissent trop réduites pour

δ	SVM+Gauss. mixt. [Murray et al., 2012]	SCNN [Lu et al., 2014b]	RD-CNN [Lu et al., 2014b]	DMA-Net [Lu et al., 2015b]	MNA-CNN [Mai et al., 2016]
0	66.7 %	71.20 %	74.46 %	75.41 %	77.4 %
1	67.0 %	68.63 %	77.70 %	-	76.5 %

δ	ILGNet [Jin et al., 2016b]	BDN [Wang et al., 2016]	MTRL-CNN [Kao et al., 2017]	A-LAMP [Ma et al., 2017]	NIMA [Talebi and Milanfar, 2017]
0	79.25 %	76.80 %	79.08 %	81.7 %	81.5 %
1	-	76.04 %	77.71 %	-	-

Table 9.2 – Résultats comparatifs de diverses méthodes d'évaluation esthétique sur la base de données AVA, appliquées à une classification à deux classes “haute qualité” / “basse qualité”. La méthode A-LAMP [Ma et al., 2017] donne les meilleurs résultats (en 2017). Le paramètre δ est le même que le delta en abscisse de la figure 7.10, d'après [Kao et al., 2017].

conduire un apprentissage de bout en bout (c'est en particulier le cas de AVA). Le recours à des apprentissages “universels” effectués sur des bases d'images banalisées pour des tâches très génériques de reconnaissance des formes est donc aujourd'hui généralisé. Contrairement à ce que l'on pourrait craindre, cela ne semble pas affecter les performances des études sur l'esthétique..

9.2.2 Quelles performances ?

La communauté utilise aujourd'hui la classification à 2 classes (“belle-image” vs. “pas belle image”) sur la base AVA comme banc de test des performances des systèmes¹.

On discutera plus loin de la pertinence de ce critère.

Ces résultats sont présentés sur le tableau 9.2 qui est repris sur la figure 9.2

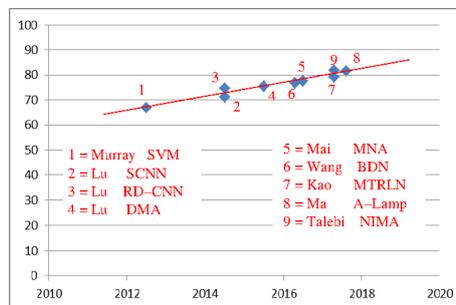


Figure 9.2 – Performances d'évaluation esthétique à 2 classes (belle / pas-belle) des systèmes, sur la base AVA, avec apprentissage automatique, depuis 10 ans. Notons que seul le système de Murray et al. a été retenu dans ce “podium” qui regroupe les meilleurs systèmes, comme système utilisant une sélection de primitives suivie d'une classification. Tous les autres utilisent des RNP.

On constate que ces performances de classification sont aujourd'hui bonnes, statistiquement, pour de grands ensembles de photos (plus de 75 % de bonne classification entre “haute qualité” et “basse qualité” sur des ensembles de plusieurs centaines de milliers d'images). Par ailleurs qu'avec quelques connaissances supplémentaires sur la catégorie de la scène observée ou sur le style de la photographie, on peut porter ce taux de reconnaissance à 85 %. Ces travaux étant encore très préliminaires, on peut s'attendre à ce qu'ils croissent encore à court terme.

On constate aussi que les méthodes proposées s'adaptent bien à une grande diversité de photos qui se distinguent aussi bien par leur style que par leur format.

1. On a cependant signalé les variantes possibles de ce critère : d'une part AVA a donné naissance à plusieurs sous-bases selon qu'une image est ou non associée à certaines méta-data, d'autre-part, le critère de “belle-image” est souvent lié à un paramètre (soit la distance entre classes, soit le pourcentage d'images retenues (cf. section 6.2).

Les études qui visent à transporter les apprentissages d'une base de données vers une autre concluent plutôt à une bonne robustesse de ces apprentissages [Kao et al., 2017] quoique des travaux donnent des résultats opposés [Kong et al., 2016].

On ne peut cependant manquer de souligner une critique majeure de ces résultats. La classification à deux classes est un problème intellectuellement peu excitant et socio-économiquement sans grand intérêt. Son attrait essentiel est d'offrir un juge de paix aux algorithmes. La plupart des travaux qui l'utilisent se fixent initialement un objectif plus ambitieux soit de conclure à une évaluation fine et non binaire de la mesure esthétique, soit de comparer entre elles un petit ensemble d'images, soit enfin de dégager des éléments permettant d'expliquer ou de justifier une notation.

Examinons ces autres objectifs et la façon dont ils ont été abordés.

Evaluation continue : hormis l'éphémère système ACQUINE [Datta and Wang, 2010], aucun système ne propose une note fiable, reflétant une évaluation continue de l'esthétique. Plusieurs systèmes se proposent de prédire une répartition de notes telle que pourrait la fournir un jury AVA et de là en déduisent une évaluation moyenne [Talebi and Milanfar, 2017, Murray and Gordo, 2017], mais ces résultats sont plus présentés comme des sous-produits, éventuellement exploitables, que comme la sortie principale du système. La comparaison, sur une base d'images tests issue d'AVA, de la note moyenne déduite du jugement du RNP et de la note moyenne de l'évaluation AVA, laisse la place à des progrès importants.

Comparaison entre images : c'est aussi un objectif souvent évoqué mais très peu traité. C'est cependant explicitement l'objectif de l'étude présentée dans [Schwarz et al., 2016]. La méthode proposée est très intéressante, les résultats préliminaires présentés sont encourageants mais au prix d'une complexité certaine du modèle mis en place. Il sera particulièrement intéressant de suivre sur ce point les résultats de la compétition SPARK² dont nous avons parlé plus haut, où, pour la première fois, un algorithme sera confronté à un jury de photographes professionnels pour dépouiller une compétition.

Justifier la notation : les techniques à base de classifieurs et de primitives sont assez bien adaptées pour dégager les causes essentielles de l'estimation esthétique puisque l'on peut tracer parmi les critères adoptés ceux qui sont effectivement actifs. Malheureusement ils ont montré des performances trop faibles pour que cette piste soit poursuivie. Les méthodes à base de RNP tournent le dos à une justification de leurs résultats puisqu'elles cachent leurs décisions dans l'épaisseur des couches successives. Traditionnellement, on essaie tout de même d'expliquer les décisions en révélant des couches proches de la sortie où s'affichent les zones activées juste avant la décision. C'est ce que font, sans grand succès³, plusieurs études [Murray and Gordo, 2017, Kairanbay et al., 2017, Jin et al., 2016a]. Plus raisonnables sont les arguments des travaux qui s'appuient sur des expressions littérales, comme ceux présentés dans [San Pedro et al., 2012, Wang et al., 2018]. Leurs résultats sont cependant encore loin de convaincre. Si la méthodologie mise en place semble tout à fait adaptée, on peut s'interroger sur leur base de travail, les avis des "experts" de la base AVA. Est-ce une matière assez fiable et assez bien formalisée pour soutenir le raisonnement proposé ?

2. "Spark" (Le concours *SPARK, a RenAIssance photographic contest*) de Huawei, en promotion de son modèle P20 Pro de téléphone portable, <https://consumer.huawei.com/uk/campaign/sparkrenaissance/> (juillet 2018).

3. Lorsque l'on examine attentivement la localisation des zones actives, il apparaît de nombreuses situations où ces zones sont "à côté" des éléments qui devraient retenir l'attention. Bien sûr ce ne sont pas les cas présentés dans les articles qui illustrent ce défaut, encore que l'on peut longuement discuter les résultats de [Kairanbay et al., 2017] ou de [Jin et al., 2016a].

9.2.3 S'interroger sur les hypothèses

On ne peut manquer, à ce point de bilan, de revenir sur les hypothèses qui soutiennent ces travaux. Nous regretterons tout d'abord que ces hypothèses soient très rarement formulées et bien sûr encore moins justifiées. L'argument de la supériorité des RNP pour les tâches de reconnaissance des formes et d'évaluation dans les domaines les plus divers tient trop souvent lieu de justificatif aux travaux entrepris en esthétique à l'aide de ces outils.

L'hypothèse la plus cachée est celle d'une esthétique objective et universelle. Il est pourtant clair dans tous ces travaux (hormis quelques très rares exceptions sur lesquelles nous reviendrons) que les auteurs mettent en place une estimation valable pour tous et en tout lieu, indépendamment de l'observateur, de sa culture, de son tempérament, de son humeur de ses émotions du moment. Ce jugement se constitue à partir d'une connaissance, reflet elle-même d'une appréciation universelle attachée à l'image et à elle seule. Ces informations sont tout entières contenues dans la base de données. Nous sommes très exactement dans un contexte que les informaticiens nomment "*data-driven*", le programmeur s'efforçant d'être aussi transparent que possible de façon que les conclusions soient universelles. Ce parti-pris, en esthétique, relève clairement de l'objectivisme tel que nous l'avons résumé au paragraphe 2.1.

Cette interprétation s'inscrit ainsi clairement à contre-sens des approches que l'on développe dans d'autres domaines de la connaissance, que ce soit en neuro-biologie (chapitre 3), en psychologie, sociale ou expérimentale (chapitre 4), en philosophie (paragraphe 2.2). Si le subjectivisme, position diamétralement opposée à l'objectivisme, n'est plus aujourd'hui la règle comme il a pu l'être en philosophie et malgré ses nombreux supporters [Zeki, 1999, Ishizu and Zeki, 2011, Changeux, 2016], il est quand même communément mis à contribution (en particulier dans l'interprétation interactionniste qui rassemble aujourd'hui beaucoup de supporters [Reber et al., 2004, Ingarden, 2011] qui prônent que l'observateur avec ses états de conscience et d'attention est un déterminant majeur de l'appréciation esthétique qui "exploite" le signal envoyé par l'objet, l'autre déterminant.

Les approches proposées en intelligence artificielle pour aborder l'évaluation de l'esthétique des photos ne laissent aujourd'hui pratiquement aucune place au subjectivisme. Elles évacuent ainsi totalement des notions pourtant reconnues comme très importantes de sensibilité, d'éducation, de culture, de contexte.

Soulignons pourtant deux tentatives de laisser une petite marge d'influence à la personnalité de l'observateur dans les travaux.

- Dans les travaux de [Kong et al., 2016], on garde trace des évaluateurs ayant apporté une note et l'on est ainsi capable de décliner une esthétique "à la façon de" tel ou tel expert. Ainsi, le subjectivisme est rattrapé en plaçant un observateur au sein d'une boîte particulière, en compétition avec d'autres observateurs dotés de sensibilités différentes. L'utilisateur dispose alors de la capacité de reconnaître "son" esthétique en identifiant le Beau qui lui convient à celui reconnu comme tel par l'un des observateurs. C'est une première façon de reconnaître des points de vue un peu différents sur un œuvre.
- Les travaux de [Park et al., 2017] vont plus loin car ils laissent la place à une préférence personnelle en mêlant, après les couches de réseau de neurones, les résultats de classification obtenue (non supervisée et universelle) et les résultats d'une évaluation subjective faite après une étape d'interaction avec l'utilisateur. C'est donc une démarche qui, comme celles qui mettaient en œuvre des classifieurs et des primitives, pourrait prendre en charge la richesse et l'autonomie de l'observateur. Notons cependant qu'il reste beaucoup de travail à faire pour atteindre des résultats convaincants avec cette approche, tant il est peu clair ce qui fait la différence entre un observateur et un autre. Il semblerait que dans cette direction, les résultats les plus convaincants devraient hériter des approches de profilage actuellement exploitant toute la richesse des données personnelles disponibles sur les réseaux sociaux. Nous examinerons cette piste plus loin.

9.2.4 De quelle image une belle image est-elle l'image ?

Reprenons ici les mécanismes mis en œuvre dans les études les plus avancées aujourd'hui qui impliquent massivement les RNP et essayons de tirer quelques conclusions des choix retenus.

C'est donc dans la base de données qui sert d'apprentissage que se trouve l'essentiel de "l'expertise esthétique" déployée dans le RNP et c'est aujourd'hui la base AVA qui est universellement retenue pour cela. Rappelons qu'elle est constituée de plus de 250 000 photos, issues du site DPChallenge. A chaque photo sont associées d'une part des méta-données, parfois sommaires, d'autre part des appréciations données par des internautes sous forme de notes et, parfois de commentaires littéraires. Cette base de données se distingue donc de bases de photos orientées vers la vision par ordinateur, comme par exemple ImageNet ou ImageCLEF, ou des entrepôts d'images à finalité sociale (Flicker, Instagram), par l'accent mis sur les qualités esthétiques des œuvres et cela est très sensible lorsqu'on consulte le site DPChallenge dont elles sont issues.

Quelques remarques sur les photos de la base AVA

A notre connaissance il n'y a pas d'étude qualifiant cette base de données (excepté l'article où elle fut présentée [Murray et al., 2012]). Comme nous l'avons dit, il est important de noter qu'en première analyse, les images proposées sont généralement de bonne qualité technique, qu'elles couvrent un très large champ de thématiques, assez compatible avec ce que l'on trouve dans les galeries d'art, les musées ou les collections : portraits, paysages, natures mortes, scènes de foule, scènes d'intérieur ou de plein air, etc. Notons cependant que les photos en couleur sont très largement majoritaires, au delà de ce que l'on trouve parmi les œuvres d'artistes, même contemporains.

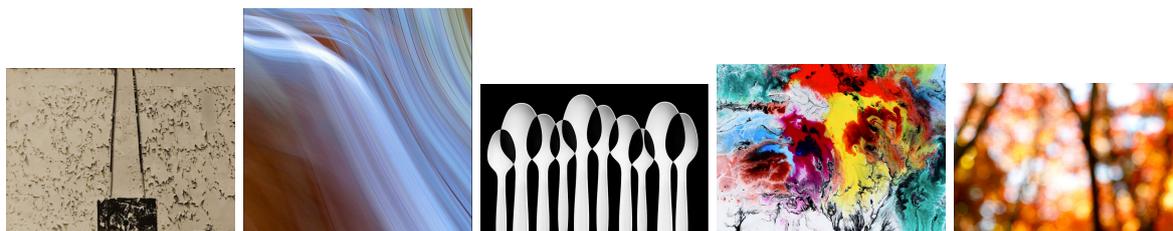


Figure 9.3 – Quelques exemples de photographies “abstraites” (ou faiblement figuratives), telles que l'on en trouve beaucoup dans les expositions actuelles : de gauche à droite, elles sont signées de Fernell Franco, Andrew Gustar, Jacqueline Hammer, Michael Lønfeldt et Les Cunliffe. Ce type de photos nous semble sous-représenté dans la base AVA et celles qui s'en approchent sont fréquemment mal notées par DPChallenge.

Il est bien évidemment très difficile de qualifier la représentativité statistique des thématiques, des styles, des compositions proposées. Il serait aussi difficile de fixer un objectif idéal de représentativité sans arrêter clairement le cadre de l'esthétique que nous souhaitons caractériser. Notons cependant que les scènes abstraites ou proches de l'abstraction (comme les œuvres de Man Ray par exemple), qui constituent une part importante des œuvres d'artistes photographes telles qu'illustrées en figure 9.3 sont trop peu représentées dans AVA. De même, il nous semble que les scènes incorporant un animal sont sur-représentées. Parmi les photos de nature, une place très grande est accordée à la thématique de l'eau et parmi les paysages, les couchers de soleil d'une part et les photos de nuit sont très présentes. Pour les scènes de la vie quotidienne, nous noterons une faible proportions d'images reflétant les situations de travail (ateliers, usines, champs, ...) tandis que les images de vie familiale sont au contraire nombreuses.

Concernant le style des photos, on ne peut manquer d'être frappé par l'importance des post-traitements numériques, visant à accroître la résolution, à étendre la palette de couleurs, à forcer les contrastes, souvent dans des proportions importantes. L'incrustation est également un artifice très utilisé, ainsi que le ralenti et la superposition d'images. C'est ce que nous appelons une esthétique technophile. Elle apparaît à la fois beaucoup plus prononcée que pour les images “réalistes” des pho-

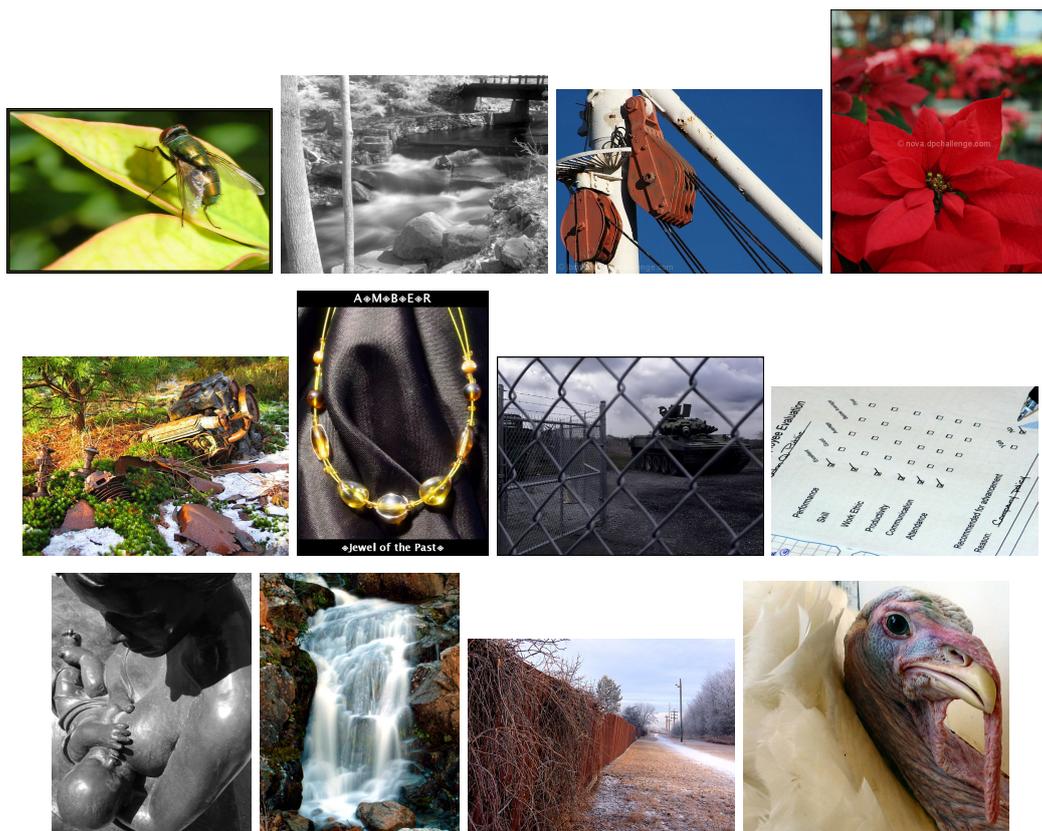


Figure 9.4 – 12 photos tirées aléatoirement dans la bibliothèque AVA parmi celles ayant reçu la note moyenne de 5,43, qui correspond à la moyenne des notes attribuées par les internautes. On note une très grande hétérogénéité parmi ces images qui rend difficile de les qualifier d'un point de vue esthétique. Une est tatouée d'une marque, une autre porte des caractères publicitaires, une troisième ne prend son sens que par l'interprétation du texte écrit, une ou deux présentent d'évidents défauts d'éclairage . . . Ces images assez critiquables sont aussi bien notées que d'autres dont les qualités esthétiques sont moins contestables.

tographes professionnels et beaucoup moins perfectionnée que chez les artistes qui s'en sont emparé.

Notons que peu d'œuvres utilisent les codes esthétiques des premiers temps de la photographie, (simplification des fonds, opposition des contrastes, linéarisation de la géométrie) comme ceux que l'on trouve dans les œuvres d'Adget, Cameron, Kertesz, Carjat, . . . dont l'esthétique est souvent dictée par les contraintes techniques, mais qui ont progressivement fait école par le dépouillement des scènes et la réduction des détails.

Il nous semble qu'il y a des enseignements à tirer de la représentativité de la base AVA sur ce point, mais que ces études sont à faire. L'étude entreprise dans [Murray et al., 2012] pour identifier chaque image dans un référentiel psychologique (en l'occurrence le modèle d'affects circomplexe de Russell⁴ [Russell, 1980]) est une excellente piste qui mériterait d'être approfondie mais devrait être conduite par des équipes pluridisciplinaires.

4. Le modèle d'affects circomplexe de Russell propose de distribuer les états psychologiques de façon ordonnée sur une circonférence organisée autour de huit affects principaux, régulièrement distribués sur le cercle : le plaisir, l'excitation, l'incitation, la détresse, le déplaisir, la dépression, la torpeur et la relaxation. Russell propose alors de rapporter toutes les autres émotions dans ce référentiel, en coordonnées polaires.

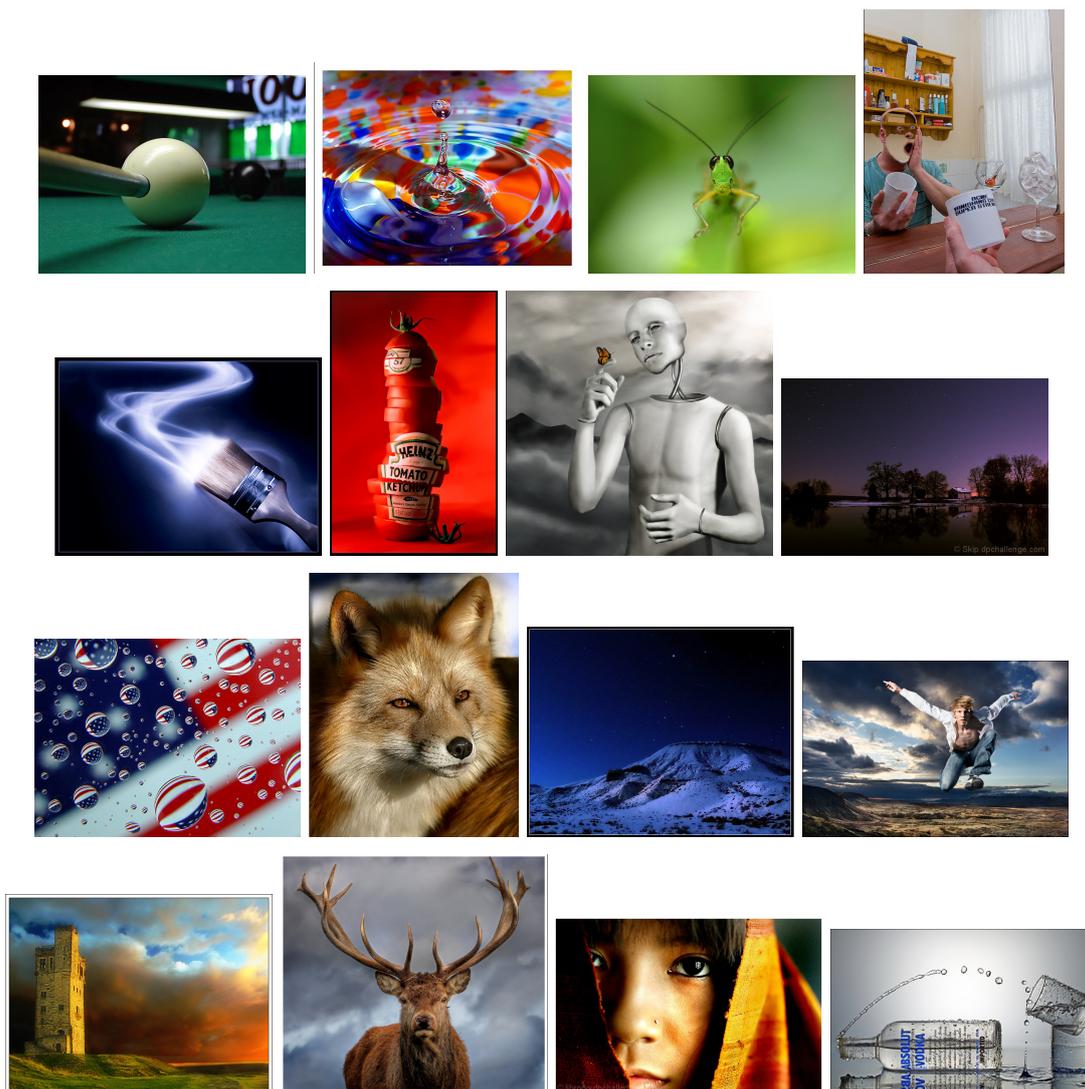


Figure 9.5 – Les 16 “plus belles” photos de la bibliothèque AVA par ordre décroissant de valeur moyenne des notes attribuées par les internautes. La meilleure photo, en haut à gauche, a pour moyenne 8,60. La suivante, à côté, 8,52. Toutes les notes moyennes de ces 12 photos sont supérieures à 8,25. On constate une grande diversité des styles et des inspirations. Etendant cet examen également aux photos suivantes du classement, on note comme sur ces 16, une très forte représentation des photos très énergiquement améliorées à l’aide d’outils de traitement des images ou par truquage. Leur inspiration pourrait être trouvée dans l’univers du design ou de la publicité (plus de la moitié de la sélection ci-dessus). On constate également une très forte représentation des photos animalières, ainsi que des paysages nocturnes et, proportionnellement à ce que l’on trouve dans les galeries et les musées, une faible représentation des portraits, des nus, de l’architecture et des scènes de rue. Elles ont pour numéros : 106, 9482, 150, 491369 - 55938, 642962, 543104, 267110 - 2892, 455890, 54599, 335951 - 106707, 455658, 111547, 957982.

Les notes de la base de données AVA

Si la base de photographies d’AVA n’appelle pas à des remarques importantes quant à son contenu, on peut être plus critique sur les évaluations qui sont associées à chaque photo. Ces évaluations se posent, dans les techniques à base de RNP, comme les sources premières d’information qui devraient reproduire le jugement humain. Remplissent-elles leur tâche ? La réponse n’est malheureusement pas très positive.

Notons tout d'abord que la plupart des études considèrent pour “vérité esthétique” la moyenne des notes attribuées. Force est de constater que cette moyenne ne mesure pas des choses comparables comme on s'en rendra compte en comparant les images, tirées aléatoirement parmi celles à qui a été attribuée la même note de 5,43 qui correspond à la moyenne des notes de la base AVA (figure 9.6). Certaines images (le questionnaire par exemple) n'ont d'autre qualité que de retenir l'attention par une pirouette humoristique ne relevant nullement du registre du beau. D'autres (le char derrière un grillage, la statue de femme allaitant) présentent d'évidents défauts de réglage qui les disqualifient d'un point de vue technique. Deux autres (le bijoux et le poinsettia) affichent des caractères superposés à l'image (volontairement ?) qui semblent incongrus dans une compétition photographique. Il reste ainsi juste la moitié des photos qui sont exemptes d'une élimination précoce que l'on trouverait dans tout concours.

Examinons maintenant les 16 photos qui obtiennent les “meilleures” notes parmi les 225 000 de la base (figure 9.5). Elles sont incontestablement toutes supérieures aux précédentes et aucune ne semble usurper sa note quoique chacune se distingue par des critères spécifiques. Nous voyons sur cette sélection l'importance des critères que nous avons qualifiés de “technophiles” précédemment et qui sont bien représentatifs de la communauté d'internautes qui participent au concours.

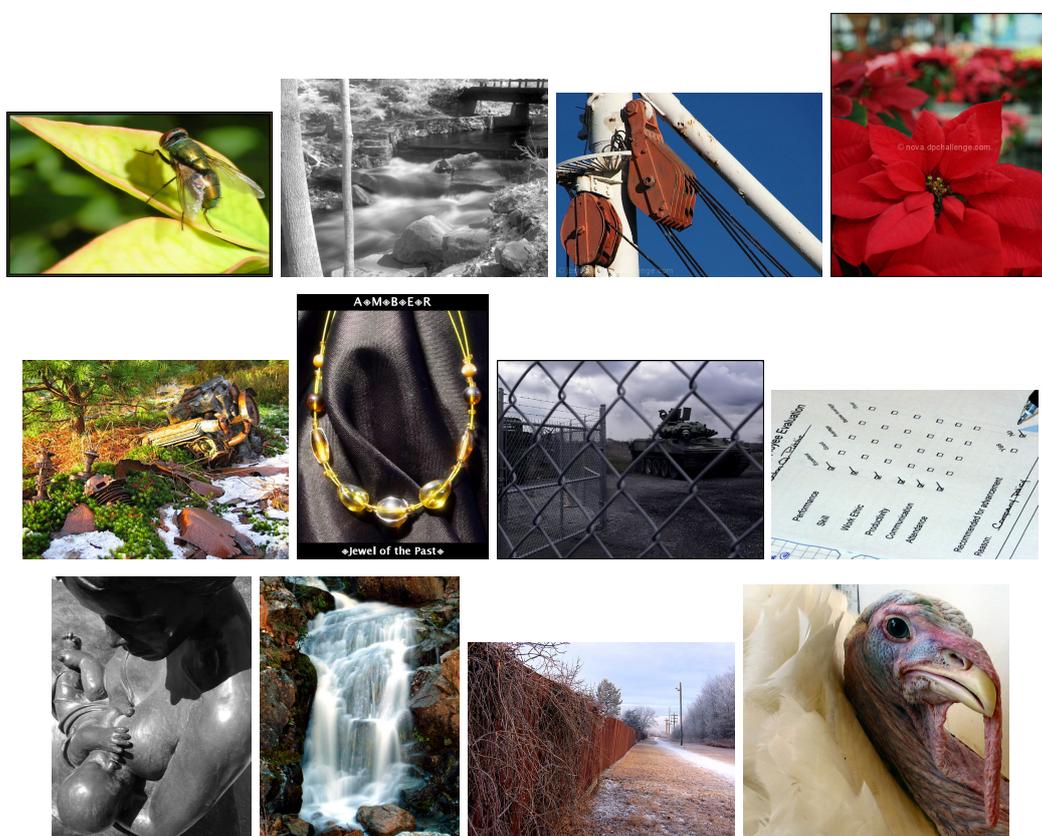


Figure 9.6 – 12 photos tirées aléatoirement dans la bibliothèque AVA parmi celles ayant reçu la note moyenne de 5,43, qui correspond à la moyenne des notes attribuées par les internautes. On note une très grande hétérogénéité parmi ces images qui rend difficile de les qualifier d'un point de vue esthétique. Une est tatouée d'une marque, une autre porte des caractères publicitaires, une troisième ne prend son sens que par l'interprétation du texte écrit, une ou deux présentent d'évidents défauts d'éclairage . . . Ces images assez critiquables sont aussi bien notées que d'autres dont les qualités esthétiques sont moins contestables. Elles ont pour numéros : 115821, 120486, 131330, 134515 - 135292, 169814, 173889, 232238 - 1133, 306322, 276462, 132804.

Nous ne présenterons pas les plus “mauvaises” photos d'AVA car plusieurs sont sans intérêt (champ vide ou à peine ombré), mais une sélection des photos les moins bien notées (figure 9.7, qui reflète

bien, là aussi, la diversité des critères qui ont pu être utilisés pour écarter les photos et l'ambiguïté d'une note unique pour caractériser l'esthétique. Cette impression de très forte diversité des façons d'interpréter la photo est renforcée lorsqu'on examine la dispersion des votes qui conduisent à la note. Ainsi, si toutes ces photos médiocres reçoivent majoritairement la note la plus faible (1), seules la photo du buisson (numero 45563) et la photo verte, abstraite (numero 92643) ne reçoivent aucune note 10 (la meilleure). Toutes les autres en reçoivent au moins 3, ce qui est pour le moins surprenant.

Nous serons plus critiques sur les évaluations associées à chaque photo. On sait qu'elles sont nombreuses (entre 100 et 500 selon la photo, plus de 200 en moyenne). On ne connaît pas individuellement ceux qui les ont formulées, et l'on ne dispose pas statistiquement de beaucoup d'information sur la communauté dont ils sont issus. On sait que DPChallenge est fréquenté par un public d'internautes amateurs souvent éclairés, parfois professionnels de la photographie. Leurs profils montrent une grande variété d'âge et de genre, mais il n'est pas clair qu'ils présentent une grande diversité d'origine géographique et culturelle. Il semble qu'il y ait une très forte représentation des villes moyennes de l'Amérique du Nord ainsi que de l'Australie, que les quadras ou quinquas, les mères de famille, les jeunes adolescents, tous disposant d'un peu de temps libre, soient bien représentés. Ils présentent une bonne familiarité avec la photographie et la plupart d'entre eux exposent également leurs photos. Ils contribuent très activement aussi bien aux divers concours qu'à l'évaluation, apportant souvent leurs appréciations à plusieurs milliers de photos. Mais les commentaires qu'ils postent sont trop souvent très vagues et peu informatifs, voire anecdotiques, laissant penser qu'ils sont peu familiers de l'argumentation esthétique où n'en maîtrisent pas les usages. Très majoritairement, les commentaires littéraux sont très positifs et donc pour les seules images bien notées. Ils se résument alors bien souvent à "J'aime!". Il est très rarement fait allusion à des critères esthétiques, ou alors sous une forme très générale : "Très belles couleurs" ou "Bonne maîtrise de l'espace". Les très belles photos de nature (couchers de soleil, forêts brumeuses, plages, sommets enneigés), de fleurs ou d'animaux, sont très sûrement plébiscitées de même que les portraits de vieillards ou d'enfants s'ils sortent des photos d'identité. Beaucoup de photos sont appréciées pour leur sujet cocasse, humoristique ou grave, plus que pour leurs qualités esthétiques comme en attestent les commentaires. Les photos abstraites dont l'interprétation n'est pas immédiate sont généralement mal cotées ou font l'objet d'avis très dispersés.

On ne peut manquer, en suivant les débats qui accompagnent les évaluations sur le site DPChallenge, d'être frappé par le rôle très important que prend le sujet de la photo au regard de la forme de cette photo. La dimension narrative de la photo est incontestablement une caractéristique importante dans la balance de l'évaluation, renforcée d'une part par le titre de la photo, qui est mis en bonne place et prend souvent une tournure très descriptive, et d'autre part par l'intitulé de la compétition à laquelle se présente la photo, cet intitulé s'imposant à l'évidence comme un cadre de lecture et d'interprétation de la photographie. Tout ceci nous renvoie aux remarques faites au paragraphe 1.1 où nous examinons les raisons d'apprécier une image ne relevant pas de l'appréciation esthétique.

Retenons enfin que l'évaluation finale dans AVA est généralement réduite à la moyenne des évaluations et l'on voit que ce qui va servir à entraîner le RNP est une "vérité" bien particulière, une sorte de beauté statistique, reflet d'un consensus raisonnable d'un petit nombre de "jurés" qui sans être tirés aléatoirement auraient du mal à être considérés comme des parangons de l'esthétique photographique.

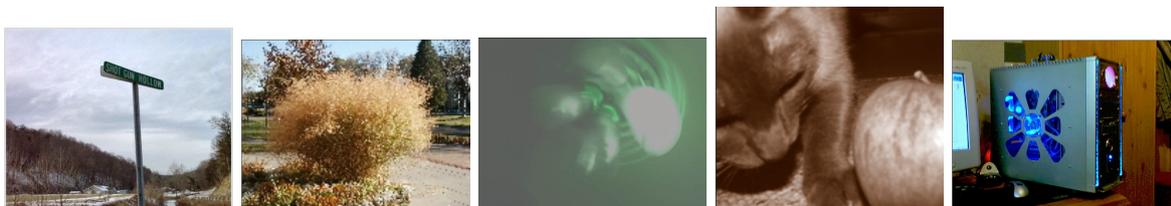


Figure 9.7 – 5 photos parmi les plus mal notées de la base AVA (leurs notes ne dépassent pas 2,15). Elles ont pour numéros : 11220, 45563, 92663, 142717, 221721.

Ainsi, le RNP, s'il fait bien son travail (et tout porte à croire qu'à brève échéance il le fera) qua-

lifera avec une grande fiabilité (95 % ? 98 % ?) les “belles photos”, c’est-à-dire des photos conformes au goût moyen d’un petit groupe d’aficionados, rassemblés autour d’une passion commune pour la photographie. Nous revoilà plongés dans l’univers de Bourdieu, plus que dans celui d’Arnheim ou de Berlyne.

9.3 Comment introduire de la “subjectivité” dans les RNP ?

Devant ces remarques, on peut chercher de trois façons différentes à échapper à ce jugement unique et “objectif” :

1. soit en poursuivant la piste de l’individualisation telle que l’ont proposée par exemple Park *et al.* [Park et al., 2017], mais nous avons vu que cette piste nécessite que l’on améliore les capacités des RNP à apprendre avec un tout petit nombre d’exemples, ce qui n’est pas envisagé à court terme ;
2. soit en multipliant les modèles de sensibilité, de façon que l’utilisateur puisse choisir “son” esthétique ;
3. soit en prenant en compte explicitement la spécificité de l’utilisateur lors de l’expérience, c’est ce que recouvre généralement le terme de “recommandation”.

Nous explorons ici les deux dernières pistes.

9.3.1 Multiplier les avis

Une première façon de briser l’objectivité esthétique du RNP est d’introduire explicitement une subjectivité à l’aide d’experts en charge de représenter des voies concurrentes d’esthétique. Ainsi, on porterait une plus grande attention, au cours de la phase d’apprentissage, à certaines qualités de l’image qui s’adresseraient en priorité à tel ou tel groupe d’observateurs en fonction de sa sensibilité, sa culture, son éducation.

Cette démarche a en fait été timidement abordée dans S. Kong *et al.* [Kong et al., 2016] qui confie l’évaluation de leur base AADB à un petit nombre de vrais experts dont ils gardent trace dans l’expertise. Chaque utilisateur est alors amené à se comparer à tous les experts et celui qui lui ressemble le plus servira seul à constituer la base d’apprentissage utilisée pour prédire la note de chaque nouvelle photo. Notons que les auteurs n’ont pas cherché à identifier des courants esthétiques précis, ni même à mettre en valeur les écarts de points de vue entre experts. Les différences observées sont le résultat du choix initial (arbitraire ?) d’experts qui manifestent naturellement des écarts non recherchés. On pourrait concevoir une attitude plus systématique de désignation d’experts, chacun reflétant un courant particulier de l’esthétique si l’on disposait d’une cartographie raisonnablement consensuelle de ces courants, ce qui à notre connaissance n’existe pas.

Des travaux cependant s’engagent dans cette voie. Ils concernent le jugement esthétique en photographie et sont particulièrement ambitieux. Ils ont pour objectif de déterminer ces communautés aux goûts plus ou moins homogènes dans des populations extrêmement variées (comme par exemple chez les abonnés à Flickr) [Hong et al., 2016]. Ces travaux, encore en cours, utilisent pour chaque image deux types de descripteurs : des descripteurs sémantiques d’une part (issus de tags fournis avec chaque image), des descripteurs de bas niveau d’autre part (couleur, contraste, contours, textures). Dans l’espace de très grande dimension de ces descripteurs, chaque image est traitée par un algorithme d’allocation latente de Dirichlet (LDA), comme un mot. Les mots sont regroupés en phrases (une phrase égale un photographe) qui fournissent donc des sujets (*topics*) latents qui définissent des communautés (un *topic* est une communauté). Les communautés partagent des sujets voisins dans l’espace des descripteurs. Ce regroupement dans le graphe des sujets se fait par un algorithme de *graph shift*. Si l’outillage mathématique mis en jeu est considérable, les premiers résultats ne sont encore aujourd’hui guère convaincants. Le poids respectifs des divers descripteurs est délicat à trouver

et il est difficile de distinguer ce qui est une tendance esthétique et ce qui est un thème photographique et bien difficile de caractériser de vraies différences esthétiques dans des distinctions encore très grossières (voir la figure 9.8).

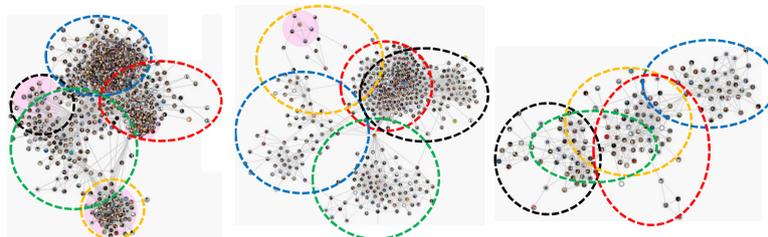


Figure 9.8 – Détection de communautés dans des nuages de descripteurs par une technique d'allocation latente de Dirichlet puis partitionnement de graphe. Les communautés sont identifiées par des ellipses de couleur. En rouge : les “designers”, en bleu : “les couleurs”, en vert : “l’architecture”, en jaune : “le Noir et Blanc”. Les 3 représentations sont des projections particulières du nuage, d’après [Hong et al., 2016]. On voit qu’il y a un important recouvrement entre ces communautés et les thématiques des photos.

Notons que dans la démarche rapportée dans [Kong et al., 2016] on peut reprocher une vision un peu brutale des préférences esthétiques où chaque utilisateur apprécie toute œuvre conformément aux préférences d’un expert et un seul. Ce schéma est probablement meilleur que celui qui aligne tout point de vue sur un point de vue unique mais il ne résiste probablement pas à une analyse très fine.

9.3.2 Les systèmes de recommandation

L’autre solution consiste à mettre en place un système de recommandation pour l’esthétique photographique. Les systèmes de recommandation qui permettent de conseiller à un client d’autres produits susceptibles de lui convenir également, sont opérationnels dans de nombreux domaines (librairie, musique, commerce de détail . . .). Ceux qui fonctionnent bien s’appuient sur des descriptions sémantiques riches des objets (telles qu’on peut les obtenir par exemple dans le résumé d’un livre ou dans un catalogue descriptif des produits offerts par un magasin). Connaissant les choix antérieurs de l’utilisateur, pour peu qu’ils soient assez nombreux et variés, on est alors capable de définir par un vecteur d’attributs, son “portrait” d’acheteur. Ce portrait est alors comparé à des portraits-types obtenus soit à l’aide de modèles théoriques, soit par apprentissage sur un grand nombre d’utilisateurs. Les portraits-types disposent d’“avis” sur tous les produits (souvent binaires : “j’aime” ou “je n’aime pas”). Par ailleurs on dispose de statistiques sur l’achat conjointement de tel produit et de tel autre. Ce sont alors les avis les mieux notés par le portrait-type le plus proche de l’utilisateur des produits les plus fréquemment acquis qui lui sont suggérés, éventuellement filtrés à partir de toute connaissance dont on dispose des écarts entre le portrait-type et l’utilisateur.

Dans le domaine du cinéma [Deldjoo et al., 2016, Elahi et al., 2017, Deldjoo et al., 2018] on trouve des travaux qui semblent assez proches du domaine de la photographie d’art. On pourrait espérer tirer profit de l’analyse qu’ils font du signal d’image (choix des contrastes, des couleurs, des cadrages), pour la transposer à l’image fixe. Malheureusement cela semble encore difficile car le rôle de l’image est très secondaire pour la recommandation en matière de films. Celle-ci repose beaucoup sur une information sémantique riche dont ne dispose pas la photographie (le script, la production, le casting, ainsi que l’enchaînement des images, des changements de plans, la bande sonore . . .).

Des travaux récents ont pris pour objet la recommandation dans le domaine de l’art et concernent surtout la vente en ligne de peintures [Dominguez et al., 2017, Messina et al., 2017, Benouaret, 2017, Messina et al., 2018]. Ils nous semblent pouvoir mieux se transposer au domaine de la photographie que ceux concernant le cinéma. Dans le domaine de l’art, on est particulièrement sensible à des aspects complémentaires concernant d’une part le style de l’œuvre, d’autre part le rôle personnel de l’artiste,

aspects qui interviennent de la même façon en photographie. Certaines particularités fortes du marché de l’art les distinguent cependant, comme par exemple le fait qu’une œuvre d’art n’existe souvent qu’à un exemplaire et qu’il n’est donc pas possible de donner sur elle l’avis de plusieurs acheteurs puisqu’elle disparaît de la vente après le premier achat. Ces travaux sur la recommandation en matière d’art sont cependant encore peu convaincants. Ces travaux, s’appuyant sur les bases de données d’une galerie d’art en ligne, ont confirmé que, comme pour l’évaluation de la qualité, les approches de recommandation à partir des propriétés de bas niveaux de l’œuvre qui utilisent des RNP sont plus efficaces que celles qui s’appuient sur des classifieurs avec primitives dédiées [Dominguez et al., 2017]. Le rôle de ces primitives visuelles demeure cependant très marginal au regard du rôle de certaines primitives symboliques. Ainsi on n’est pas surpris que la méta-donnée “nom de l’artiste” prenne une place beaucoup plus importante dans la recommandation finale [Messina et al., 2018].

Un système de recommandation a également été développé pour assister les recherches au sein de la base de données BAM! du site artistique Behance (voir paragraphe 6.2) [He et al., 2016]. Il utilise deux modèles concurremment, l’un exploite une chaîne de Markov personnalisée pour maintenir au cours des requêtes le profil des préférences de l’utilisateur, l’autre modélise les apparences visuelles des divers éléments de la scène. Ils créent ainsi un système qu’ils appellent VISTA+. Là encore ce sont des algorithmes de recherche de connaissances latentes qui sont mis en œuvre. Elles doivent permettre de définir : la relation d’un observateur à une œuvre, la relation d’un observateur aux autres observateurs, la relation de dépendance entre les œuvres, et enfin l’évolution temporelle des préférences de l’utilisateur au cours de ses visites.

Signalons enfin, le domaine de la photographie, des travaux assez spécialisés qui se sont intéressés à un aspect très particulier de la recommandation, celle qui préside au choix des filtres d’amélioration des photographies sur les réseaux sociaux. Ces filtres qui agissent sur le contraste et la colorisation de l’image sont en petit nombre (une vingtaine que l’on trouve dans les boîtes à outils des réseaux sociaux), la tâche de recommandation est alors raisonnablement complexe, puisqu’elle consiste à proposer à un utilisateur donné, pour une image particulière, un ou deux filtres adaptés à ses goûts et ses images [Sun et al., 2017]. Le logiciel proposé qui fonctionne par un apprentissage par RNP de préférence sur des couples d’images traitées par des logiciels différents, fonctionne relativement bien dans 80 % des cas.

Chapitre 10

En forme de conclusion

Comment aborder aujourd’hui la relation de l’esthétique et de la photographie, alors même que la photographie vit en ce début du *XXI^e* siècle un cycle nouveau de son histoire brève mais mouvementée, cycle lié à l’évolution de la technique de prise de vue et de diffusion ?

Constatons tout d’abord que l’image photographique est de plein droit objet de l’esthétique et que nul ne lui dénie aujourd’hui, comme l’a fait P. Bourdieu il y a 50 ans, une place parmi les autres arts : peinture, sculpture, musique Selon de nombreux critères, la photographie est même devenue l’un des arts les plus dynamiques et les plus respectés [Fried, 2008, Hariman and Lucaites, 2016, Cotton, 2009].

Notons ensuite que, plus que les autres modalités de l’art, la photographie se prête aux techniques modernes d’investigation, d’une part parce qu’elle est aisément archivable et manipulable par l’ordinateur, d’autre part parce qu’elle a trouvé, au sein de la société du numérique, une place qui la rend très facilement accessible en très grandes quantités et une forme directement intégrée aux chaînes de traitement de l’information.

Nous avons présenté une analyse de l’esthétique en photographie selon divers registres, espérant pouvoir réunir ces champs dans une même conclusion. Ce n’est malheureusement pas le cas car il y a bien peu de passerelles entre ces divers pans de connaissance, où chacun se retranche volontiers derrière son expertise (voir par exemple [Gombrich, 2000, Ramachandran and Freeman, 2001, Bullo et Reber, 2013]) et c’est en déclinant chacun des registres où nous l’avons présentée que l’on peut analyser les progrès des relations entre photographie et esthétique.

Dans le domaine philosophique, l’esthétique est confrontée, depuis 50 ans au moins, au bouleversement de la notion même d’art [Arnheim, 1986, Danto, 1992]. Ce bouleversement a amené à l’avant-scène des œuvres où la notion d’esthétique était pour le moins secondaire, si ce n’est même étrangère. Il s’est accompagné d’une grande confusion dans la pensée occidentale [Danto and Goehr, 2014]. Parallèlement, l’irruption des arts “primitifs”, puis l’ouverture internationale du marché de l’art et son exposition à des traditions et des cultures très variées ont alimenté une réflexion sur la contingence de la beauté et sur le poids considérable des cultures dans sa définition [Wascheck, 2000, Brown and Dissanayake, 2009], cette intrusion de la culture ne rendant pour autant la notion de beauté plus malléable. Cet état de chose s’étend à l’image tout autant qu’à la peinture ou la sculpture.

Dans le domaine de la neuro-biologie, le champ de la neuro-esthétique est toujours en quête de réponses convaincantes à la chronologie des mécanismes de l’appréciation esthétique [Zeki, 1999]. Si un petit nombre d’hypothèses ont été écartées (comme celles du traitement par les aires dédiées d’un hypothétique cortex hédonique), de grandes incertitudes demeurent sur la sélection des flux montants du cortex visuel, ainsi que sur les relations avec les aires en charge du souvenir, du raisonnement ou de l’émotion, ce qui laisse l’interprétation très largement spéculative [Changeux, 2016]. La nécessité de disposer de meilleurs outils d’introspection fonctionnelle (comme par exemple l’IRM de diffusion,

qui met en évidence les connexions entre aires actives) se fait cruellement sentir. De même des outils rapides permettant de relier les signaux issus des voies visuelles aux activations des zones préfrontales seraient indispensables pour étayer des hypothèses de mécanismes d'évaluation esthétique convaincants.

Les expériences de psycho-physiologie expérimentale, très bien adaptées à de nombreuses études sur l'appareil visuel, sur la perception et sur l'interprétation de formes simples sont aujourd'hui beaucoup trop simplifiées dans leurs protocoles expérimentaux pour pouvoir débroussailler utilement le foisonnement des interrogations du praticien ou du théoricien de la photographie confronté à la notion du beau. Les édifices théoriques élaborés par la Gestalt Theorie [Wertheimer, 1938] peinent à se retrouver dans des recommandations faites en termes sémantiquement trop généraux ("zone d'intérêt", "sujet principal", "personnage central" . . .), ou dans des concepts par trop géométriques ("règle du tiers", "centre de fixation" . . .). La vérification à posteriori de ces recommandations (lorsqu'elles ont pu être formulées) par le biais de statistiques conduites après traitement d'images serait certainement un acquis expérimental très appréciable. Ces travaux manquent cependant aujourd'hui de protocoles universellement reconnus, de corpus de données partagés et acceptés et souffrent donc de capacités modestes de conviction.

Enfin, les dernières venues des études sur l'esthétique des photographies, celles qui s'appuient sur les méthodes d'apprentissage automatique et les bases de données du web, s'installent de façon iconoclaste et exagérément aculturée dans le paysage, fortes de potentiels développements sur le marché des nouvelles technologies et de la réputation soigneusement entretenue d'ultime progrès d'une "intelligence artificielle", qui a du mal à échapper à son image de propagande socio-économique. Ces approches sont encore bien loin de convaincre malgré les énormes ressources qu'elles ont su drainer. Les travaux encore préliminaires sont, au mieux, "prometteurs" (ce qui dans le domaine scientifique indique qu'ils doivent faire leurs preuves), au pire, bien décevants. Il est plus juste de dire qu'ils offrent des résultats intéressants mais intrigants dont on ne sait pas bien quoi faire en l'état.

S'il nous faut identifier des raisons de ces blocages dans nos démarches, un élément semble devoir être pointé du doigt que la plupart des auteurs se gardent bien d'aborder, il s'agit de la relation très intime de l'esthétique et de la culture. Niée dans les approches objectivistes, cette relation est immanente dans les approches subjectivistes, mais elle introduit une nouvelle variable, en l'occurrence la culture elle-même dont la nature est au moins aussi complexe et protéiforme que l'esthétique. Si la notion de culture semble détenir de nombreuses clefs pour expliquer l'esthétique, elle apparaît elle-même terriblement hermétique à nos outils actuels d'introspection et semble en l'occurrence remplacer un mystère très grand par un mystère plus grand encore.

Annexe I : Petite revue des esthétiques

Voyons comment diverses conceptions du Beau se sont rangées dès l'Antiquité sous l'étiquette objectiviste. On trouvera dans [Tatarkiewicz, 1970] de très larges développements d'où sont issues ces lignes.

L'esthétique dans le monde antique : l'objectivisme fait loi

L'esthétique de la période classique grecque : la Grèce hellénique

La période classique athénienne (V^e et IV^e siècle av JC) porte toutes les références ultimes à l'esthétique. Le Beau y est doté de vertus autant esthétiques que morales. Il est le produit de trois sources : *“la matière, fournie par la nature, le savoir, fruit de la tradition et le travail donné par l'artisan”*¹ ([Tatarkiewicz, 1970], livre 1). La notion de création y est quasiment absente, la fidélité à la tradition y est dominante, tradition exprimée par les “canons” qui fixent des formes précises à l'objet (figure 10.1). Les canons ont le plus souvent de lointaines sources religieuses et liturgiques, mais d'autres semblent bien n'avoir que des fondements esthétiques.

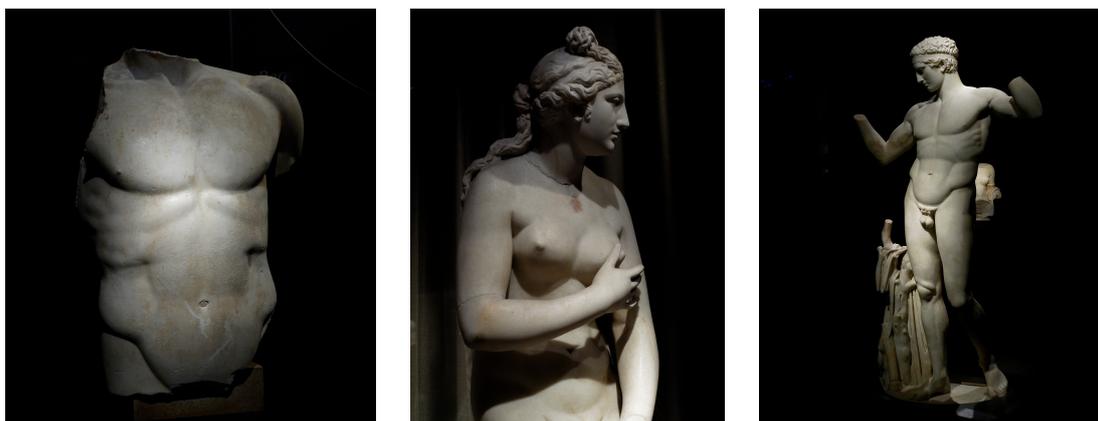


Figure 10.1 – Nous ne disposons plus de peintures datant de l'époque dite “hellénique” et il faut nous figurer les canons qui présidaient à l'œuvre peinte à la lumière des statues. Les artistes grecs ont ouvert l'art à la représentation naturelle du corps humain, abandonnant les schémas rigides qui prévalaient dans les mondes antérieurs (art égyptien, cycladique, grec archaïque ...) : à gauche Torse d'homme, copie romaine en marbre d'un bronze du 5^e siècle av. JC, au centre : La naissance d'Aphrodite, à droite, Ephèbe peignant ses cheveux (diadoumenos). Tous provenant du musée archéologique d'Athènes.

1. Pour les Grecs, les arts plastiques : peinture, architecture, sculpture ... sont œuvres d'artisans et non d'artistes, alors que l'expression orale, dotée du pouvoir quasi-divinatoire de préparer l'avenir, est œuvre d'artiste.

Les canons nous sont assez bien connus dans les domaines de l'architecture et de la statuaire où de nombreuses œuvres sont encore disponibles. Ils ont été très largement repris et commentés par Vitruve², à Rome, au premier siècle. Nous avons beaucoup moins de traces en peinture. Ils s'expriment par des règles, de simplicité³, d'ordre⁴, d'égalité⁵ et de symétrie, en particulier et par des chiffres, rapports entre les grandeurs qui expriment des proportions et reflètent une harmonie⁶. La simplicité vise à garantir la "juste mesure", c'est-à-dire l'utilisation des seuls éléments nécessaires à l'expression. Plus les rapports sont simples, plus grand est l'effet esthétique.

Aristote introduit un concept supplémentaire qui résume les propriétés ci-dessus, celle de perceptibilité qui prône que les qualités du bel objet doivent apparaître naturellement et spontanément, en adaptant la taille de l'objet au champ de vue de l'observateur et en éliminant les artefacts superflus. D'ailleurs Aristote soutient que la perception du beau s'impose à l'esprit sans justification, que ce n'est pas une propriété des sens ou du jugement⁷.

L'esthétique utilise les mathématiques, mais elle n'est pas d'essence abstraite. Elle puise ses sources dans la nature et dans l'univers et en particulier dans le corps humain qui, partie de l'univers, en reflète l'harmonie. Aujourd'hui, nous dirions que cette harmonie est fractale car elle concerne les proportions à toutes les échelles : celle du buste au corps entier, celle de la main au membre, celle des phalanges à la main, celle du nez au visage et ainsi de suite⁸. Les rapports ainsi préférés sont $1/2$ ou $1/3$, ou ceux construits à partir du triangle rectangle "parfait" de côtés (3,4,5) dont on déduit le nombre d'or $\frac{1+\sqrt{5}}{2} = 1,618$ ou le rapport approché $8/5 = 1,6$. Mais ces proportions sont dans des rapports finalement assez nombreux, car, par exemple en architecture, elles doivent se conformer aux exigences de la mécanique pour assurer la répartition de charges. Dans le domaine des couleurs, l'harmonie se trouve dans la composition de quatre primaires : le blanc, le noir, le rouge et le jaune (Démocrite, Empédocle)⁹.

Si les canons sont des contraintes supérieures, de petits écarts aux canons sont tolérés et même appréciés de certains¹⁰. Ils ont pour but de mieux faire percevoir les règles esthétiques par l'observateur. Ainsi il avait été noté que le parallélisme des colonnes d'un temple élevé donnent l'impression de diverger. Il est alors conseillé de rendre ces colonnes légèrement convergentes de façon qu'un observateur, au centre de la place, les voit parallèles. Vitruve rapporte que Démocrite usait de tels artifices

2. Marcus Vitruvius Pollio, dit Vitruve, est un architecte romain du 1^{er} siècle. Il a exprimé les règles de l'esthétique par six lois :

1. Ordre, (*ταξίς, taxis*),
2. Disposition, (*διαθεσις, diathesis*),
3. Art du mouvement ou eurythmie, (*ευρυθμος, eurhythmos*),
4. Symétrie, (*συμμεθρία, symmetria*),
5. Propriété,
6. Economie, (*οικονομία, Oikonomia*).

3. "Si quelqu'un dépasse la mesure, les choses les plus agréables deviennent déplaisantes.", Démocrite, rapporté par J. Stobaeus, Flor, III, 17, 38, frg. B 23 Diels, cité par [Tatarkiewicz, 1970].

4. "L'ordre et les proportions sont beaux et utiles, tandis que le désordre et l'absence de proportions sont laids et inutiles.", Ecole de Pythagore, rapporté par J. Stobaeus IV I, 40 H. frg. D 4 Diels, cité par [Tatarkiewicz, 1970].

5. "Dans toute chose, l'égalité est juste, l'excès et le défaut ne le sont pas à mon avis.", Démocrite, Democrites, Sent. 68, frg. B 102 Diels, cité par [Tatarkiewicz, 1970].

6. Cette vision pythagoricienne de la beauté est reprise dans les textes chrétiens écrits au début de notre ère : "Omnia in mensura et numero et pondere disposuisti. (Tu as fait toute chose en accordant mesure, nombre et poids.)", Livre de la Sagesse, XI, 21.

7. Diogène soutient cette conception a priori du Beau universel. Il affirme que seul un aveugle peut s'interroger sur la raison de perdre du temps sur la notion de beau.

8. Pour son rapport au corps humain, cette esthétique a été qualifiée d'"anthropométrique" par E. Panofsky [Panofsky, 1983].

9. "Il y a quatre couleurs correspondant au nombre des éléments : blanc, noir, rouge et jaune." d'après Empédocle, rapporté par Aëtius, Plac. I, 15, 3; frg A 92 Diels, cité par [Tatarkiewicz, 1970].

10. Mais ce n'est pas le cas de Platon qui recommande le plus grand accord avec la tradition : "... en bref, ceux qui sont en charge de notre communauté doivent prendre garde à ne pas négliger la moindre consigne envers l'innovation à travers notre système d'éducation du corps ou de l'esprit ...", Platon, la République, 424 B, cité par [Tatarkiewicz, 1970].

pour exprimer le relief dans ses peintures.

La Grèce hellénistique et Rome

La Grèce hellénique a fait place, après les conquêtes d'Alexandre, à la Grèce hellénistique et au monde romain. Les arts et la notion de beau s'inscrivent les uns et les autres dans la tradition classique. Quoique le contexte social soit complètement bouleversé par l'émergence d'empires immenses, l'ouverture des frontières à de lointaines cultures très élaborées et l'apparition de nouveaux régimes politiques, les grandes lignes de l'esthétique de Platon et d'Aristote demeurent au centre de la *doxa* esthétique.

Néanmoins trois grands courants philosophiques s'affirment qui traitent différemment de l'esthétique :

- les hédonistes (Epicure) : ils assimilent art et plaisir et symétriquement plaisir et art. Il n'y a pas d'autonomie de l'art et pas de règles esthétiques ;
- les sceptiques (Pyrrhon) soulignent les désaccords et les contradictions des jugements sur la beauté et l'art ; en conséquence, théoriser sur l'art est impossible, inutile et dangereux ; ils réfutent les arguments platoniciens d'universalité de l'art ;
- les stoïciens (Zénon, Sénèque, Cicéron) subordonnent le beau aux valeurs morales ; ils reconnaissent une beauté morale et une beauté des sens, la seconde ayant généralement moins de valeur. Ils s'accordent aux critères platoniciens et leur font écho.

Dans la pratique, se sont les règles platoniciennes et péripatéticiennes qui sont le plus généralement suivies. L'influence des stoïciens se traduit par une plus grande souplesse dans la mise en œuvre des canons, qui régissent les grandes lignes des œuvres, tandis que les détails s'accommodent d'un *decorum*¹¹ qui peut être interprété comme une déclinaison locale et partielle du canon spécifique à chaque objet représenté.

Le Haut Moyen-Âge : Europe de l'Ouest et Byzance

Reprise par le monde chrétien naissant, l'esthétique platonicienne fait migrer la beauté de la nature au monde transcendant (Plotin, Pseudo-Denys l'Aréopagite). C'est le monde divin qui est paré de tous les attributs de la beauté et c'est dans sa représentation que l'artiste met en œuvre les recommandations d'harmonie, d'équilibre, de simplicité . . . C'est la période du néo-platonnisme.

De cette dimension divine, la beauté gagne une vertu nouvelle. Alors qu'elle s'imposait passivement à l'observateur, maintenant elle rayonne et participe à la diffusion de la parole. La notion de lumière y prend ainsi une place particulière avec ce qui l'accompagne de brillant. En contrepartie, elle perd en sensualité.

Plus à l'Est, dans cette autre partie de l'empire romain qui commence à se séparer de l'Occident, une figure nouvelle de la peinture émerge : l'icône prend à Byzance une place centrale. Seule l'image de Dieu ou de ses saints est digne d'effort. Elle reproduit un monde formellement beau et dépourvu des troubles terrestres qu'exprimerait une expression faciale. Les codes (transpositions des canons) s'imposent : œil en position centrale, figures allongées, couleurs artificielles, cadre dépourvu de contexte matériel ou réduit à des fragments symboliques de végétation, si ce n'est idéalisé par un disque d'or.

Enfin, en Afrique du Nord, Augustin, évêque d'Hippone, partant des textes platoniciens revus par Plotin, jette les bases de ce qui fera l'esthétique de l'Occident jusqu'à la Renaissance : la beauté est objective et indépendante de nous ; elle est créée par l'harmonie des parties, leur ordre et leur unité et cela doit être réglé par des nombres¹². Elle culmine dans la beauté divine, qui surpasse et inspire

11. *“Ut enim in vita, sic in oratione nihil est difficilius quam quid deceat videre. Πρέπον appellant hoc Graeci; nos dicamus sane decorum. De quo praeclare et multa praecipiuntur et res est cognitione dignissima.”* (*“Dans l'oratoire, comme dans la vie, il n'est rien de plus difficile que de reconnaître ce qui convient. Les Grecs l'appellent “prepon” tandis que nous disons “decorum”. Il y a de nombreuses et excellentes règles qui méritent d'être connues.”*), Cicéron, Orator, 21, 70, cité par [Tatarkiewicz, 1970].

12. *“Omnia enim quanto magis moderata, speciosa, ordinata sunt, tanto magis utique bona sunt; quanto autem*

toute beauté. C'est aussi probablement lui qui introduit l'idée qu'il n'y a pas de valence négative du beau (le contraire du beau), mais simplement une absence de beau dans ce que nous désignons par laid¹³, un point sur lequel nous aurons l'occasion de revenir.

Le Bas Moyen-Âge

A partir de l'an 1000, c'est avec l'architecture religieuse que se manifeste la résurgence d'une vigoureuse expression esthétique : à travers le style roman tout d'abord, puis du style gothique. Art d'artisans bien sûr, mais aussi d'ingénieurs ; et donc art du Nombre dans la tradition pythagoricienne, mais aussi dans l'enseignement et le secret des ateliers et des guildes. Les règles de proportion, d'harmonie, de rythme¹⁴ sont retenues et transmises, mais les sources sont généralement omises ; les chiffres s'imposent par eux-mêmes comme un prérequis de beauté et non plus, comme à l'origine platonicienne, par leur rapport au corps humain, à la sagesse de la pensée sociale et collective. Les prouesses de la technique aidant, la lumière est amenée dans les édifices ; les arts ont leur inspiration au firmament. C'est R. Grosseteste qui fait le mieux la synthèse entre l'esthétique des nombres, de la géométrie et de la lumière¹⁵, dans une approche qui vise à donner une vision mathématique unifiée de la beauté [Tatarkiewicz, 1970].

Une distinction nouvelle dans l'esthétique est apportée par Alhazem¹⁶, philosophe et mathématicien du monde arabo-musulman au virage de l'an 1000. Alhazem, fidèle propagateur des écrits helléniques et hellénistiques est aussi un excellent opticien. Son message diffuse lentement en Europe et il faut plusieurs siècles pour qu'il soit repris par la nouvelle école scholastique (E. Vitellion, R. Bacon). Il défend l'idée que l'œil ne perçoit directement que la couleur et la lumière. Nous n'apprécions les autres composantes de l'esthétique (forme, composition, disposition) que par d'autres actions qui composent la perception dont la mémoire et la réflexion. C'est cette forme nouvelle - et singulièrement moderne - qui est à la base de la perception esthétique qui peut distinguer une beauté simple, immédiatement perçue par l'œil seul, puis une plus complexe s'adressant à des formes elles-mêmes plus complexes¹⁷. Si la distinction de deux types de beau est classique dans la pensée antique, elle y est fondée sur des principes philosophiques transcendants, tandis que la distinction proposée ici repose sur des mécanismes psycho-physiologiques inconnus mais pressentis. Il n'y a que deux propriétés qui participent au premier type de beauté (la lumière et la couleur), mais Alhazem et Vitellion recensent près de 20 propriétés contribuant à la seconde ([Tatarkiewicz, 1970], II, p. 266). Regrettons cependant que ces propriétés soient bien loin de fournir des règles opérationnelles de beauté. Notons enfin que ces philosophes demeurent fondamentalement objectivistes dans leur conception de la beauté. S'ils prennent en compte l'observateur, c'est pour comprendre les mécanismes qui le rendent sensible au beau.

G. d'Ockham traite peu de l'esthétique mais il se penche sur la nature de l'image et y décèle une

minus moderata, minus speciosa, minus ordinata sunt, minus bona sunt. (En toute chose, plus il y a de mesure, de forme, de nombre, meilleure elle est ; et moins il y a de mesure, de forme, de nombre, moins elle est bonne.)", Augustin, De l'origine du Bien, 3, PL, 42, c. 554, cité par [Tatarkiewicz, 1970].

13. Il semble qu'il lui était difficile d'affirmer que dans un monde créé par Dieu il y ait eu création du laid. Ainsi, il écrit : "*Cuique naturae non est malum nisi minui bono. (Aucune nature n'a de mal en elle, mais seulement un déficit de bien.)*", Augustin, De l'origine du Bien, 14-17, PL, 42, c. 555-6, cité par [Tatarkiewicz, 1970].

14. "*Paritas autem dimensionis secundum aequalitatem, similitudinem, compositionem et modificatam et commensuratam congruentiam artium non minima pars pulchritudinis est. (L'unité de dimension établie sur les principes d'égalité, de similarité, d'arrangement approprié, d'adaptation et de commensuration des parties n'est pas le moindre facteur de beauté.)*", Baudouin de Cantorbéry, Tractatus de salutatione angelica, PL 204, c. 469, cité par [Tatarkiewicz, 1970].

15. "*Solae quinque proportionales repertae his quattuor numeris : unum, duo, tris, quattuor, aptantur compositioni et concordiae stabilienti omne compositum. (La composition et l'harmonie de toute chose composée découle seulement des cinq proportions que l'on trouve entre les quatre nombres : un, deux, trois, quatre.)*", Grosseteste, de la Lumière, (Baur, 58) et "*lux est maxime pulchrificativa et pulchritudinis manifestiva. (La lumière rend les choses belles et montre leur beauté au plus haut point.)*", Grosseteste, Comm. in Div. Nom., IV (Pouillon 320), cités par [Tatarkiewicz, 1970].

16. Alhazem a pour nom complet Abu Ali al-Hasan ibn al-Hasan ibn al-Haytham.

17. "*Omnis vera comprehensio formarum visibilium aut est per solam intuitionem, aut per intuitionem cum scientia praecedente. (Toute appréhension des formes visibles, se fonde soit sur une observation directe, soit sur une observation à travers une connaissance antérieure.)*", Vitellion, Optique, III, 62 (Alhazem, II, 69), cité par [Tatarkiewicz, 1970].

dimension originale qui échappe à l'objet représenté et qui est le fruit de l'intention du dessinateur¹⁸. C'est là aussi un concept nouveau qui sera particulièrement pertinent dans l'esthétique de la photographie où l'objet qui est fixe et le même pour tous pourra être "transcendé" en objet d'art par l'œil du photographe. La subjectivité s'introduit alors, non pas dans l'observateur où nous l'attendions naturellement mais dans le créateur. Ces idées seront très largement adoptées par Hegel.

La Renaissance

C'est ce mouvement qui a porté la revisite de la pensée hellénique, puis sa reformulation en termes modernes et ensuite sa diffusion en l'Europe. Elle en a hérité pour partie son nom¹⁹. Elle s'est développée aux XIV^e et XV^e siècles dans le Nord de la péninsule italienne, mais surtout à Florence, Venise puis Rome. L'héritage platonicien est le point de départ de la pensée esthétique, appuyé par les textes de Plotin et de Vitruve dans le domaine architectural et complété de notions issues d'Aristote.

Marsile Ficin, à Florence, se fait le héraut de la Beauté platonicienne revisitée par l'esprit renaissant. L'harmonie, la symétrie, l'équilibre du cosmos, créés par Dieu se reflètent dans l'œuvre de l'artiste et donnent naissance au Beau comme à l'Amour : "*La beauté du corps ne consiste pas dans l'ombre de la matière, mais dans la clarté et la grâce de la forme (...) dans le nombre et la mesure convenable*"²⁰, et Leon Baptista Alberti renchérit : "*La beauté est un certain accord et si l'on peut dire une conspiration des parties dans le tout où elles s'établissent selon un nombre, un ordre qualitatif et une place définis comme le requiert l'harmonie principe absolu et premier de la nature.*"²¹. Ils inspireront Vinci, Raphaël, Michel-Ange et bien d'autres

L'esthétique se décline sous la Renaissance vers de nombreux autres domaines de la vie courante (orfèvrerie, marquetterie, tapisserie, ...). Le rôle des sciences est particulièrement renforcé pour soutenir la notion de forme dans les arts visuels. La nature reste le modèle premier de la beauté et l'artiste doit s'efforcer de s'en approcher. Sa contribution se place souvent dans la subtilité et la pertinence des écarts qu'il apporte dans cette reproduction. La notion d'harmonie globale préside à la composition des œuvres picturales²². La beauté est universelle et tout le monde doit en convenir²³.

C'est aussi au cours de la Renaissance que la place de l'artiste est vraiment reconnue à l'égal de celle d'un créateur²⁴. C'est aussi à cette période que l'on reconnaît que l'objectif de l'artiste est bien

18. "(a) *Strictissime et sic imago est substantia formata ab artifice ad similitudinem alterius ... Sic accipiendo imaginem de ratione imaginis est, quod fiat ad imitationem illius, cuius est imago; [...]*.

(b) *Alio modo accipitur imago pro tali formato, sive fiat ad imitationem alterius, sive non.*

((a) *Au sens strict, l'image est une substance faite par l'artiste à la ressemblance d'une autre ... Dans ce sens, il est important que l'image soit à l'imitation de ce qu'elle représente; [...]. (b) Dans un autre sens, l'image est conçue comme toute chose faite par l'artiste, comme l'imitation d'une autre ou non.*"), Ockham, Questions en IV, Sententiarum libros, I, d., 3, q. 10B, (Baudry), cité par [Tatarkiewicz, 1970].

19. "Renaissance" a deux sens, le premier exprime la mise au monde de l'homme nouveau, libéré des structures mentales du Moyen-Âge, le second la remise à l'honneur de la philosophie hellénique au sein de la société (voir [Tatarkiewicz, 1970], III, p. 32).

20. Marsile Ficin, *Lettre à A. Canisano*, citée par A. Chastel [Chastel, 1996].

21. J.B. Alberti *De re aedificatoria*, cité dans [Chastel, 1996].

22. "*La bellezza non consiste solo nelli colori, ma è una qualità que resulta dalla proporzione e corrispondenzia delli membri e delle altre parti del corpo; tu non dirai che una donna sia bella per avere uno bello naso o belle mani, ma quando vi sono tutte le proporzioni. (La beauté ne consiste pas en couleurs, mais en une qualité qui résulte de la proportion et des correspondances des parties du corps. Tu ne diras pas qu'une femme est belle parce qu'elle a un beau nez ou de belles mains, mais seulement si tout est en proportion.)*", J. Savonarole, Sermon sur Ézéchiël, XXVIII, cité par [Tatarkiewicz, 1970].

23. "*Ma e'ci sono alcuni che non approvano simili cose et che dicono che ella è una certa varia oppenione, con la quale noi facciamo giudicio della bellezza, et di tutte le muraglie, e che la forma degli edificij se muta secundo il diletto, e il piacere di ciascuno, non si ristriugnendo dentro ad alcuni comandamenti de la arte. Comune diffeto de gli Ignoranti, é il dire che quelle cose che non sanno loro, non sieno. (Mais il y a ceux qui ne reconnaissent pas cela et qui disent qu'il y a diverses vues concernant la beauté et l'architecture et que la forme des édifices évolue selon les goûts et les préférences de chacun, sans être gouvernées par les commandements de l'art. C'est une erreur commune des ignorants qui disent que ce qu'ils ignorent n'existe pas.)*", L. B. Alberti, *Della pittura*, Lib. II, cité par [Tatarkiewicz, 1970].

24. "*La deità, ch'a la scientia del pittorefa che la mente del pittore si transmuta in una similitudine di mente divina. (La divinité qui est la science de la peinture, transmute l'esprit du peintre en une ressemblance avec l'esprit*

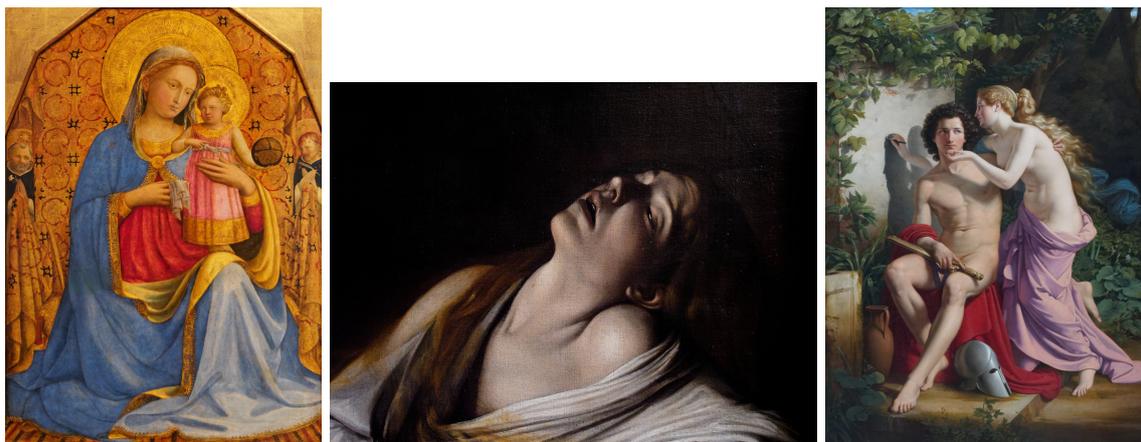


Figure 10.2 – La Renaissance a revitalisé les canons de la beauté antique pour figurer ses personnages religieux en majesté, en les ornant des symboles hérités du Moyen-Âge et de l'époque byzantine (ici les auréoles) : à gauche la “Madone à l'Enfant avec les saints martyrs Dominique et Pierre”, de Fra Angelico, de 1435, au musée Bode de Berlin. Plus tard, les artistes sont revenus à l'expression de la nature seule, dépourvue de tout symbolisme (au centre, détail de “la Madeleine Klain en extase”, de Caravage, en 1606, collection particulière). Cette esthétique s'est figée aux 18^e et 19^e siècles dans un classicisme scolaire (à droite “La naissance de la peinture”, Eduard Daege, 1832, Alte Nationalgalerie, Berlin), qui débouchera sur l'art pompier.

le plaisir qu'il tire d'une part de sa création, d'autre part de son partage avec les spectateurs.

Le monde Moderne : de l'Objectivisme au Subjectivisme

Le XVII^e siècle s'ouvre sur un monde considérablement bouleversé. Les philosophes anglais et écossais férus de méthode scientifique et expérimentale ont proposé des modèles hardis de description des fonctions perceptives et des fonctions cognitives. Les physiciens, Copernic, Kepler puis Gallilée viennent bouleverser le modèle ptoléméen de l'univers et portent au passage le doute sur le legs platonicien. Mais c'est Descartes qui se charge d'établir un modèle résolument neuf qui ouvre la porte aux Lumières.

René Descartes

Descartes ne rompt pas totalement avec la philosophie platonicienne. Il en conserve très largement la trame : un univers bâti autour des concepts mathématiques, un univers rationnel et accessible à l'entendement, reflétant des vérités universelles et temporellement stables ; un univers ou physique et métaphysique sont étroitement liés. Mais, à l'inverse des Grecs anciens, il considère que ce monde est entièrement accessible à l'entendement rationnel. S'il laisse une place au “Créateur”, c'est celle de “l'impulsion initiale” qui a donné naissance au monde. Mais le monde ensuite poursuit sa course selon le mouvement initial et le Créateur, dans sa toute puissance, est devenu indifférent à ce destin et ne se pique pas d'en modifier le cours, laissant à l'homme éclairé, le loisir d'en percer les mystères²⁵. Descartes n'a pas beaucoup écrit sur l'art et l'esthétique qu'il tient pour occupations de salon, mais pour lui, si le Beau et le Bien sont universels et intemporels, comme pour Platon, il n'en fait pas des propriétés de l'univers qui seraient innées dans l'homme, mais bien des expériences individuelles propres au sujet pensant. Ainsi le Beau, ressenti par tous, n'est pas une propriété de l'âme universelle qui baigne la création, mais une expérience vécue de l'être individuel et le jugement esthétique est

divin.), Leonardo Da Vinci, Traité de la Peinture, frag. 280, cité par [Tatarkiewicz, 1970].

25. “Certes Dieu a tout créé, mais il s'est absenté de sa création. Une entière indifférence en Dieu est une preuve très grande de sa toute puissance.” Descartes, Principes de Philosophie, II, cité par [Dézarnaud-Dandine and Sevin, 2007].

le résultat d'une déduction rationnelle. Pour Descartes, le monde perçu est régi par les lois de la physique. Il nous est accessible par la vision et la perception qui créent un monde sensible. Ce monde sensible est distinct du monde rationnel que traite notre esprit.

C'est cette distinction "monde sensible *vs.* monde rationnel" qui servira pendant deux siècles de champ de disputation entre philosophes.

Jean-Pierre de Crousaz

Parmi les réflexions objectivistes les plus abouties, est celle de Jean-Pierre de Crousaz, philosophe suisse, auteur d'un *Traité du Beau* [DeCrousaz, 1724] qui reprend le partage cartésien et en distribue les propriétés dans le champ de l'esthétique : le Beau a deux natures, l'une qui relève des idées (et donc du monde platonicien), l'autre des sentiments²⁶. Cette double hérédité accordée au Beau fait dire à Syliane Malinowski-Charles : "*le pas franchi par Crousaz est significatif et il peut être vu comme l'emblème d'une transition majeure entre le paradigme rationaliste (qu'on pourrait dire aussi objectiviste) et le paradigme empiriste ou subjectiviste*" [Malinowski-Charles, 2004]. De Crousaz, explorant plus à fond la notion d'esthétique, s'accorde sur des critères très proches de ceux que mettra en équation Birkhoff, qui mêlent intimement complexité et simplicité²⁷.

Les idées "translationnelles" de Crousaz seront reprises en France par Jean-Baptiste Du Bos et le père André²⁸ et en Irlande par Francis Hutcheson, [Jullien, 2017].

William Hogarth

Un peu postérieur à de Crousaz, William Hogarth, peintre très reconnu en son temps, a rédigé en 1753 une "*Analysis of Beauty, written with the view of fixing the fluctuating ideas of taste*" où il s'inscrit nettement dans une esthétique objectiviste. Il propose ainsi la "courbe de beauté" (figure 10.3) qui reflète à ses yeux l'excellence esthétique. Par ailleurs il fait l'inventaire des critères de beauté, bien au delà de ceux retenus dans l'esthétique platonicienne (simplicité, régularité, symétrie, harmonie).

Hogarth identifie dix critères de beauté pour la peinture²⁹ ; il en ajoute deux spécifiquement pour la beauté du corps humain et deux pour le mouvement qu'il destine à qualifier la beauté de la danse. Notons que le premier critère, dénommé "*fitness*" exprime l'adaptation fonctionnelle de l'objet à son usage, un critère bien peu platonicien ; un autre, dénommé "*intricacy*" s'attache à la faculté de l'objet de défier notre compréhension immédiate, une sorte de prime à la réflexion qui aura des échos chez les philosophes du siècle futur.

David Hume

David Hume s'exprime sur l'esthétique dans plusieurs essais entre 1740 et 1760. Il adopte une position nettement subjectiviste et expérimentaliste. Le Beau est un ressenti de l'esprit, comme le Juste, issu de notre perception, et généralement résultat d'un rapprochement positif entre l'objet et sa fonction, voire entre l'animal et sa nature : les membres d'un cheval sont beaux si leur forme est

26. De Crousaz écrit ainsi : "*Quand on demande ce que c'est que le Beau, on ne prétend pas parler d'un objet qui existe hors de nous et séparé de tout autre, comme quand on demande ce que c'est qu'un cheval ou ce que c'est un arbre.*", dans [DeCrousaz, 1724].

27. "*le beau est créé par l'unité au sein de la diversité*", dans [DeCrousaz, 1724], prenant ses inspirations dans la beauté des figures mathématiques : triangle, carré, polygone

28. "*Je dis qu'il y a un beau essentiel, & indépendant de toute institution, même divine : qu'il y a un beau naturel, & indépendant de l'opinion des hommes : enfin qu'il y a une espèce de beau, d'institution humaine, & qui est arbitraire jusqu'à un certain point.*", dans [André, 1759], pp 4.

29. Les critères de Hogarth sont les suivants : *intricacy, variety, uniformity, simplicity, quantity* (ceci correspond plus ou moins à l'harmonie du tout et des parties), *fitness* (qui correspond à l'accord de la forme et de son usage), ainsi que quatre qualités des lignes qui permettent de définir la ligne serpentine comme idéale.



Figure 10.3 – La ligne serpentine figure en couverture de l’ouvrage original de W. Howarth “L’analyse de la Beauté”. Elle représente pour l’auteur le summum de l’esthétique graphique et mérite cet attribut aussi bien en tant qu’image fixe que comme ligne en mouvement, où elle permettra de reconnaître par exemple les plus belles phases de la danse.

utile aux fonctions de transport et de vitesse. Transmise à notre esprit par perception, la forme suscite spontanément notre sympathie, c’est-à-dire une résonance positive et désintéressée, qui prend place en amont de la réflexion. Les formes qui plaisent relèvent de règles que Hume n’explicité pas, mais il leur reconnaît une certaine universalité et une certaine destination qui rappellent les penseurs grecs : “Certaines formes ou qualités particulières, de par la structure originale de la constitution interne de l’homme, sont calculées pour plaire et d’autres pour déplaire³⁰”. Le réceptacle des impressions (le cerveau hédonique ?) qu’est la nature humaine est fondamentalement le même en tous les hommes et à toutes les époques.

Cependant, pour D. Hume, le sentiment de goût n’est pas égal entre les hommes, mais une petite élite partage le sens du beau, constituant ce qui a été baptisé “aristocratie esthétique” ; elle fixe la norme du goût et se distingue par des qualités particulières³¹.

Alexander Baumgarten

Alexander Gotlieb Baumgarten est le premier d’une longue lignée de philosophes allemands qui mettront l’art au centre de leur philosophie. Il est l’auteur, en 1750, de l’*Esthétique*, ouvrage qui crée véritablement le terme qui couvrira la science du goût, mêlant dans une approche identique tous les arts et séparant nettement le Bien et le Beau. Il fait apparaître clairement que sont concernées sous ce chapeau aussi bien la perception (et selon les diverses modalités des diverses sensations), les représentations et les associations mentales aussi variées que les illusions, les rêves, les hallucinations ou les raisonnements.

Baumgarten attribue une grande importance à l’intuition sensible dont il précisera la définition en se dégageant de ses prédécesseurs, Descartes et Leibnitz : c’est une perception directe des propriétés de l’œuvre qui s’installe en parallèle du raisonnement mais permet souvent d’accéder plus rapidement

30. David Hume, *Sur la Norme du goût* in *Essais esthétiques* II Vrin 1974.

31. Pour D. Hume, le jugement de beau ne sert de référence que s’il est issu d’experts dotés d’un certain nombre de vertues :

1. une nature saine,
2. des organes délicats et affinés,
3. une culture travaillée par la fréquentation des œuvres d’art,
4. un avis naturel, non affecté ou maniéré, gage d’un bon sens et d’un entendement commun,
5. une intelligence suffisamment aiguisée qui permet de créer une esthétique de la perfection.

et plus efficacement, quoique de manière confuse, à l'essence des choses [Malinowski-Charles, 2005]. Baumgarten l'affirme : *“L'intuition est nécessaire à la beauté”*³².

Emmanuel Kant

E. Kant est reconnu comme le grand théoricien moderne de l'esthétique³³. Avec son traité dédié spécifiquement à l'esthétique, il a laissé un legs inestimable qui demeure encore aujourd'hui très abondamment commenté.

Comme nous l'avons indiqué déjà, s'il considère bien, comme Platon, que le Beau est universel, il met cette vertu non tant au compte de l'objet lui-même, qu'au compte du jugement du sens esthétique de l'observateur. Il considère ce jugement comme un don universel, don commun à tous les hommes et également partagé qu'il appelle le “sens commun”. Pour lui, le Beau est étroitement attaché à la notion du Bien. C'est un concept qui est à la fois perçu par nos sens et entendu par notre conscience, un jeu libre de notre imagination et de notre compréhension. Il n'a pas de finalité biologique mais vise à l'harmonie de nos représentations internes qu'il ne faut pas confondre avec le simple plaisir car ce ressenti est désintéressé de toute finalité. En cela c'est bien une propriété de l'observateur et non de l'objet. Néanmoins, le subjectivisme de l'esthétique kantienne s'accorde en de nombreux points à l'objectivisme platonicien : universalité (qu'il appelle “inter-subjectivité”) et atemporalité. Il y ajoute des attributs qu'il défend intensément, en particulier “l'autonomie de jugement”.

Kant, s'il est prêt à reconnaître l'individualisme du sens du goût sur la beauté qu'il appelle “adhérente”, c'est-à-dire celle qui s'attache à un objet doté d'une finalité utilitaire (une église, un habit), ne reconnaît pas cette liberté de jugement pour la beauté qu'il appelle “libre”, qui est en action dans une œuvre d'art.

Les positions de Kant sont longuement décrites en section 2.1. On trouvera par ailleurs dans [Ferry, 1990] une discussion approfondie de l'apport de Kant à la pensée moderne sur l'esthétique et dans [Schaper, 1964, Hopkins, 2001, Crowther, 1976, Aquila, 1970] des discussions en termes contemporains sur l'interprétation que l'on peut apporter à la pensée kantienne.

Georg William Friedrich Hegel

Dans son très long traité, intitulé “L'Esthétique”, Hegel remet à plat tous les concepts de Beau et enrôle l'esthétique dans une conception globale et rationnelle de l'individu. Partant d'un raisonnement dialectique, dégagé des schèmes antérieurs, il conclut cependant comme les platoniciens que c'est une influence objective de l'objet sur l'observateur qui crée le Beau et non une interprétation subjective comme le suggérerait Locke par exemple³⁴. L'originalité de la pensée de Hegel est de considérer le Beau comme une expérience intellectuelle, fruit de notre raisonnement³⁵ et non résultat de nos sensations. Il prend le contrepied des Subjectivistes et se méfie de nos sensations³⁶. Mais parallèlement, il est

32. Baumgarten, *Esthétique* §37, rapporté dans [Malinowski-Charles, 2005].

33. Les textes d'Emmanuel Kant se rapportant à l'esthétique sont en grande partie issus de *Critique de la faculté de juger*, de 1790 [Kant, 1985].

34. “Tout ce qui existe n'a donc de vérité qu'autant qu'il est l'idée passée à l'état d'existence. Car l'idée seule est la véritable réalité. Autrement dit, tout ce qui apparaît aux sens n'est pas vrai simplement parce qu'il a une existence interne ou externe, ou parce qu'il est réel comme tel, mais simplement parce que cette réalité correspond au concept. C'est seulement par là que l'existence acquiert l'actualité et la vérité. Cette vérité n'est pas du tout à prendre dans le sens subjectif selon lequel il y aurait un accord entre une existence et mes idées, mais dans le sens objectif selon lequel moi ou un objet extérieur, une action, un événement, une situation dans sa réalité constituent en eux-mêmes une réalisation du concept.” [Hegel, 1997], p. 176.

35. “... la beauté artistique est la beauté engendrée et ré-engendrée par l'esprit et, de la même façon que l'esprit et ses productions sont plus élevés que la nature et ses phénomènes, de même le beau artistique est plus élevé que la beauté de la nature.” [Hegel, 1997], p. 52.

36. “La pire façon d'appréhender les choses, celle qui convient le moins à l'esprit, est la façon purement sensible.” [Hegel, 1997], p. 91.

amené à condamner l'approche des experts en "Bel Art"³⁷. Dans la critique artistique il a cependant une vision très classique de la hiérarchie esthétique et ses arguments sont peu argumentés, ressortant surtout des canons classiques³⁸. En peinture, c'est la lumière qui porte l'essentiel du message du peintre. C'est par elle que le peintre fait ressortir tel ou tel objet car il est par ailleurs astreint à respecter formes et couleurs données par le modèle. A ce titre, la lumière est la partie "subjective" de l'œuvre créatrice³⁹, mais c'est une subjectivité du créateur, non de l'observateur. Enfin, comme pour Kant, pour lui, "*Le beau doit lui aussi être reconnu universellement.*" [Hegel, 1997], p. 117.

Arthur Schopenhauer

A la différence des précédents auteurs, Arthur Schopenhauer n'a pas écrit d'ouvrage exclusivement consacré à l'esthétique. Ses contributions à ce sujet sont à retrouver essentiellement dans deux textes : *Sur la vue et les couleurs*, de 1815 et *Le monde comme volonté et comme représentation*, de 1818 et 1844.

Dans le premier texte, qu'il rédige après des discussions avec Goethe dont il emprunte la théorie de la perception chromatique, il reprend la distinction kantienne entre perception et vision ainsi que la conception phénoménologique du monde extérieur. La couleur n'est que le produit de notre subjectivité et ne saurait être rattachée à l'objet concerné. La couleur n'est pas présente dans la lumière⁴⁰. Les teintes de référence (le jaune, l'orange, le bleu et le violet⁴¹) sont des innés de notre conscience. Elles sont issues d'une subdivision de l'intervalle noir/blanc en des fractions pythagoriciennes simples, à la façon de la gamme musicale. Cet ouvrage se veut comme une déconstruction de l'approche newtonnienne qui, à son avis, en ne s'intéressant qu'à la physique, passe à côté de la notion profondément subjective de la perception colorée. On trouve dans ce texte des éléments qui s'appliquent à l'esthétique : un subjectivisme de principe qui crée phénoménologiquement les propriétés de l'objet par la seule représentation qu'en a le spectateur, un référentiel qualitatif inné, qui privilégiera la spontanéité et l'intuition sur le raisonnement conscient, un socle de concepts platoniciens : simplicité, harmonie, proportions, largement universels et transcendants.

Dans *Le monde comme volonté et comme représentation*, Schopenhauer consacre un long chapitre à l'art et au génie⁴². Dans une approche qui doit beaucoup à Kant, il distingue la perception d'instinct (qu'il nomme intuition) de la déduction raisonnée (le principe de raison) qu'il met sous la coupe de la volonté. Le génie n'est donné qu'à ceux qui sont capables de cette perception impulsive, expression profonde de l'état de nature (il appelle en cela Nietzsche qui poussera à son paroxysme cette idée). C'est cette intuition naïve et native qui conduit l'artiste à une représentation plus exacte de la nature et qui donne à son œuvre ses qualités universellement reconnues⁴³. Les arts et les mathématiques sont, pour Schopenhauer, deux domaines où le principe de raison, qu'il considère pourtant comme l'outil privilégié pour acquérir les concepts indispensables à l'élaboration de la pensée, se révèle inférieur à l'intuition⁴⁴. Il y a également dans cette dualité intuition/volonté un message à l'attention de celui

37. "*Là où se déclenchent les grandes passions et les émotions d'une âme profonde, il n'y a plus de place pour les subtiles distinctions du goût et son petit commerce de minuties ... à l'homme de goût ou au juge doué de goût en matière d'art, a succédé le connaisseur.*" [Hegel, 1997], p. 89.

38. Il définit ainsi, en termes platoniciens, la beauté de la forme abstraite : "*régularité, symétrie, conformité aux lois et harmonie des formes et des couleurs.*" [Hegel, 1997], p204.

39. "*Si nous demandons maintenant quel est l'élément physique dont se sert la peinture, c'est la lumière, qui rend visibles les objets du monde extérieur en général. (...) La lumière, par cette identité idéale, offre le seul côté qui réponde aux principes de la subjectivité. Sous ce rapport, elle a la propriété de rendre visibles les objets.*" [Hegel, 1997], II- p. 225.

40. Comme Goethe, Schopenhauer est très marqué par les couleurs rémanentes qui persistent après une vision colorée, ou celles qui apparaissent spontanément dans l'obscurité totale.

41. En d'autres places, il suggère que les couleurs principales sont au nombre de six, par paires : jaune/violet, orange/bleu, rouge/vert.

42. Il est intitulé : *Livre troisième : le monde comme représentation* et s'ouvre par une citation de Platon.

43. "*(...) ces merveilleux peintres hollandais qui ont contemplé d'une intuition si objective les objets les plus insignifiants et qui nous ont donné dans leurs tableaux d'intérieur une preuve impérissable de leur objectivité, de leur sérénité d'esprit ; un homme de goût ne peut contempler leur peinture sans émotion (...)*" (in [Schopenhauer, 1966], p. 240.)

44. On ne pourra manquer de rapprocher ces idées de celles qui ont été débattues avec beaucoup de vigueur depuis

qui observe une œuvre d'art ⁴⁵.

Comme Kant et Hegel, Schopenhauer disserte longuement sur les propriétés spécifiques du sublime, dont l'importance était en son temps beaucoup plus importante qu'aujourd'hui. Le sublime est pour lui un état de ravissement qui exclut toute capacité à mobiliser le principe de raison. Le spectacle qui conduit au sublime exprime un état d'hostilité franche à la volonté.

Friedrich von Schelling

Il nous faut ici évoquer le nom de Friedrich Wilhelm Joseph von Schelling, auteur d'une *Philosophie de l'art* écrite entre 1802 et 1805 mais publiée après sa mort, qui s'inscrit dans la lignée kantienne. Pourtant cette œuvre concerne très peu notre propos car elle dégage l'art de l'influence de l'esthétique [David, 2002] et ne le considère que sous l'aspect de sa genèse et de sa place dans la philosophie : pour lui, "*tout art est mythologie*", comme le dit J.F. Marquet [Marquet, 1983].

La bascule romantique

Friedrich Nietzsche

C'est probablement le philosophe qui a eu l'impact le plus sensible sur des générations d'artistes, et pourtant ses leçons sont complexes et changeantes, et l'on peut sans peine tirer de son œuvre des recommandations contradictoires. Friedrich Nietzsche s'est exprimé sur l'esthétique dans "*Le crépuscule des idoles*" de 1888 [Nietzsche, 2005], mais il a semé des messages sur ce thème dans de nombreux autres écrits, en particulier dans *la naissance de la tragédie*, *La volonté de puissance* ou *L'Antéchrist*. Sa contribution à l'esthétique est amplement discutée dans [Ferry, 1990].

Disons tout d'abord pour fixer l'enjeu de la réflexion esthétique de Nietzsche qu'il considère l'art comme le summum des activités humaines, et la seule activité qui mérite véritablement un investissement total.

Nietzsche dénonce la cérébralisation de l'esthétique classique, avec ses notions d'ordre, de règles, d'intelligence (ce qu'il appelle l'art apollonien), cérébralisation qui a pour ressorts la dialectique platonicienne largement défendue par Hegel ⁴⁶, qui pour lui n'est qu'une duperie intellectuelle. Il revendique au contraire un art de la liberté, de la démesure, de la jouissance irraisonnée, de l'ivresse (un art qu'il appelle dionysiaque) ⁴⁷. Il revendique une esthétique individuelle, égoïste, élitiste ; il prône la participation aux fêtes orgiaques et païennes, loin des cérémonies religieuses et compassées des religions chrétiennes. Il recherche un sentiment primitif sans référence au rationnel et à la réflexion.

1950 dans le champ de la production artistique et qui ont accompagné les succès d'artistes comme J. Pollock, G. Matthieu ou JM. Basquiat, ou celles que défendait J. Dubuffet sous les termes d'art brut : "*Nous entendons par là des ouvrages exécutés par des personnes indemnes de culture artistique, dans lesquels le mimétisme, contrairement à ce qui se passe chez les intellectuels, aït peu ou pas de part, de sorte que leurs auteurs y tirent tout (...) de leur propre fond et non pas des poncifs de l'art classique ou de l'art à la mode. Nous y assistons à l'opération artistique toute pure, brute, réinventée dans l'entier de toutes ses phases par son auteur, à partir seulement de ses propres impulsions.*" (Jean Dubuffet, *L'art brut préféré aux arts culturels*, Gallimard, 1967).

45. "*Ce n'est que par cette contemplation pure et toute entière absorbée dans l'objet que l'on conçoit les idées; l'essence du génie consiste dans une aptitude éminente à cette contemplation; elle exige un oubli complet de la personnalité et de ses relations; ainsi la génialité n'est pas autre chose que l'objectivité la plus parfaite, c'est-à-dire la direction objective de l'esprit, opposée à la direction subjective qui aboutit à la personnalité, c'est-à-dire à la volonté. Par suite, la génialité consiste à se maintenir dans l'intuition pure et à s'y perdre.*" (in [Schopenhauer, 1966], p. 240.)

46. D'après Deleuze, la position esthétique compliquée de Nietzsche ne peut s'interpréter qu'en la rapportant à son objectif premier : démonter le message philosophique de Hegel, [Deleuze, 1962], cité dans [Ferry, 1990].

47. "*Que signifie les oppositions d'idées entre apollinien et dionysien (...) ? L'ivresse apollinienne produit avant tout l'irritation de l'œil qui donne à l'œil la faculté de vision. Le peintre, le sculpteur, le poète épique sont des visionnaires par excellence. Dans l'état dionysien, par contre, tout le système émotif est irrité et amplifié : en sorte qu'il décharge d'un seul coup tous ses moyens d'expression, en expulsant sa force d'imitation, de reproduction, de transfiguration, de métamorphose, toute espèce de mimique et d'art d'imitation.*" [Nietzsche, 2005].

Il fait confiance à la nature première de l'homme - pourvu qu'il ne soit pas vulgaire - pour débusquer le beau véritable au sein de la nature.

Et pourtant Nietzsche refuse aussi la démarche esthétique des Romantiques (qui ne lui en garderont pas rigueur!) en particulier lorsqu'il critique Wagner auquel il reproche d'être trop attentif aux plaisirs physiologiques vulgaires. Il se retourne alors vers une esthétique classique et vient soutenir que Corneille est préférable à Victor Hugo.

(à suivre)

Annexe II : L'esthétique en Chine

“La Mer possède le déferlement immense, la Montagne possède le recel latent. La Mer engloutit et vomit, la Montagne se prosterne et s'incline. La Mer peut manifester une âme, la Montagne peut véhiculer un rythme. [...] La Montagne, c'est la Mer et la Mer c'est la Montagne. Montagne et Mer connaissent la vérité de ma perception : tout réside en l'Homme par le libre élan du Pinceau, de la seule Encre.”

Shitao, dans la traduction de [Cheng, 1991], pp. 94-95

L'image du lettré chinois

Consacrons ici quelques lignes à l'esthétique chinoise, définitivement différente de l'esthétique occidentale. Pour cela nous emprunterons largement à quelques ouvrages :

- à l'œuvre de François Jullien tout d'abord. *“La grande image n'a pas de forme”* a dit Lao Tseu ; F. Jullien, philosophe et sinologue, reprend cette phrase en titre de son œuvre sur l'esthétique chinoise [Jullien, 2003],
- à l'ouvrage de François Cheng, écrivain, poète et calligraphe, membre de l'Académie, pétri des deux cultures [Cheng, 1991],
- aux mémoires de Françoise Verdier, peintre formée à la dure école classique chinoise [Verdier, 2001, Verdier, 2003],
- aux conférences et aux textes de Xun Jiang, écrivain et historien d'art contemporain de Taiwan [Xun, 2015, Xun, 2016], aimablement traduits par le Pr. Sun Hong,
- au recueil de traductions de textes anciens, commentés par Yolaine Escande [Escande, 2003], sinologue et philosophe de l'esthétique.

Nous l'avons vu, l'esthétique occidentale est essentiellement portée par les travaux philosophiques de la Grèce du V^e siècle avant notre ère. Elle repose sur une imitation fidèle de la nature, privilégiant des critères d'harmonie tant dans la construction de l'œuvre, que dans la reproduction des détails ou le choix des couleurs.

L'esthétique chinoise de la grande tradition artistique de l'Empire du Milieu, trouve ses origines à peu près à la même période que l'esthétique grecque et comme elle, elle a fait l'objet, très tôt de traités qui jalonnent 25 siècles d'une intense pratique. A la différence de l'art occidental, elle ne semble pas avoir subi de révolution, les œuvres les plus anciennes demeurant des références respectées que

n'affecte que lentement l'évolution des goûts jusqu'à l'achèvement de la dynastie Qing et l'émergence des républiques⁴⁸. Les règles qui président à la construction de la "Grande Image" sont à rechercher dans l'émergence des trois grands mouvements de pensée : le confucianisme du V^e au III^e siècle avant notre ère, le taoïsme aux III^e et II^e siècles avant notre ère, puis le bouddhisme à partir du I^{er} siècle de notre ère. Ces influences philosophiques, radicalement différentes de la logique platonicienne, ont conduit à prescrire des pratiques définitivement différentes de celles que nous avons rencontrées tout au long de cet ouvrage.

L'esthétique chinoise s'est beaucoup concentrée sur les arts visuels : la peinture d'une part⁴⁹, la calligraphie d'autre part qui tient en Chine une place éminente parmi les beaux-arts⁵⁰. Elle est moins concernée par la musique et les arts plastiques : statuaire, porcelaine ... qui relèvent pour elle d'un autre champ que les arts graphiques. Nous ne nous intéresserons qu'aux arts graphiques et plus particulièrement à la peinture et au dessin pour ce qu'ils pourraient nous aider à élaborer des règles de beauté pour la photographie.

Mais ce terme même de beauté semble mal adapté pour caractériser une peinture chinoise réussie. Le peintre chinois ne vise pas la beauté il vise à faire partager une expérience intime à l'observateur, à l'intégrer au monde qu'il crée⁵¹. Plus qu'une expérience esthétique, la peinture chinoise est une mystique [Cheng, 1991].



Figure 10.4 – *Rendre visible l'invisible, faire s'exprimer la fluidité de l'eau, la stabilité de la montagne, la mouvance des bambous. A gauche : "Une représentation en image du poème de Du Fu" par Zhao Kui (1185-1266), dynastie Song du Sud - A droite : "Pêche retirée dans la rivière" par Wen Boren (1502-1575), dynastie Ming - Musée de Shanghai.*

Si l'art grec est tout entier porté par la distinction aristotélicienne entre les objets qui nous entourent et la pensée qui les appréhende, s'il s'efforce de construire des images de ces objets de

48. L'empire Qing s'effondre en 1912, suivi de périodes de grands troubles politiques. Sur le plan culturel, l'esthétique profondément ancrée dans la tradition s'ouvre progressivement aux influences occidentales. Mais il faut la guerre civile (1927) puis la guerre sino-japonaise (en 1931, en Mandchourie, puis en 1937, sur l'ensemble du territoire), pour que bascule la civilisation chinoise dans une longue parenthèse qui a brutalement rompu le fil de la tradition antique ; il faut se tourner vers Taïwan pour trouver maintenu un filet persistant de cette tradition bimillénaire.

49. "En Chine, de tous les arts, la peinture occupe la place suprême." écrit F. Cheng ([Cheng, 1991], p. 11), en l'étayant de citations.

50. Il faut considérer la calligraphie chinoise comme l'aboutissement de la peinture et il n'est pas une règle, valable pour la calligraphie qui ne soit transposable à la peinture.

51. Cela est bien exprimé à propos de la calligraphie par Cyrille Javary : "Pourtant un fait est avéré : jamais on ne dit en chinois d'une calligraphie qu'elle est "belle". Ce serait la qualifier d'inerte. En poésie, en peinture aussi, l'esthétique chinoise est passée du concept de beauté pure, celle que la raison grecque rêvait au ciel des Idées, pour lui préférer celui, plus fugace d'une beauté vive." (Cyrille J.-D. Javary : *L'écrire étincelant* in [Verdier, 2001]). On lit également : "Le critère en art n'est pas le beau, notion subjective qui varie suivant les lieux et les époques, mais la sincérité, l'authenticité.", par la bouche du peintre Huang Yang, rapporté dans [Verdier, 2003], et dans [Escande, 2003] : "l'esthétique en tant que science du beau n'existe pas dans l'art chinois".

façon que la pensée les retrouve à partir de la toile (dont le rideau de Parrhasios est le suprême exemple⁵²), l'art chinois s'appuie sur une conception opposée du monde, dans laquelle le peintre est partie prenante de la scène qu'il observe. L'objectif est alors de rendre compte de cette immersion de façon que l'observateur la partage également. Cette immersion qui réunit le peintre inspiré et le paysage qu'il s'apprête à reproduire, est matérialisée par le "souffle-énergie" (le *qi*), souffle vital alterné *yin/yang* qui traverse toute chose et donne à l'univers sa forme⁵³. Saisir le *qi* de façon qu'il irrigue l'observateur est le but de l'art chinois, comme saisir la beauté est le but de l'art occidental. Ceci est moins affaire d'habileté que de communion avec le paysage. Ce n'est pas non plus affaire de perception ou de perspective. Comme le rappelle François Jullien, "*l'évidence objective est détruite, on n'a plus besoin d'une nature préétablie supposant un objet déjà donné.*"⁵⁴. Les détails fins pourront être ignorés et la ressemblance gagnera à être lointaine, car, et ainsi parle Jing Hao⁵⁵ : "*la ressemblance, c'est atteindre la forme et laisser échapper le souffle énergie.*" ou encore : "*quand le souffle-énergie est perdu au niveau de l'image phénomène, c'est la mort de l'image.*"⁵⁶, et F. Cheng rajoute : "*manquer de souffle, c'est le signe même d'une peinture médiocre.*"⁵⁷.



Figure 10.5 – Dessiner avec le vide, choisir le trait simple et efficace pour évoquer l'idée. Le trait jaillit sans travail, rendu nécessaire par l'objet. A droite : "Paysage (détail)" par Hong Ren (1610-1663), dynastie Qing - Au centre : "Fleurs (détail d'un rouleau)" par Chen Chun (1483-1544), dynastie Ming. Musée de Shanghai.

Plus encore qu'une immersion, la peinture chinoise vise un partenariat entre la scène et le peintre⁵⁸ : "*il n'y a plus objet perçu et sujet percevant, mais corrélation et échange entre les pôles - hôtes accueillant/accueilli*" dit F. Jullien. Comme l'exprime Shitao⁵⁹ à la fin de sa vie, maintenant qu'il sait peindre, "*le paysage me charge de parler à sa place...*".

52. Le grand peintre Zeuxis (considéré comme le plus grand de la Grèce antique), lors d'un concours contre Parrhasios avait peint des raisins merveilleux que les moineaux venaient picorer. Convaincu par cette expérience d'avoir atteint le suprême beau, il s'attendait à gagner le prix mais voulut que Parrhasios au préalable, retirât le rideau qui cachait son œuvre. Ce rideau était peint et Zeuxis dut reconnaître qu'il était battu puisque lui-même avait été berné tandis qu'il n'avait pu tromper que des moineaux.

53. Ainsi, on lit dans François Cheng : "*On sait que la pensée esthétique chinoise, fondée sur une conception organiciste de l'univers, propose un art qui tend depuis toujours à recréer un microcosme total où prime l'action unificatrice du Souffle-Esprit, où le vide même, loin d'être synonyme de flou ou d'arbitraire, est le lieu interne où s'établit le réseau des souffles vitaux.*" [Cheng, 1991], p. 10.

54. Cité dans [Jullien, 2003] p. 107-108

55. Jing Hao (855-915) est un peintre réputé et un lettré raffiné du Henan.

56. Les deux références citées dans [Jullien, 2003] p. 227)

57. Cité dans [Cheng, 1991], p. 72

58. Cette vision orientale de l'intime union de l'homme à son paysage n'est pas sans écho chez nos philosophes occidentaux. Dans [Merleau-Ponty, 1964], l'un des maîtres de la phénoménologie s'exprime ainsi : "*Visible et mobile, mon corps est au nombre des choses, il est l'une d'elles, il est pris dans le tissu du monde et sa cohésion est celle d'une chose. Mais, puisqu'il voit et se meut, il tient les choses en cercle autour de soi, elles sont une annexe ou un prolongement de lui-même, elles sont incrustées dans sa chair, elles font partie de sa définition pleine et le monde est fait de l'étoffe même de son corps. (...) la vision est prise (...) là ou persiste (...) l'indivision du sentant et du senti.*"

59. Shitao est le célèbre Moine Citrouille-amère (1642-1708), du Sud de la Chine. Son œuvre est connue en Europe par la traduction qu'en a faite Pierre Ryckmans.

Rien n'est plus stimulant pour le peintre chinois que de tenter de rendre ce que l'on ne perçoit pas : la disparition de la montagne dans la brume, la trace du poisson dans l'eau, la fin de la pluie sur des pêcheurs, le passage matinal de la nuit au jour. La peinture chinoise se nourrit de la polarité "être/ne pas être" et développe pour cela un outil qu'utilise moins la peinture occidentale : le vide⁶⁰. Le vide (*hsü* ou *t'ai-hsü*) est le lieu privilégié d'interaction entre les natures bipolaires ; c'est lui qui permet que ces polarités fusionnent au lieu d'entrer en conflit ([Cheng, 1991] p. 51). Le vide, dans la peinture chinoise, se trouve à tous les niveaux : dans la composition en réservant de grandes masses à l'absence, dans le tracé entre les brins chargés ou allégés de couleur du pinceau, dans l'opacité de l'encre ou de la peinture qui s'éteint en fumée, dans les branchages des arbres, suspendus loin du tronc, dans les tissus des vêtements, emportés par une brise. Selon Chang Shih, il est bon que seul un tiers du papier seulement soit couvert. L'esquisse est en cela la forme la plus accomplie de l'art chinois, qui laisse les formes inachevées et laisse ainsi amplement le spectateur envahir et se faire envahir par l'espace décrit.

Au contraire du peintre occidental, qu'il soit portraitiste florentin, intimiste flamand ou paysagiste français, le peintre chinois ne cherche pas à remplir sa toile. Il joue avec l'espace, les pleins et les creux et laisse à penser ce qu'il ne peint pas. Ne cherchez pas une carte fidèle des lieux, une représentation précise des personnes. Toute face a des yeux, pourquoi les représenter s'ils ne doivent rien apporter ? Le format même de la toile défie nos pratiques. Le rouleau s'étale très loin autour du lac, franchit les montagnes, accompagne un fleuve tranquille, comme le tour d'horizon du promeneur sur la montagne, ou, à l'inverse, la grappe de glycine orne en cascade les lances des bambous tandis que les vols des hirondelles percent les nuages et rejoignent les têtes escarpées des falaises dans une interminable proclamation verticale.



Figure 10.6 – *Simplifier le trait et laisser deviner l'absent ; le regard du peintre ne décrit pas, il exprime : "Scène poétique au bord du lac Shihu" par Wen Zhengming (1470-1559), dynastie Ming. Musée de Shanghai.*

Le mélange des couleurs procède lui-même de cette vision cosmologique de l'univers. On en trouve les principes déjà mentionnés dans le *Mémoire sur les artisans* du *Rituel des Zhou* de la dynastie Han (453-221 av JC), cité dans [Escande, 2003], bien avant notre ère. Ces principes sont encore intacts aujourd'hui quoiqu'ils furent parfois discutés⁶¹. Les couleurs premières sont au nombre de cinq. A chaque couleur est associée une composante originelle de l'univers⁶² : le bois (vert/bleu⁶³), le métal (blanc), la terre (jaune), l'eau (noir) et le feu (rouge). Les associations possibles de couleurs suivent les affinités de la nature où l'eau ni le bois ne se mélangent au feu. Si dans les temps anciens, les

60. On trouve pourtant dans V. Jankélévitch : "C'est dans l'inachevé qu'on laisse la vie s'installer.", cité dans [Verdier, 2003], p. 131.

61. En particulier par Zhuang Zi, le philosophe du taoïsme, qui leur reproche de brider l'imagination.

62. La relation couleur/matière est discutée en détail dans *La pensée chinoise* de M. Granet (1968), Albin Michel, cité dans [Escande, 2003].

63. Le vert/bleu, dénommée *qing*, qui n'est pas exactement notre cyan ne doit pas être confondu avec le vert (*lu*), ni avec le bleu (*lan*).

couleurs sont très souvent utilisées sous des formes éclatantes, il fut progressivement recommandé qu'elles s'expriment dans des teintes plutôt rabattues qui prédominent dans la grande période et s'imposent sur les œuvres qui remplissent majoritairement les musées.

Objective ou subjective ?

Nous serions bien en peine de replacer maintenant l'esthétique chinoise dans le débat qui a été développé dans cet ouvrage. Il est clair que l'on ne peut nulle part plaquer les termes de l'analyse platonicienne et de l'objectivité esthétique à l'œuvre chinoise. Les objets ne sont pas beaux en eux, ils sont beaux dans leur relation à leur contexte et le beau émane de la réunion et de la conjonction des présences par une grâce délocalisée, le *qi*. Les plus beaux objets de la Grèce ancienne : le corps humain, le nu, n'évoquent aucun charme pour le connaisseur chinois, et au contraire, le summum de l'art chinois, dans ses représentations des paysages où la majesté des montagnes se mêle à la fluidité des rivières et la mouvance des forêts, dans une balance subtile de *yin* et de *yang*, aurait laissé totalement indifférent l'esthète occidental, au moins jusqu'à ce que les peintres florentins et lorrains n'en fassent matière d'étude. Copier exactement la nature pour s'approcher de son modèle, percer les règles de l'éclairage, des ombres, de la perspective, c'est indigne du peintre chinois qui apprend au contraire le trait rare et signifiant, l'ellipse, l'illusion. Le "canon" également n'aurait aucun sens car si l'exemple des anciens est vénéré, copier les méthodes est indigne du peintre⁶⁴.

Il nous faut pourtant souligner que l'esthétique chinoise partage avec l'esthétique occidentale antique la plus objectiviste la conviction que le beau (ou ce qui en tient lieu dans un monde qui ne lui a pas attribué la même valeur) est également universellement partagé comme ultime trame cosmique, reflet ici d'une harmonie, là d'un souffle vital. Cette trame ultime est unique et ouverte à tous qui la percevront identique pour peu qu'ils acquièrent les vertus nécessaires (c'est vrai autant pour le *kalos kagathos* athénien que pour le lettré chinois). A ce titre, le chef d'œuvre chinois ne souffre pas de critique. Il s'impose et s'imposera durablement.

Si pour la majorité des critères, l'art chinois ne s'accorde pas aux préceptes objectivistes de l'esthétique, est-ce qu'il est en accord avec les conceptions subjectivistes ? Nous en sommes loin. Certes l'objectif du peintre est bien de faire ressentir intimement à l'observateur l'instant vécu lors de la peinture par l'artiste ; certes il y a bien l'idée que l'œuvre d'art ne prend pas forme tant qu'elle n'est éprouvée par le spectateur et que le succès de l'artiste se mesure à l'ampleur de l'adhésion du spectateur. Mais l'esthétique chinoise refuse de reconnaître la promotion artistique aux capacités et aux facultés de l'observateur à répondre à la perception par une contribution active et instrumentée par son tempérament, sa cognition, son raisonnement. Au contraire, l'effet esthétique ne peut exister que par la passivité complète de l'individu qui se laisse traverser par le souffle-énergie : "*Que votre cœur soit vide et dégagé sans la moindre poussière, et le paysage surgira du plus intime de votre âme.*", dit Wang Yu⁶⁵, cité dans [Verdier, 2001], et Xun Jiang⁶⁶, aujourd'hui, exclut toute action consciente dans la perception de la beauté qui n'est que "*le réveil de l'âme*" et dans le beau qui est "*l'équilibre du funambule entre sentiment et ressenti d'une part, et entre excitation et plaisir d'autre part*" [Xun, 2015].

Objectivisme / Subjectivisme, ni l'un ni l'autre semble répondre aux attentes de l'esthète chinois. On s'approcherait en adoptant une démarche phénoménologiste qui accorderait une grande place à l'intuition (Schopenhauer⁶⁷, Bergson, Merleau-Ponty). Mais ceci est encore une autre histoire.

64. Sur ce point : doit-on ou non copier les maîtres, la littérature chinoise est partagée comme l'atteste Xie He, peintre et critique d'art des dynasties Liu Song et Qi du Sud, donc vers 550, qui recommande, en dernier point de ses "Six principes de peinture" de peindre non seulement en s'inspirant de la nature, mais aussi en copiant les maîtres.

65. Wang Yu, qui signait ses œuvres Dongzhuang, est un peintre chinois né dans la province de Jiangsu vers 1650 et mort en 1729.

66. Xun Jiang est un éminent historien et critique d'art contemporain, taiwanais, qui a publié de nombreux ouvrages traçant des parallèles entre art oriental et occidental.

67. Rappelons ici ces mots de Schopenhauer, cités plus complètement dans la note 28 du chapitre 2 : "*(...) la condition subjective du plaisir esthétique (...) consiste à affranchir la connaissance que la volonté asservirait, à oublier le moi individuel, à transformer la conscience en un sujet connaissant pur et affranchi de la volonté, du temps, de toute*

Que pourra faire l'Intelligence artificielle pour l'esthétique chinoise ?

Il n'échappera pas au lecteur que cette peinture chinoise traditionnelle résiste aux analyses que nous avons suivies dans les derniers chapitres de cet ouvrage. Est-il possible d'imaginer une façon de réintégrer cette esthétique dans les schémas utilisés par les scientifiques pour mécaniser l'évaluation esthétique d'une œuvre ?

Il nous faut à ce point séparer la réflexion en deux niveaux, selon que l'on procède à la détection préalable de critères de beauté (ce qui a été dénommé *handcrafted*, au chapitre 7) ou que l'on confie de bout-en-bout le traitement à un réseau de neurones.

Handcrafted primitives Dans cette approche il nous faudrait identifier tous les critères quantifiables du beau au sens de l'esthétique chinoise. Il conviendrait donc de faire un inventaire, comparable à celui du chapitre 7 à partir de la littérature chinoise.

Il se trouve que tout l'enseignement de la peinture en Chine, repose sur des règles de production d'une belle œuvre : les six lois de Xie He⁶⁸. Ces lois sont considérées comme un guide absolu pour le peintre, intangibles et éternelles et tout apprenti doit s'y conformer rigoureusement⁶⁹. Tous les enseignements débutent ainsi par les règles de Xie He. Mais elles nous sont parvenues sous forme de six mots de quatre caractères chacun, et ces mots en calligraphie ancienne, très hermétiques, ont donné lieu à de nombreuses traductions et des interprétations parfois fort éloignées⁷⁰. Malheureusement, leur forme très abstraite nous laisse très démunis pour en tirer des enseignements pour notre objectif. Il nous faut rechercher des traits plus objectifs que nous pourrions mesurer dans la photo.

A notre connaissance un inventaire de ces propriétés n'a pas été fait ni pour la peinture et ni pour la photo.

Pourtant, à partir des textes concernant la peinture, il serait possible de relever des traits importants concernant, par exemple, la place laissée au blanc à divers niveaux, de l'attache de la feuille sur la branche, au vide des prairies ou à l'ouverture des montagnes et des nuées. Il serait possible aussi d'identifier les couples en polarité *yin/yang* : montagne/rivière⁷¹, arbre/pierre, barque/pêcheur, chemin/torrent, pont/bateau, etc. Il faudrait aussi repérer les traces de disparition, le trait qui s'épuise, le fondu du chemin dans la forêt, l'évanouissement de la montagne dans la brume, l'évanescence des lointains, l'incomplétude des personnages ... Au contraire, et dans une opposition notable, le trait trouve en Chine une vigueur exceptionnelle, il balaie le détail et exprime crûment la ligne principale avec la force du geste longuement réfléchi, "cette impulsion motrice procédant du souffle vital (*qi ji*)"

relation."

68. Xie He (ou Sheikh) est un des premiers théoriciens de l'esthétique chinoise. Il a vécu au 6^e siècle. Il était aussi peintre, mais ses œuvres n'ont pas atteint le niveau de célébrité de son ouvrage principal : *Guhua Pinlu, Appréciation des peintures anciennes*.

69. On lit ainsi sur le site de l'Académie centrale des beaux arts de Beijing : "*La proposition des six lois a eu une influence décisive sur la théorie de la peinture postérieure [à Xie He]. Les six lois sont toujours comme la norme permettant de décider du niveau de qualité d'une peinture.*"

70. Nous en donnons ici la traduction du sinologue Yolande Escande ([Escande, 2003] p. 296) qui nous semble la plus aisément compréhensible :

1. la résonance des souffles donne vie et mouvement,
2. la méthode de l'os [i.e. du squelette, de l'ossature] guide le maniement du pinceau,
3. respecter la fidélité à l'objet en figurant les formes,
4. assurer la conformité aux genres en appliquant les couleurs,
5. porter attention à l'arrangement convenable dans la composition,
6. permettre la transmission par la copie [des maîtres].

la hiérarchie de ces règles est stricte, les deux premières prévalent sur les deux suivantes et à fortiori sur les deux dernières. Certains ont vu dans ces 6 règles un écho des 6 recommandations de Vitruve : ordre, arrangement, eurhythmie, symétrie, propriété et économie.

71. "Montagne rivière", traduit littéralement *shan shui*, le couple qui désigne génériquement le paysage, faisant porter au mot collectif la marque de la polarité.

comme le dit F. Jullien, citant Jiang He ([Jullien, 2003], p.279). Cela s'exprime probablement dans son contraste, ses courbures, ses orientations.

Il est certainement souhaitable d'accorder une grande attention à la peinture à l'encre, qui tient la première place dans l'esthétique chinoise. Prédominance de l'encre sur la peinture et puissance du pinceau (le "*bi mo*"⁷²), qui amènent à reconnaître 6 "nuances" : le noir, le blanc, le sec, le mouillé, l'épais et le fluide comme le rappelle Tang Dai⁷³ cité dans [Jullien, 2003].

Il serait probablement également souhaitable de s'attacher au détail élémentaire, dont l'importance est primordiale dans la peinture chinoise : le trait. La qualité du trait de pinceau est le fruit d'une éducation de nombreuses années⁷⁴. La singularité du "beau trait" qui n'échappe pas au lettré chinois pourrait probablement se révéler à une analyse morphologique fine par ordinateur. Nous n'avons pas trouvé de travaux allant dans ce sens.

Enfin, l'absence d'ombre est un trait obligé de la peinture chinoise qui y voit la manifestation d'un *yin* qu'il faut exclure des images fortes. Cette contrainte, très facilement mise en œuvre en peinture (et pour cela peu discriminante dans l'évaluation de sa beauté) devrait, en photographie, prendre une place plus importante si elle fait aussi partie des critères tacites d'esthétique. Néanmoins il ne nous semble pas que cela ait jamais fait l'objet d'études systématiques.

Réseaux neuromimétiques profonds Nous avons vu, dans les chapitres précédents, que la question du choix des primitives est rapidement devenue un problème secondaire en raison des progrès considérables faits par les méthodes par RNP qui font l'impasse sur ces primitives. On imagine donc qu'il y aurait moyen d'éviter l'inventaire amorcé au paragraphe précédent en nourrissant directement un réseau par des images réputées conformément aux prescriptions des lettrés chinois.

A notre connaissance aucune approche n'a été faite pour éduquer un ordinateur à reconnaître le beau au sens de la Grande Image. Bien plus, il semble que ce soit justement de l'Empire du Milieu que nous viennent les travaux les plus nombreux et souvent les plus avancés⁷⁵, s'appuyant sur une approche platonicienne de l'esthétique. Faut-il croire que le monde numérique tourne le dos au monde sensible dont s'abreuve la peinture chinoise traditionnelle ? Probablement pas, mais tout laisse à penser que c'est poussés par la possibilité facile de mettre en œuvre simplement les outils conformes à une esthétique objectiviste que les informaticiens se sont engouffrés dans la brèche offerte par les réseaux sociaux et les bases de photographies en ligne. Simple opportunité donc.

S'il nous fallait apprendre à un ordinateur à distinguer le Beau conformément aux critères esthétiques chinois, il ne serait certainement pas très satisfaisant de remplacer pièce pour pièce, d'une part la base de données AVA, d'autre part le recueil des évaluations issues de DPChallenge, pour entraîner une architecture de RNP à l'esthétique chinoise. D'une part il faudrait tout d'abord collecter quelques centaines de milliers de "belles images" issues des meilleures collections des musées chinois, puis, pour chacune de ces œuvres, recueillir un nombre raisonnable d'évaluations fiables de façon à les engager dans la piste de l'apprentissage par RNP. Suggérons alors que l'on profite de l'expérience menée sur la photographie "occidentale" pour remettre en cause l'approche résolument objectiviste adoptée jusqu'à ces jours.

Peut-être est-il encore trop tôt pour concevoir des pistes nouvelles. Il est probablement souhaitable d'attendre que la connaissance claire des limites de l'approche objectiviste se dégage des expérimentations en cours pour que soient développées d'autres pistes qui porteront peut-être des fruits plus riches. Il ne serait pas alors vraiment surprenant que l'on doive reprendre à zéro les prin-

72. Le *bi mo* exprime à la fois le petit attirail de préparation de l'encre par broyage de la poudre et un style de peinture d'atelier, comme nous disons "aquarelle" ou "estampe".

73. Tang Dai est un peintre du XVIII^e siècle, originaire de Mandchourie et auteur d'un important traité pédagogique sur la peinture.

74. On lit ainsi le programme pédagogique du maître Huang Yuan, dans [Verdier, 2003], : "*Tu vas commencer par tracer des traits, seulement des traits pendant plusieurs mois. En peinture chinoise, tout se construit avec des traits (...) Il faut d'abord que tu possèdes une base solide. Cette base, c'est le trait horizontal, nous ne passerons pas aux autres traits ...*".

75. Un examen attentif des auteurs cités en référence aux chapitres 7 et 8 nous en convaincra.

cipes de l'apprentissage et les mécanismes du raisonnement afin de mieux respecter la subtilité de l'esthétique du lettré chinois.

Bibliographie

- [Alberti, 1992] Alberti, L. (1992). *De pictura*. Editions Macula.
- [Alvarez et al., 1999] Alvarez, L., Gousseau, Y., and Morel, J. M. (1999). The size of objects in natural and artificial images. *Avances in Imaging & Electron Physics*, 111 :167–242.
- [Amengual et al., 2015] Amengual, X., Bosch, A., and de la Rosa, J. L. (2015). Review of methods to predict social image interestingness and memorability. In Springer, editor, *International Conference on Computer Analysis of Images and Patterns, Cham.*, pages 64–76.
- [Amirshahi et al., 2013] Amirshahi, S., Denzler, J., and Redies, C. (2013). Jenaesthetics—a public dataset of paintings for aesthetic research. In *European Conference on Computer Vision (Workshop)*, Jena, Germany.
- [Amirshahi et al., 2014] Amirshahi, S., Hayn-Leichsenring, G., Denzler, J., and Redies, C. (2014). Evaluating the rule of thirds in photographs and paintings. *Art & Perception*, 2(1–2) :163–182.
- [André, 1759] André, J. (1759). *Essai sur le Beau*. Schneider, J.H., Amsterdam.
- [Aquila, 1970] Aquila, R. (1970). A new look at Kant’s aesthetic-judgement. *Kant Studien*, 1-4 :17–34.
- [Arbellini, 2017] Arbellini, J. (2017). Les photographies d’art ont-elles un format compatible avec le nombre d’or? Student project, Institut Villebon-Georges Charpak, Telecom-ParisTech Dept. IDS.
- [Aristote, 1990] Aristote (1990). *Poétique*. Le livre de poche, (Paris).
- [Arnheim, 1954] Arnheim, R. (1954). *Art and Visual Perception : A Psychology of the Creative Eye*. University of California Press, Los Angeles (CA).
- [Arnheim, 1983] Arnheim, R. (1983). *The power of the center. A study of composition of the visual arts*. Univ. of California Press.
- [Arnheim, 1986] Arnheim, R. (1986). *New essays on the psychology of art*. University of California Press, Oakland (CAL).
- [Attewell and Baddeley, 2007] Attewell, D. and Baddeley, R. (2007). The distribution of reflectances within the visual environment. *Vision Research*, 47(4) :548–554.
- [Augustin et al., 2011] Augustin, M. D., Carbon, C., and Wagemans, J. (2011). Measuring aesthetics impression. In *ECVP’11*, Toulouse.
- [Augustin et al., 2018] Augustin, M. D., Wagemans, J., Carbon, C. C., Holmes, T., Kapoula, Z., and Roberts, M. N. (2018). Measuring aesthetic impressions. (under revision).
- [Barsalou, 1985] Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of experimental psychology : learning, memory, and cognition*, 11(4) :629.
- [Batteux, 1747] Batteux, C. (1747). *Les Beaux Arts réduits à un même principe*. Librairie Durand, Paris.
- [Beardsley, 1966] Beardsley, M. (1966). *Aesthetics from classical Greece to the present*. McMillan, New-York.
- [Bell, 1914] Bell, C. (1914). *Art*. Chatto & Windus (London).

- [Bell et al., 1991] Bell, S. S., Holbrook, M. B., and Solomon, M. R. (1991). Combining esthetic and social value to explain preferences for product styles with the incorporation of personality and ensemble effects. *Bell, S. S., Holbrook, M. B., & Solomon, M. R. (1991). Combining esthetic and social value to explain preferences for product styles with the incorporation of personality and ensemble effects. Journal of Social behavior and Personality, 6(6), 243., 6(6) :243.*
- [Bénard, 1877] Bénard, C. (1877). L'esthétique du laid. *Revue philosophique de la France et de l'étranger*, Année 2 - tome IV :233–265.
- [Benouaret, 2017] Benouaret, I. (2017). *Un système de recommandation contextuel et composite pour la visite personnalisée de sites culturels*. PhD thesis, Thèse de l'Université de Technologie de Compiègne (France).
- [Bense, 1969] Bense, M. (1969). *Einführung in die informationstheoretische Ästhetik., Grundlegung und Anwendung in der Texttheorie*. Rowoldt Taschenbuch Verlag.
- [Bense, 2007] Bense, M. (2007). *Aesthetica*. Les Editions du Cerf, Passages, Paris.
- [Bergson, 2009] Bergson, H. (2009). *La pensée et le mouvant*. Presses Universitaires de France (Paris).
- [Berlyne, 1970] Berlyne, D. (1970). Novelty, complexity and hedonic values. *Perception and Psychophysics*, 8 :279–290.
- [Berlyne, 1971] Berlyne, D. (1971). *Aesthetics and Psychobiology*, volume 336. Appleton-Century Croft, New-York.
- [Berridge and Kringelbach, 2013] Berridge, K. and Kringelbach, M. (2013). Neuroscience of affect : brain mechanisms of pleasure and displeasure. *Current Opin Neurobiology*, 23(3) :294–303.
- [Birkhoff, 1933] Birkhoff, G. (1933). *Aesthetic Measure*. Harvard University Press, Cambridge (USA).
- [Blijlevens et al., 2017] Blijlevens, J., Thurgood, C., Hekkert, P., Chen, L. L., Leder, H., and Whitfield, T. W. (2017). The aesthetic pleasure in design scale : The development of a scale to measure aesthetic pleasure for designed artifacts. *Psychology of Aesthetics, Creativity, and the Arts*, 11, 11(1) :86–97.
- [Boccia et al., 2016] Boccia, M., Barbetti, S., Piccardi, L., Guariglia, C., Ferlazzo, F., Giannini, A. M., and Zaidel, D. W. (2016). Where does brain neural activation in aesthetic responses to visual art occur ? meta-analytic evidence from neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, 60 :65–71.
- [Bouleau, 2014] Bouleau, C. (2014). *The Painter's Secret Geometry : A Study of Composition in Art*. Allegro editions.
- [Boulez et al., 2014] Boulez, P., Changeux, J., and Manoury, P. (2014). *Les neurones enchantés*. Odile Jacob, Paris.
- [Bourdieu, 1965] Bourdieu, P. (1965). *Un art moyen : essai sur les usages sociaux de la photographie*. Les Editions de Minuit, Paris.
- [Brady and Field, 2000] Brady, N. and Field, D. (2000). Local contrast in natural images : normalisation and coding efficiency. *Perception*, 29 :1041–1055.
- [Brown and Dissanayake, 2009] Brown, S. and Dissanayake, E. (2009). The Arts are more than Aesthetics : Neuroaesthetics as narrow aesthetics. *Neuroaesthetics*, pages 43–57.
- [Brown et al., 2011] Brown, S., Gao, X., Tisdelle, L., Eickoff, S., and Lotti, M. (2011). Naturalizing aesthetics : Brain areas for aesthetic appraisal across sensory modalities. *Neuroimages*, 58 :250–258.
- [Bréhin, 2007] Bréhin, Y. (2007). Une lecture de la "beauté réelle" d'Eddy Zemach. *Marges : revue d'art contemporain*, 6 :112–119.
- [Bullot and Reber, 2013] Bullot, N.-J. and Reber, R. (2013). The artful mind meets art history : Towards a psycho-historical framework for the science of art appreciation. *Behavioural and brain sciences*, 36 :123–180.
- [Cela-Conde et al., 2011] Cela-Conde, C., Agnati, L., Huston, J., Mora, F., and Nadal, M. (2011). The neural foundations of aesthetic appreciation. *Progress in Neurobiology*, 94 :39–48.

- [Chandler, 2013] Chandler, R. (2013). Seven challenges in image quality assessment : past, present and future research. *ISRN Signal Processing*, 2013 :1–53.
- [Changeux, 2008] Changeux, J. (2008). *Du vrai, du beau, du bien : une nouvelle approche neuronale*. Odile Jacob.
- [Changeux, 2016] Changeux, J. (2016). *La beauté dans le cerveau*. Odile Jacob, Paris.
- [Charrier et al., 2012] Charrier, C., Lezoray, O., and Lebrun, G. (2012). Machine learning to design full-reference image quality assessment algorithm. *Signal Processing : Image Communications*, 27-3 :209–219.
- [Chastel, 1996] Chastel, A. (1996). *Marsile Ficin et l'Art*. Droz, Genève.
- [Chatterjee, 2003] Chatterjee, A. (2003). Prospects for a cognitive neuroscience of visual aesthetics. *Bulletin of Psychology and the Arts*, 4(2) :55–60.
- [Chatterjee, 2004] Chatterjee, A. (2004). Prospects for a cognitive neuroscience of visual aesthetics. *Bulletin of psychology and the arts*, 4(2) :55–60.
- [Chatterjee and Vartanian, 2016] Chatterjee, A. and Vartanian, O. (2016). Neuroscience of aesthetics. *Annals of the New-York Academy of Sciences*, 1369 :172–194.
- [Cheng, 1991] Cheng, F. (1991). *Vide et plein, le langage pictural chinois*. Editions du Seuil, (Paris).
- [Chevreul, 1864] Chevreul, E. (1864). *Des couleurs et leurs applications aux arts industriels à l'aide des cercles chromatiques*. J.B. Baillières et Fils, Paris.
- [Chopra et al., 2005] Chopra, S., Hadsell, R., and LeCun, Y. (2005). Learning a similarity metric discriminatively with application to face verification. In Soc., I. C., editor, *CVPR'05*, volume 1.
- [Cometti et al., 2000] Cometti, J. P., Morizot, J., and Pouivet, R. (2000). *Questions d'esthétique*. Presses universitaires de France.
- [Cotton, 2009] Cotton, C. (2009). *The photograph as contemporary art*, volume 1. London : Thames & Hudson.
- [Crettez, 2017] Crettez, J. (2017). *Les supports de la géométrie interne des peintres*. ISTE Editions, London.
- [Crowther, 1976] Crowther, P. (1976). Fundamental ontology and transcendent beauty : an approach to Kant's aesthetics. *Kant Studien*, 1-4 :55–71.
- [Cupchnik et al., 2009] Cupchnik, G., Vartanian, O., Crawley, A., and Mikulis, D. (2009). Viewing artworks : Contributions of cognitive control and perceptual facilitation for aesthetic experience. *Brain and Cognition*, 70 :84–91.
- [Dalal and Triggs, 2005] Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *CVPR conference Computer Vision and Pattern Recognition*, pages 886–893, San-Diego Etats-Unis.
- [Damasio, 1994] Damasio, A. (1994). *Descartes' error : emotion, reason and the human brain*. Grosset/Putnam.
- [Damasio, 1999] Damasio, A. (1999). *The feeling of what happens :Body, emotion and consciousness*. Harcourt Brace (New-York).
- [Danto, 1964] Danto, A. (1964). The artworld. *The journal of philosophy*, 61(19) :571–584.
- [Danto, 1992] Danto, A. (1992). *Beyond the Brillo Box : The Visual Arts in Post-Historical Perspective*. University of California Press.
- [Danto and Goehr, 2014] Danto, A. and Goehr, L. (2014). *After the End of Art : Contemporary Art and the Pale of History*. Princeton Classics.
- [Datta et al., 2006] Datta, R., Joshi, D., Li, J., and Wang, J. (2006). Studying aesthetics in photographic images using a computational approach. In *Computer Vision, ECCV 2006*, volume 3953 of *Lecture Notes in Computer Science*, pages 288–301.
- [Datta et al., 2008] Datta, R., Li, J., and Wang, J. Z. (2008). Algorithmic inferencing of aesthetics and emotion in natural images : An exposition. In IEEE, editor, *15th IEEE International Conference on Image Processing.*, pages 105–108.

- [Datta and Wang, 2010] Datta, R. and Wang, J. (2010). ACQUINE : Aesthetic quality inference engine - real-time automatic rating of photo-aesthetics. In ACM, editor, *MIR'10*.
- [David, 2002] David, P. (2002). Schelling : construction de l'art et récusation de l'esthétique. *Revue de Métaphysique et de morale*, 34(2) :29–41.
- [de Tommaso et al., 2008] de Tommaso, M., Pecoraro, C., Sardaro, M., Serpino, C., Lancioni, G., and Livrea, P. (2008). Influence of aesthetic perception on visual event-related potentials. *Consciousness and cognition*, 17(3) :933–945.
- [DeCrousaz, 1724] DeCrousaz, J. (1715-1724). *Traité du beau, où l'on montre en quoi consiste ce que l'on nomme ainsi, par des exemples tirés de la plupart des arts et des sciences*. Fayard (janvier 1985).
- [Delahaye, 2015] Delahaye, J. P. (2015). La beauté mise en formules. *Pour la Science*, 455 :78–83.
- [Deldjoo et al., 2018] Deldjoo, Y., Constantin, M. G., Ionescu, B., Schedl, M., and Cremonesi, P. (2018). Mmtf-14k : A multifaceted movie trailer feature dataset for recommendation and retrieval. In *Proceedings of the 9th ACM Multimedia Systems Conference*, pages 450–455.
- [Deldjoo et al., 2016] Deldjoo, Y., Elahi, M., Cremonesi, P., Garzotto, F., Piazzolla, P., and Quadrona, M. (2016). Content-based video recommendation system based on stylistic visual features. *Journal on Data Semantics*., 5(2) :99–113.
- [Deleuze, 1962] Deleuze, G. (1962). *Nietzsche et la philosophie*. PUF, (Paris).
- [Deng et al., 2017a] Deng, Y., Loy, C. C., and Tang, X. . (2017a). Aesthetic-driven image enhancement by adversarial learning. arXiv preprint arXiv :1707.05251.
- [Deng et al., 2017b] Deng, Y., Loy, C. C., and Tang, X. (2017b). Image aesthetic assessment : An experimental survey. *IEEE Signal Processing Magazine*, 34(4) :80–106.
- [Dennett, 1993] Dennett, D. (1993). *La conscience expliquée*. Odile Jacob, (Paris).
- [Descartes, 1991] Descartes, R. (1991). *Discours de la méthode suivi de La Dioptrique*. Gallimard Folio (Paris).
- [Dessalles, 2008] Dessalles, J. (2008). *La pertinence et ses origines cognitives*. Hermès- Sciences, Paris.
- [Dessalles, 2013] Dessalles, J. (2013). Algorithmic simplicity and relevance. In Dowe, D., editor, *Algorithmic probability and friends*, volume LNAI, pages 119–130, Berlin. Springer-Verlag.
- [Dewar, 1937] Dewar, H. (1937). A comparison of tests of artistic appreciation. *British Psychological Society*, pages 29–49.
- [Dézarnaud-Dandine and Sevin, 2007] Dézarnaud-Dandine, C. and Sevin, A. (2007). *Symétrie m'était contée ... histoires de symétries*. Ellipses, Paris.
- [Dhar et al., 2011] Dhar, S., Ordonez, V., and Berg, T. (2011). High level describable attributes for predicting aesthetics and interestingness. In on, I. C., editor, *Computer Vision and Pattern Recognition (CVPR), 2011*, pages 1657–1664.
- [Di Dio et al., 2011] Di Dio, C., Canessa, N., Cappa, S., and Rizzolatti, G. (2011). Specificity of esthetic experience for artworks : an fmri study. *Frontiers in human Neuroscience*, 5(139) :1–14.
- [Di Dio et al., 2007] Di Dio, C., Macaluso, E., and Rizzolatti, G. (2007). The golden beauty : brain response to classical and renaissance sculptures. *PloS one*, 2(11) :e1201.
- [Di Dio and Vittorio, 2009] Di Dio, C. and Vittorio, G. (2009). Neuroaesthetics : a review. *Current Opinion in Neurobiology*, 19 :682–687.
- [Diderot, 1769] Diderot, D. (1769). *Correspondance littéraire*. Furne, Paris.
- [Dominguez et al., 2017] Dominguez, V., Messina, P., Parra, D., Mery, D., Trattner, C., and Soto, A. (2017). Comparing neural and attractiveness-based visual features for artwork recommendation. In *Proceedings of the 2nd ACM Workshop on Deep Learning for Recommender Systems*, pages 55–59.
- [Dufrenne, 1980] Dufrenne, M. (1980). *Esthétique et Philosophie*. Klincksleck esthétique, Paris.
- [Dupont, 2015] Dupont, J. (2015). Que faire de l'imagerie cérébrale ? - Territoires anciens et nouveaux d'une technologie. *Comptes Rendus Biologie*, 8-9 :607–612.

- [Elahi et al., 2017] Elahi, M., Deldjoo, Y., Bakhshandegan Moghaddam, F., Cella, L., Cereda, S., and Cremonesi, P. (2017). Exploring the semantic gap for movie recommendations. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, pages 326–330.
- [Elton, 1954] Elton, W. R. (1954). *Aesthetics and language*. Philosophical Library.
- [Escande, 2003] Escande, Y. (2003). *Traité chinois de peinture et de calligraphie. Tome 1 : les textes fondateurs (des Han aux Sui)*. Klincksieck (Paris).
- [Evans et al., 1992] Evans, A., Collins, D., and Milner, B. (1992). An MRI-based stereotactic atlas from 250 young normal subjects. *Journ of Soc. Neurosci.*, 18 :412.
- [Eysenck, 1939] Eysenck, H. (1939). The general factor in aesthetic judgements. *British Psychological Society*, pages 94–102.
- [Fang and Zhang, 2017] Fang, H. and Zhang, M. (2017). Creatism : A deep-learning photographer capable of creating professional work. arXiv preprint arXiv :1707.03491.
- [Fechner, 1871] Fechner, G. (1871). *Zur Experimentalen Aesthetik*. Hirtzel, S., Leipzig.
- [Fernandez and Wilkins, 2008] Fernandez, D. and Wilkins, A. (2008). Uncomfortable images in art and nature. *Perception*, 37(7) :1098–1113.
- [Ferry, 1990] Ferry, L. (1990). *Homo aestheticus, l'invention du goût à l'âge démocratique*. Grasset, Le Livre de Poche (Essais).
- [Fize, 2004] Fize, D. (2004). *La catégorisation visuelle rapide*, chapter 4 in *Imagerie cérébrale fonctionnelle électrique et magnétique* (Renault, B. ed), pages 69–93. Lavoisier (Paris).
- [Freedberg and Gallese, 2007] Freedberg, D. and Gallese, V. (2007). Motion, emotion and empathy in esthetic experience. *Trends in Cognitive Sciences*, 11(5) :197–203.
- [Fried, 2008] Fried, M. (2008). *Why photography matters as art as never before*. New Haven : Yale University Press.
- [Fu et al., 2014] Fu, Y., Hospedales, T. M., Xiang, T., Gong, S., and Yao, Y. (2014). Interestingness prediction by robust learning to rank. In Springer, editor, *European conference on computer vision*, pages 488–503.
- [Gärdenfors, 2000] Gärdenfors, P. (2000). *Conceptual Spaces, the Geometry of Thought*. MIT Press, a Bradford book, Cambridge (Mass.).
- [Ghosal et al., 2019] Ghosal, K., Rana, A., and Smolic, A. (2019). Aesthetic image captioning from weakly-labeled photographs. arXiv :1908.11310v1.
- [Ghyka, 1931] Ghyka, M. (1931). *Le nombre d'or*. Gallimard, Paris.
- [Gibson, 1986] Gibson, J. (1986). *The ecological approach to visual perception*. Lawrence Erlbaum Associates (Hillsdale).
- [Gombrich, 1960] Gombrich, E. H. (1960). *Art and Illusion : a study in the psychology of pictorial representations*. Princeton University Press, Wahington D.C.
- [Gombrich, 2000] Gombrich, E. H. (2000). Concerning ‘the science of art’ : Commentary on ramachandran and hirstein. *Journal of consciousness studies*, 7(8-9) :17–17.
- [Gousseau and Roueff, 2007] Gousseau, Y. and Roueff, F. (2007). Modeling occlusion and scaling in natural images. *SIAM Journal of Multiscale Modeling and Simulation*, 6(1) :105–134.
- [Gygli et al., 2013] Gygli, M., Grabner, H., Riemenschneider, H., Nater, F., and Van Gool, L. (2013). The interestingness of images. In *IEEE International Conference on Computer Vision (ICCV)*, pages 1633–1640.
- [Hariman and Lucaites, 2016] Hariman, R. and Lucaites, J. L. (2016). Photography : The abundant art. *Photography and Culture*, 9(1) :39–58.
- [Hayn-Leichsenring et al., 2017] Hayn-Leichsenring, G. U., Lehmann, T., and Redies, C. (2017). Subjective ratings of beauty and aesthetics : correlations with statistical image properties in western oil paintings. *i-Perception*, 8(3).
- [He et al., 2015a] He, K., Zhang, X., Ren, S., and Sun, J. (2015a). Deep residual learning for image recognition. ArXiv arXiv :1512.03385, Microsoft Beijing.

- [He et al., 2015b] He, K., Zhang, X., Ren, S., and Sun, J. (2015b). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans on PAMI*, 37(9) :1904–1907.
- [He et al., 2016] He, R., Fang, C., Wang, Z., and McAuley, J. (2016). Vista : a visually, socially, and temporally-aware model for artistic recommendation. In *Proceedings of the 10th ACM Conference on Recommender Systems*, pages 309–316.
- [Hegel, 1838] Hegel, G. (1835-1838). *Vorlesungen über die Ästhetik*. recueilli par H.G. Hohto.
- [Hegel, 1997] Hegel, G. (1997). *Esthétique ou Philosophie de l'art*, volume 1 & 2. Le Livre de Poche, Paris.
- [Henry, 1885] Henry, C. (1885). *Introduction à une esthétique scientifique*. Revue contemporaine, Paris.
- [Henry, 1889] Henry, C. (1889). *Éléments d'une théorie générale de la dynamogénie, autrement dit, du contraste, du rythme et de la mesure avec applications spéciales aux sensations visuelles et auditives*. Charles Verdin, Paris.
- [Henry, 1891] Henry, C. (1891). *Physiologie générale des sensations et esthétique*, pages 1–71. Notices sur les travaux scientifiques de M. Charles Henry. Imprimerie des sciences mathématiques et physiques, Via Lombardia, Rome.
- [Henry, 1895] Henry, C. (1895). Quelques aperçus sur l'Esthétique des formes. *La Revue Blanche*, pages 6–61.
- [Hii et al., 2017] Hii, Y. L., See, J., Kairanbay, M., and Wong, L. K. (2017). Multigap : Multi-pooled inception network with text augmentation for aesthetic prediction of photographs. In *IEEE International Conference on Image Processing (ICIP)*, pages 1722–1726.
- [Hillyard and Anllo-Vento, 1998] Hillyard, S. A. and Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences*, 95(3) :781–787.
- [Hoffer and Ailon, 2015] Hoffer, E. and Ailon, N. (2015). Deep metric learning using triplet network. In *Lectures Notes in Computer Sciences N 9370*.
- [Hong et al., 2016] Hong, R., Zhang, L., and Tao, D. (2016). Unified photo enhancement by discovering aesthetic communities from flickr. *IEEE transactions on Image Processing*, 25(3), 25(3) :1124–1135.
- [Hopkins, 2000] Hopkins, R. (2000). Beauty and testimony. In *Philosophy, the Good, the True and the Beautiful*, number 47 in Royal Institute of Philosophy, pages 209–236. O'Hear, A.
- [Hopkins, 2001] Hopkins, R. (2001). Kant, quasi-realism and the autonomy of aesthetic judgement. *European Journal of Philosophy*, 9(2) :166–189.
- [Howard et al., 2017] Howard, A., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and H., A. (2017). Mobilenets : Efficient convolutional neural networks for mobile vision applications. arXiv :1704.04861.
- [Hsieh et al., 2014] Hsieh, L. C., Hsu, W. H., and Wang, H. C. (2014). Investigating and predicting social and visual image interestingness on social media by crowdsourcing. In IEEE, editor, *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference*, pages 4309–4313.
- [Hsu, 2009] Hsu, L. (2009). *Le visible et l'expression. Etude sur la relation intersubjective entre perception visuelle, sentiment esthétique et forme picturale*. PhD thesis, Ecole des Hautes Etudes en Sciences Sociales, Paris. HAL id : tel00401739.
- [Hubner and Fillingner, 2016] Hubner, R. and Fillingner, M. (2016). Comparaison of objective measures for predicting perceptual balance and visual aesthetic preferences. *Frontiers in Psychology*, 7(335) :1 :15.
- [Huneman and Kulich, 1997] Huneman, P. and Kulich, E. (1997). *Introduction à la phénoménologie*. Armand Collin (Paris).

- [Hunter and Askarinejad, 2015] Hunter, M. R. and Askarinejad, A. (2015). Designer’s approach for scene selection in tests of preference and restoration along a continuum of natural to manmade environments. *6, 1228*, 6 :1228.
- [Hurlberg and Ling, 2007] Hurlberg, A. and Ling, Y. (2007). Biological components of sex differences in color preference. *Current Biology*, 17(16) :R623–R625.
- [Hurlberg and Ling, 2012] Hurlberg, A. and Ling, Y. (2012). Understanding colour perception and preference. In Best, J., editor, *Colour design, theories and applications*, volume 12, chapter 5, pages 129–158. Woodhead Publishing.
- [Ibarra et al., 2017] Ibarra, F. F., Kardan, O., Hunter, M. R., Kotabe, H. P., Meyer, F. A., and Berman, M. G. (2017). Image feature types and their predictions of aesthetic preference and naturalness. *Frontiers in psychology*, 8 :632–648.
- [Ingarden, 2011] Ingarden, R. (2011). *Esthétique et ontologie de l’œuvre d’art. Choix de textes 1937-1969*. J. Vrin, Paris.
- [Iseminger, 1981] Iseminger, G. (1981). Aesthetic appreciation. *The Journal of Aesthetics and Art Criticism*, 39, 4 :389–397.
- [Ishizu and Zeki, 2011] Ishizu, T. and Zeki, S. (2011). Toward a brain-based theory of beauty. *PLOS One*.
- [Isola et al., 2011] Isola, P., Xiao, J., Torralba, A., and Oliva, A. (2011). What makes an image memorable? In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 145–152.
- [Isola et al., 2017] Isola, P., Zhu, J. Y., Zhou, T., and Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. arXiv preprint.
- [Itten, 1973] Itten, J. (1973). *The art of color*. Wiley, New-York.
- [Jacobsen et al., 2006] Jacobsen, T., Schubotz, R., Höfel, L., and Cramon, D. (2006). Brain correlates of aesthetic judgment of beauty. *Neuroimage*, 29 :276–285.
- [Jin et al., 2016a] Jin, B., Ortiz-Segovia, V., and Süsstrunk (2016a). Image aesthetic predictors based on weighted CNNs. In *ICIP, Int. Conf. on Image Processing 2016*, volume 1, pages 2291–2295.
- [Jin et al., 2009] Jin, W., Ho, H., and Srihari, R. (2009). OpinionMiner : a novel machine learning system for web opinion mining and extraction. In *Proc ACM SIGKDD, KDD’09*, pages 1195–1204, New-York.
- [Jin et al., 2016b] Jin, X., Chi, J., Peng, S., Tian, Y., Ye, C., and Li, X. (2016b). Deep image aesthetics classification using inception modules and fine-tuning connected layer. In *8th IEEE International Conference on Wireless Communications & Signal Processing (WCSP)*, pages 1–6.
- [Jullien, 2017] Jullien, C. (2017). Les mathématiques : le langage de la beauté. *Maths Langages express*, pages 93–98.
- [Jullien, 2003] Jullien, F. (2003). *La grande image n’a pas de forme*. Editions du Seuil.
- [Kairanbay et al., 2017] Kairanbay, M., See, J., Wong, L. K., and Hii, Y. L. (2017). Filling the gaps : Reducing the complexity of networks for multi-attribute image aesthetic prediction. In *IEEE International Conference on Image Processing (ICIP)*, pages 3051–3055.
- [Kandinsky, 1954] Kandinsky, W. (1954). *Du spirituel dans l’art, et dans la peinture en particulier*. Denoël, (Paris).
- [Kant, 1985] Kant, E. (1985). *Critique de la faculté de juger*. Gallimard, Folio-Essais, Paris.
- [Kao et al., 2017] Kao, Y., He, R., and Huang, K. ((2017)). Deep aesthetic quality assessment with semantic information. *IEEE Transactions on Image Processing*, 26(3) :1482–1495.
- [Kao et al., 2016] Kao, Y., Huang, K., and Maybank, S. (2016). Hierarchical aesthetic quality assessment using deep convolutional neural networks. *Signal Processing : Image Communication*, 47 :500 :510.
- [Kaplan et al., 1972] Kaplan, S., Kaplan, R., and Wendt, J. S. (1972). Rated preference and complexity for natural and urban visual material. *Perception & Psychophysics*, 12(4) :354–356.

- [Kardan et al., 2017] Kardan, O., Demiralp, E., Hout, M. C., Hunter, M. R., Karimi, H., Hanayik, T. and Yourganov, G., Jonides, J., and Berman, M. G. (2017). Is the preference of natural versus man-made scenes driven by bottom-up processing of the visual features of nature? *Frontiers in psychology*, 6 :471–484.
- [Kawabata and Zeki, 2004] Kawabata, H. and Zeki, S. (2004). Neural correlates of beauty. *Journal of Neurophysiology*, 91(4) :1699–1705.
- [Ke et al., 2006] Ke, Y., Tang, X., and Jing, F. (2006). The design of high-level features for photo quality assessment. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, volume 1, pages 419–426.
- [Kemp, 1990] Kemp, M. (1990). *The Science of Art, Optical themes in western art from Brunelleschi to Seurat*. Yale university Press.
- [Khosla et al., 2012] Khosla, A., Xiao, J., Isola, P., Torralba, A., and Oliva, A. (2012). Image memorability and visual inception. In ACM, editor, *SIGGRAPH Asia 2012 Technical Briefs*, page 35.
- [Kim et al., 2013] Kim, J., Yoon, S., and Pavlovic, V. (2013). Relative spatial features for image memorability. In ACM, editor, *Proceedings of the 21st ACM international conference on Multimedia*, pages 761–764.
- [Kim et al., 2015] Kim, W. H., Choi, J. H., and Lee, J. S. (2015). Subjectivity in aesthetic quality assessment of digital photographs : Analysis of user comments. In *Proceedings of the 23rd ACM international conference on Multimedia*, pages 983–986.
- [Kirsch et al., 2016] Kirsch, L. P., Urgesi, C., and Cross, E. S. (2016). Shaping and reshaping the aesthetic brain : Emerging perspectives on the neurobiology of embodied aesthetics. *Neuroscience & Biobehavioral Reviews*, 62 :56–68.
- [Kivy, 1968] Kivy, P. (1968). Aesthetic aspects and aesthetic qualities. *The journal of philosophy*, LXV(4) :85 :93.
- [Koch et al., 2010] Koch, M., Denzler, J., and Redies, C. (2010). 1/f² characteristics and isotropy in the fourier power spectra of visual art, cartoons, comics, mangas, and different categories of photographs.
- [Koelsch et al., 2015] Koelsch, S., Jacobs, A. M., Menninghaus, W., Liebal, K., Klann-Delius, G., von Scheve, C., and Gebauer, G. (2015). The quartet theory of human emotions : an integrative and neurofunctional model. *Physics of life reviews*, 13 :1–27.
- [Kong et al., 2016] Kong, S., Shen, X., Lin, Z., Mech, R., and Fowlkes, C. (2016). Photo aesthetics ranking network with attributes and content adaptation. In *ECCV European Conference on Computer Vision*, pages 662–679. Springer International Publishing.
- [Koshelev et al., 1998] Koshelev, M., Kreinovich, V., and Yann, Y. (1998). Towards the use of aesthetics in decision making : Kolmogorov complexity formalizes Birkhoff’s idea. In *Digitalcom-mons@UTEP*.
- [Kotabe et al., 2017] Kotabe, H. P., Kardan, O., and Berman, M. G. (2017). The nature-disorder paradox : A perceptual study on how nature is disorderly yet aesthetically preferred. *Journal of Experimental Psychology : General*, 146(8), 1126., 145(8) :1126–1142.
- [Kringelbach, 2005] Kringelbach, M. (2005). The human orbito-frontal cortex : linking reward to hedonic experience. *Nat. Rev Neurosciences*, 6 :691–702.
- [Krizhevsky et al., 2012] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105.
- [Kuehn and Gallinat, 2012] Kuehn, S. and Gallinat, J. (2012). The neural correlates of subjective pleasantness. *Neuroimage*, 61 :289–294.
- [Kurdi et al., 2017] Kurdi, B., Lozano, S., and Banaji, M. R. (2017). Introducing the open affective standardized image set (oasis). *Behavior research methods*, 49(2) :457–470.
- [Lacey et al., 2011] Lacey, S., Hagtveldt, H., Patrick, V. M., Anderson, A., Stilla, R., Deshpande, G., HU, X., Sato, J., Reddy, S., and Sathian, K. (2011). Art for reward’s sake : Visual art recruits the ventral striatum. *Neuroimage*, 55(1) :420–433.

- [Lamarque, 2010] Lamarque, P. (2010). The uselessness of art. *The Journal of Aesthetics and Art Criticism*, 68(3) :205–214.
- [Lang et al., 1999] Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (1999). International affective picture system (iaps) : Instruction manual and affective ratings (tech. rep. no. a-4). Technical report, University of Florida, Gainesville, USA.
- [Lange and James, 1922] Lange, C. and James, W. (1922). *The emotions*. Williams & Wilkins.
- [Laughlin, 1987] Laughlin, S. (1987). Form and function in retinal processing. *Trends in Neurosciences*, 10(11) :478–483.
- [Le Bihan, 2012] Le Bihan, D. (2012). *Le cerveau de cristal*. Odile Jacob, Paris.
- [Leder et al., 2004] Leder, H., Belke, B., Oeberst, A., and Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgements. *British Journal of Psychology*, 95 :489–508.
- [Leder et al., 2006] Leder, H., Carbon, C. C., and Ripsas, A. L. (2006). Entitling art : Influence of title information on understanding and appreciation of paintings. *Acta psychologica*, 121(2) :176–198.
- [Leder and Nadal, 2014] Leder, H. and Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments : The aesthetic episode—developments and challenges in empirical aesthetics. *British Journal of Psychology*, 105(4) :443–464.
- [LeDoux, 2000] LeDoux, J. (2000). Emotion circuits in the brain. *Ann. Rev. Neurosciences*, 23 :155–184.
- [Lengger et al., 2007] Lengger, P. G., Fischmeister, F. P. S., Leder, H., and Bauer, H. (2007). Functional neuroanatomy of the perception of modern art : A DC-EEG study on the influence of stylistic information on aesthetic experience. *Brain research*, 1158 :93–102.
- [Levinson, 1980] Levinson, J. (1980). Aesthetic uniqueness. *The Journal of Aesthetics and Art Criticism*, 38(4) :435–449.
- [Ling and Hurlberg, 2007] Ling, Y. and Hurlberg, A. (2007). A new model for color preference : Universality and individuality. In *Color and Imaging Conference*, pages 8–11.
- [Lipps, 1903] Lipps, T. (1903). *Ästhetik : Psychologie des Schönen und der Kunst*. Voss, L. Hamburg.
- [Liu et al., 2017] Liu, J., Lughofer, E., and Zeng, X. (2017). Toward model building for visual aesthetic perception. *Computational intelligence and neuroscience*, pages 1–13.
- [Liu et al., 2018] Liu, J., Lughofer, E., Zeng, X., and Li, Z. (2018). The power of visual texture in aesthetic perception : An exploration of the predictability of perceived aesthetic emotions. *Computational Intelligence and Neuroscience*, pages 1–9.
- [Liu et al., 2010] Liu, L., Chen, R., Wolf, L., and Cohen-Or, D. (2010). Optimizing photo composition. *Computer Graphics Forum*.
- [Livingstone, 2002] Livingstone, M. (2002). *Vision and Art : the biology of Seeing*. Harry N. Abrams. Inc., Publishers.
- [Livio, 2008] Livio, M. (2008). *The golden ratio : the story of Phi, the world’s most astonishing number*. Broadway Books.
- [Lo et al., 2012] Lo, K., Liu, K., and Chen, C. (2012). Assessment of photo aesthetics with efficiency. In *IEEE International Conference on Pattern Recognition (ICPR)*,, pages 2186–2189.
- [Lo et al., 2013] Lo, K., Liu, K., and Chen, C. (2013). Intelligent photographing interface with on-device aesthetic quality assessment. In Parks, J. and Kim, J., editors, *ACCV 2012 Workshop*, volume LNCS 7729, pages 533–544, Berlin. Springer Verlag.
- [Locke, 2009] Locke, J. (2009). *Essai sur l’entendement humain (traduction P. Coste)*. Le Livre de Poche (les Classiques de la Philosophie).
- [Lu et al., 2014a] Lu, P., Kuang, Z., Peng, X., and Li, R. (2014a). Discovering harmony : a hierarchical colour harmony model for aesthetics assesment. In *Asian Conference on Computer Vision*, pages 452–467. Springer Int.
- [Lu et al., 2015a] Lu, P., Peng, X., Li, R., and Wang, X. (2015a). Towards aesthetics of image : A Bayesian framework for color harmony modeling. *Signal Processing : Image Communication*, 39 :487–498.

- [Lu et al., 2016] Lu, P., Peng, X., Zhu, X., and Li, R. (2016). An EL-LDA based general color harmony model for photo aesthetics assessment. *Signal Processing*, 120 :731–745.
- [Lu et al., 2014b] Lu, X., Lin, Z., Jin, H., Yang, J., and Wang, J. (2014b). RAPID : Rating pictorial aesthetics using Deep Learning. In ACM, editor, *22nd Int. Conf on Multimedia MM'14*, pages 457–466, Orlando, FL.
- [Lu et al., 2014c] Lu, X., Lin, Z., Jin, H., Yang, J., and Wang, J. Z. (2014c). Rating image aesthetics using deep learning. *IEEE Transactions on Multimedia*, 17(11) :2921–2034.
- [Lu et al., 2015b] Lu, X., Lin, Z., Shen, X., Mech, R., and Wang, J. Z. (2015b). Deep multi-patch aggregation network for image style, aesthetics, and quality estimation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 990–998.
- [Luffarelli et al., 2019] Luffarelli, J., Stamatogiannakis, A., and Yang, H. (2019). The visual asymmetry effect : An interplay of logo design and brand personality on brand equity. *Journal of Marketing Research*, 56(1).
- [Luo and Tang, 2008] Luo, Y. and Tang, X. (2008). Photo and video quality evaluation : Focusing on the subject. In Forsyth, D. and Zisserman, A., editors, *ECCV 2008*, volume LNCS 5304, pages 386–399, Berlin-Heidelberg. Springer-Verlag.
- [Ma et al., 2017] Ma, S., Liu, J., and Chen, C. (2017). A-lamp : Adaptive layout-aware multi-patch deep convolutional neural network for photo aesthetic assessment. arXiv :1704.00248.
- [Machado and Cardoso, 1998] Machado, P. and Cardoso, A. (1998). Computing aesthetics. In Springer, editor, *Brazilian symposium on Artificial Intelligence*, pages 219–228, Berlin, Heidelberg.
- [Machajdik and Hanbury, 2010] Machajdik, J. and Hanbury, A. (2010). Affective image classification using features inspired by psychology and art theory. In *MM'10, ACM Conference*, pages 83–92, Firenze (It.).
- [Mai et al., 2011] Mai, L., Le, H., Niu, Y., and Liu, F. (2011). Rule of thirds detection from photograph. In *IEEE Int. Symp on Multimedia (ISM)*, pages 91–96.
- [Mai et al., 2016] Mai, L., Niu, Y., and Liu, F. (2016). Composition preserving deep photo aesthetic assessment. In *Proc. IEEE Conf. Comp. Vision and Pattern Recognition*, pages 497–506.
- [Maitre, 2015] Maitre, H. (2015). *Du photon au pixel : l'appareil photographique numérique*. ISTE Editions, Londres, Royaume Uni.
- [Malinowski-Charles, 2004] Malinowski-Charles, S. (2004). Entre rationalisme et subjectivisme : l'esthétique de Jean-Pierre de Crousaz. *Revue de théologie et de philosophie*, 136 :7–21.
- [Malinowski-Charles, 2005] Malinowski-Charles, S. (2005). Baumgarten et le rôle de l'intuition dans les débuts de l'esthétique. *Etudes philosophiques*, 75(4) :537–558.
- [Marchesotti et al., 2011] Marchesotti, L., Perronnin, F., Larlus, D., and Czurka, G. (2011). Assessing the aesthetic quality of photographs using generic image descriptors. *Int. Conf. on Computer Vision*, pages 1784–1791.
- [Marchesotti et al., 2013] Marchesotti, L., Perronnin, F., and Meylan, F. (2013). Learning beautiful (and ugly) attributes. In *BMVC*, volume 7, pages 1–11.
- [Markowsky, 1992] Markowsky, G. (1992). Misconceptions about the golden ratio. *The College of Mathematics*, 23(1) :2–19.
- [Marquet, 1983] Marquet, J. (1983). *Contribution à l'histoire de la philosophie moderne*. Presses Universitaires de France, (Paris).
- [Marr, 1982] Marr, D. (1982). *Vision : A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman and Company, 1982.
- [McManus et al., 1985] McManus, I., Edmonson, D., and Rodger, J. (1985). Balance in picture. *British J. of Psychology*, 76 :311–324.
- [McManus et al., 1981] McManus, I., Jones, A., and Cottrell, J. (1981). The aesthetic of colour. *Perception*, 10 :651–666.

- [McManus et al., 2011] McManus, I., Stöver, K., and Kim, D. (2011). Arnheim’s gestalt theory of visual balance : Examining the compositional structure of art photographs and abstract images. *i-Perception*, 2(6) :615–647.
- [Mehrabian, 1977] Mehrabian, A. (1977). Individual differences in stimulus screening and arousability. *J. Personality and Social Psychology*, 45 :237–250.
- [Merleau-Ponty, 1945] Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. NRF, Gallimard, Paris.
- [Merleau-Ponty, 1964] Merleau-Ponty, M. (1964). *L’œil et l’esprit*. Folio, Gallimard (Paris).
- [Messina et al., 2017] Messina, P., Dominguez, V., Parra, D., Trattner, C., and Soto, A. (2017). Exploring content-based artwork recommendation with metadata and visual features. arXiv preprint arXiv :1706.05786.
- [Messina et al., 2018] Messina, P., Dominguez, V., Parra, D., Trattner, C., and Soto, A. (2018). Content-based artwork recommendation : integrating painting metadata with neural and manually-engineered visual features. *User Modeling and User-Adapted Interaction*, Springer, pages 1–40.
- [Mittal et al., 2013] Mittal, A., Soundararajan, R., and Bovik, A. C. (2013). Making a ”complete blind” image quality analyzer. *IEEE Signal Processing Letters*, 20 (3) :209–212.
- [Moles, 1957] Moles, A. A. (1957). Théorie de l’information et perception esthétique. *Revue Philosophique de la France et de l’Etranger et Flammarion*, pages 233–242.
- [Molnar, 1974] Molnar, F. (1974). Experimental aesthetics or the science of art. *Leonardo*, 7(1) :23–26.
- [Molnar, 1977] Molnar, F. (1977). Mouvement des yeux et l’hypothèse des explorations épistémiques et diversives. *Journal National de la Société Française de Psychologie*.
- [Molnar, 1981] Molnar, F. (1981). *Towards Science in Art*, pages 385–413. Advances in intrinsic motivation and aesthetics. Hill, A., ed., Plenum Press, New-York.
- [Molnar, 1997] Molnar, F. (1997). A science of vision for visual art. *Leonardo*, 30(3) :225–232.
- [Moon and Spencer, 1944a] Moon, P. and Spencer, D. (1944a). Aesthetic measure applied to color harmony. *Journal Optical Society of America*, 34 :234–242.
- [Moon and Spencer, 1944b] Moon, P. and Spencer, D. (1944b). Area in color harmony. *Journal Optical Society of America*, 34 :93–103.
- [Moon and Spencer, 1944c] Moon, P. and Spencer, D. (1944c). Geometric formulation of classical color harmony. *Journal of the Optical Society of America*, 34(1) :46–63.
- [Moorthy and Bovik, 2011] Moorthy, A. K. and Bovik, A. C. (2011). Blind image quality assessment : From natural scene statistics to perceptual quality. *IEEE transactions on Image Processing*, 20 (12) :3350–3364.
- [Moshagen and Thielsch, 2010] Moshagen, M. and Thielsch, M. T. (2010). Facets of visual aesthetics. *International journal of human-computer studies*, 66(10) :689–709.
- [Motoyoshi et al., 2007] Motoyoshi, I., Ishida, S., Sharan, L., and Adelson, E. (2007). Image statistics and the perception of surface qualities. *Nature*, 447 :206–209.
- [Murray and Gordo, 2017] Murray, N. and Gordo, A. (2017). A deep architecture for unified aesthetic prediction. arXiv preprint arXiv :1708.04890.
- [Murray et al., 2012] Murray, N., Marchesotti, L., and Perronnin, F. (2012). AVA : a large -scale database for for aesthetic visual analysis. In *CVPR : Intern. Conf. on Comp. Vision and Pattern Recognition*, pages 2408–2415.
- [Murray et al., 2013] Murray, N., Marchesotti, L., and Perronnin, F. (2013). Methods and systems for ranking images using semantic and aesthetic models (us patent -20140351264).
- [Neveux, 1995] Neveux, M. (1995). *Le nombre d’or, radiographie d’un mythe*. Editions du Seuil, Paris.
- [Nietzsche, 2005] Nietzsche, F. (2005). *Crépuscule des idoles, (traduction par Patrick Wotlin)*. Garnier-Flammarion.

- [Nodine et al., 1993] Nodine, C., Locher, P., and Krupinski, E. (1993). The role of formal art training on perception and aesthetic judgement of art compositions. *Leonardo*, 26(3) :219–227.
- [Obrador et al., 2010] Obrador, P., Schmitt-Hackenberg, L., and Oliver, N. (2010). The role of image composition in image aesthetics. In *Proc of the 17th IEEE Int. Conf. on Image Processing*, pages 3185–3188, Hong-Kong.
- [O’Connor, 2010] O’Connor, Z. (2010). Colour harmony revisited. *Color Research & Applications*, 35(4) :267–273.
- [Ou et al., 2004] Ou, L., Luo, M., Woodcock, A., and Wright, A. (2004). A study of colour emotion and colour preference : Part I. colour emotions of single colours. *Color research and applications*, 29(3) :232–240.
- [Palmer, 2010] Palmer, S. E. and Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proceedings of the National Academy of Sciences, USA*, 107(19) :8877–8882.
- [Panofsky, 1983] Panofsky, E. (1983). *Meaning in the visual arts*. University of Chicago Press.
- [Park et al., 2017] Park, K., Hong, S., Baek, M., and Han, B. (2017). Personalized image aesthetic quality assessment by joint regression and ranking. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1206–1214.
- [Pascual-Marqui, 1999] Pascual-Marqui, R. D. (1999). Review of methods for solving the EEG inverse problem. *International journal of bioelectromagnetism*, 1(1) :75–86.
- [Pearce et al., 2016] Pearce, M. T., Zaidel, D. W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., and Nadal, M. (2016). Neuroaesthetics : the cognitive neuroscience of aesthetic experience. *Perspectives on Psychological Science*, 11(2) :265–279.
- [Pennington et al., 2014] Pennington, J., Socher, R., and Manning, C. (2014). Glove : Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.
- [Pepperell, 2005] Pepperell, R. (2005). The Posthuman Manifesto. *Kritikos*, 2 :1552–5112. ISSN 1552-5112.
- [Petitot, 2008] Petitot, J. (2008). *Neurogéométrie de la vision*. Les Editions de l’Ecole Polytechnique.
- [Platon, 2011] Platon (2011). *Platon, œuvres complètes*. Flammarion.
- [Pouivet, 2006] Pouivet, R. (2006). *Le réalisme esthétique*. Presses universitaires de France, Paris.
- [Ramachandran and Freeman, 2001] Ramachandran, V. S. and Freeman, A. (2001). Sharpening up ‘the science of art’. *Journal of Consciousness Studies*, 8(1) :9–30.
- [Reagle, 2013] Reagle, J. (2013). Revenge rating and tweak critique at photo.net. personal website : <http://reagle.org/joseph/2013/photo.net.html>.
- [Reber et al., 2004] Reber, R., Schwarz, N., and Winkielman, P. (2004). Processing fluency and aesthetic pleasure : Is beauty in the perceiver’s processing experience? *Personality and Social Psychology Review*, 8(4) :364–382.
- [Redi et al., 2017] Redi, M., Liu, F., and O’Hare, N. (2017). Bidding the aesthetics gap : The wild beauty of web imaginary. In ACM, editor, *ICMR’17 conference*, pages 242–250, Bucarest.
- [Redies, 2015] Redies, C. (2015). Combining universal beauty and cultural context in a unifying model of visual aesthetic experience. *Frontiers in human neuroscience*, 9 :218.
- [Renault, 2004] Renault, B. (2004). *Imagerie cérébrale fonctionnelle électrique et magnétique*. Hermes-Lavoisier, Paris.
- [Rigau et al., 2008] Rigau, J., Feixas, M., and Sbert, M. (2008). Information aesthetic measures. *IEEE Computer Graphics and Applications*, 2 :24–34.
- [Rosenberg and Klein, 2015] Rosenberg, R. and Klein, C. (2015). *Art, Aesthetics and the brain*, chapter 5 - The moving eye of the beholder : Eye tracking and the perception of paintings, pages 79–108. Oxford University Press.
- [Rozenkranz, 1853] Rozenkranz, K. (1853). *Ästhetik des Häßlichen*. Heidelberg.

- [Russell, 1980] Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, 39(6) :1161–1178.
- [Saint Girons, 2005] Saint Girons, B. (2005). *Le sublime de l'antiquité à nos jours*. Desjonquères (Paris).
- [San Pedro and Siersdorfer, 2009] San Pedro, J. and Siersdorfer, S. (2009). Ranking and classifying attractiveness of photos in folksonomies. In *Proc. ACM conf on World Wide Web WWW'09*, pages 771–780, New-York.
- [San Pedro et al., 2012] San Pedro, J., Yeh, T., and Oliver, N. (2012). Leveraging user comments for aesthetic aware image search reranking. In ACM, editor, *Proceedings of the 21st international conference on World Wide Web*, pages 439–448.
- [Schaper, 1964] Schaper, E. (1964). Kant on aesthetics appraisals. *Kant Studien*, 1-4 :431–449.
- [Schendan and Kutas, 2007] Schendan, H. E. and Kutas, M. (2007). Neurophysiological evidence for transfer appropriate processing of memory : processing versus feature similarity. *Psychonomic bulletin & review*, 14(4) :612–619.
- [Schifanella et al., 2015] Schifanella, R., Reddi, M., and Aiello, L. (2015). An image is worth more than thousand favourites : surfacing the hidden beauty of Flickr pictures. ArXiv : 1505.03358v2.
- [Schloss and Palmer, 2011] Schloss, K. B. and Palmer, S. E. (2011). Aesthetic response to colour combinations : preference, harmony and similarity. *Attention Perceptual Psychophysics*, 73 :551–571.
- [Schmidhuber, 1997] Schmidhuber, J. (1997). Low-complexity art. *Leonardo, Journal of the International Society for the Arts, Sciences, and Technology*, 30(2) :97–103.
- [Schmidhuber, 2007] Schmidhuber, J. (2007). Simple algorithmic principles of discovery, subjective beauty, selective attention, curiosity and creativity. arXiv : 0709.0674.
- [Schopenhauer, 1966] Schopenhauer, A. (1966). *Le monde comme volonté et comme représentation*. PUF (Paris).
- [Schupp et al., 2004] Schupp, H., Cuthbert, B., Bradley, M., Hillman, C., Hamm, A., and Lang, P. (2004). Brain processes in emotional perception : Motivated attention. *Cognition and emotion*, 18(5) :593–611.
- [Schuster et al., 2010] Schuster, R., Mörzinger, R., Haas, W., Grabner, H., and Van Gool, L. (2010). Real-time detection of unusual regions in image streams. In ACM, editor, *Proceedings of the 18th ACM international conference on Multimedia*, pages 1307–1310.
- [Schwarz et al., 2016] Schwarz, K., Wieschollek, F., and Lensch, H. (2016). Will people like your image? arXiv 1611-05203, University of Tübingen.
- [Schweinhart and Essock, 2013] Schweinhart, A. and Essock, E. (2013). Structural content in paintings : Artists overregularize oriented content of paintings relative to the typical natural scene bias. *Perception*, 42 :1311–1332.
- [Sibley, 1959] Sibley, F. (1959). Aesthetic concepts. *Philosophical review*, LXVIII :421–450.
- [Simond et al., 2015] Simond, F., Arvanitopoulos, N., and Süsstrunk, S. (2015). Image aesthetics depends on context. In *IEEE International Conference on Image Processing*, pages 3788–3792. IEEE.
- [Skov, 2009] Skov, M. (2009). Neuroaesthetic problems : a framework for neuroaesthetic research. In Baywood, editor, *Neuroaesthetic*, pages 9–36, Amityville (N.Y).
- [Smith, 2008] Smith, D. (2008). Color-person-environment relationship. *Color Research & Applications*, 33 :312–319.
- [Solso, 1996] Solso, R. (1996). *Cognition and the visual Arts*. The MIT Press, Cambridge (Mass.), 5th edition.
- [Sperber and Wilson, 2004] Sperber, D. and Wilson, D. (2004). Relevance theory. In Horn, L.R. & Ward, G., editor, *The Handbook of Pragmatics*, pages 607–632. Oxford : Blackwell.

- [Srivastava and Kant, 2018] Srivastava, M. M. and Kant, S. (2018). Visual aesthetic analysis using deep neural network : model and techniques to increase accuracy without transfer learning. arXiv preprint arXiv :1712.03382.
- [Su et al., 2012] Su, H. H., Chen, T. W., Kao, C. C., Hsu, W. H., and Chien, S. Y. (2012). Preference-aware view recommendation system for scenic photos based on bag-of-aesthetics-preserving features. *IEEE Transactions on Multimedia*, 14(3) :833–843.
- [Sun et al., 2017] Sun, W. T., Chao, T. H., Kuo, Y. H., and Hsu, W. H. (2017). Photo filter recommendation by category-aware aesthetic learning. *IEEE Transactions on Multimedia*, 19(8) :1870–1880.
- [Swami, 2013] Swami, V. (2013). Context matters : Investigating the impact of contextual information on aesthetic appreciation of paintings by Max Ernst and Pablo Picasso. *Psychology of Aesthetics, Creativity, and the Arts*, 7(3) :285–296.
- [Szegedy et al., 2015] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2015). Rethinking the inception architecture for computer vision. arXiv :1512.00567.
- [Séailles, 1877] Séailles, G. (1877). L'esthétique de Hartmann. *Revue philosophique de la France et de l'étranger*, IV, p. , nov, 1877, IV :483–495.
- [Talairach and Tournoux, 1988] Talairach, J. and Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain :3-dimensional proportional system-An approach to cerebral imaging*. Thieme Medical Publishers, New-York.
- [Talebi and Milanfar, 2017] Talebi, H. and Milanfar, P. (2017). NIMA : Neural image assessment. arXiv : 1709.0541v1.
- [Tatarkiewicz, 1970] Tatarkiewicz, W. (1970). *History of aesthetics*. Mouton ed.
- [Taylor et al., 2013] Taylor, C., Clifford, A., and Franklin, A. (2013). Color preferences are not universal. *Journal of experimental psychology*, 142(4) :1015–1027.
- [Thorpe et al., 1996] Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of processing in the human visual system. *Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. nature.*, 381(6582) :520–521.
- [Thumfart et al., 2011] Thumfart, S., Jacobs, R. H., Lughofer, E., Eitzinger, C., Cornelissen, F. W., Groissboeck, W., and Richter, R. (2011). Modeling human aesthetic perception of visual textures. *ACM Transactions on Applied Perception (TAP)*, 8(4) :27.
- [Ticini et al., 2014] Ticini, L. F., Rachman, L., Pelletier, J., and Dubal, S. (2014). Enhancing aesthetic appreciation by priming canvases with actions that match the artist's painting style. *Frontiers in human neuroscience*, 8 :391.
- [Trujillo et al., 2014] Trujillo, L. T., Jankowitsch, J. M., and Langlois, J. H. (2014). Beauty is in the ease of the beholding : A neurophysiological test of the averageness theory of facial attractiveness. *Cognitive, Affective, & Behavioral Neuroscience*, 14(3) :1061–1076.
- [Tversky and Kahneman, 1981] Tversky, A. and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211 :451–458.
- [Ulrich, 1983] Ulrich, R. S. (1983). *Behavior and the natural environment*, chapter 3 - Aesthetic and affective response to natural environment, pages 85–125. Springer, Boston, MA.
- [Uttal, 2002] Uttal, W. (2002). Precis of the new phrenology : the limits of localizing processes in the brain. *Brain and Mind*, 3 :221–228.
- [Van Dongen et al., 2016] Van Dongen, N. N., Van Strien, J. W., and Dijkstra, K. (2016). Implicit emotion regulation in the context of viewing artworks : ERP evidence in response to pleasant and unpleasant pictures. *Brain and Cognition*, 107 :48–54.
- [van Gogh, 1992] van Gogh, V. (1992). *Dernières lettres*. Mille et une nuits, Paris, n 212.
- [Vartanian and Skov, 2014] Vartanian, O. and Skov, M. (2014). Neural correlates of viewing paintings : evidence from a quantitative meta analysis of fmrimaging data. *Brain Cogn.*, 87 :52–56.
- [Verdier, 2001] Verdier, F. (2001). *L'unique trait de pinceau - Calligraphie, peinture et pensée chinoise*. Albin Michel, Paris.

- [Verdier, 2003] Verdier, F. (2003). *Passagère du silence*. Albin Michel (Paris).
- [Veryzer Jr and Hutchinson, 1998] Veryzer Jr, R. W. and Hutchinson, J. W. (1998). The influence of unity and prototypicality on aesthetic responses to new product designs. *Journal of consumer research*, 24(4) :374–394.
- [Vidal, 2011] Vidal, F. (2011). La neuroesthétique, esthétique scientifique. *Revue d'histoire des sciences humaines*, 25(2) :239–264.
- [Vidal, 2012] Vidal, F. (2012). Neuroaesthetics : Getting rid of art and beauty. *Biosocieties*, 7(2) :209.
- [Vienot and Le Rohellec, 2012] Vienot, F. and Le Rohellec, J. (2012). Colorimétrie et physiologie : la spécification LMS. In Fernandez-Maloigne, C., Robert-Inacio, F., and Macaire, L., editors, *Couleur numérique : acquisition, perception, codage et rendu*, chapter 1, pages 19–40. Hermès-Lavoisier, Paris.
- [Vinyals et al., 2015] Vinyals, O., Toshev, A., Bengio, S., and Erhan, D. (2015). Show and tell : A neural image caption generator. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3156–3164.
- [Wallis, 2007] Wallis, J. (2007). Orbitofrontal cortex and its contribution to decision making. *Annu. Rev. Neurosciences*, 30 :31–56.
- [Wang et al., 2014] Wang, W., Liu, J., Zhao, W., and Li, J. (2014). A system of image aesthetic classification and evaluation using cloud computing. In Li, S., Liu, C., and Wang, Y., editors, *Chinese Conference on Pattern Recognition. CCPR 2014.*, volume 443, Communications in Computer and Information Science, pages 183–195. Springer, Berlin, Heidelberg.
- [Wang et al., 2018] Wang, W., Yang, S., Zhang, W., and Zhang, J. (2018). Neural aesthetic image reviewer. arXiv preprint arXiv :1802.10240.
- [Wang et al., 2016] Wang, Z., Chang, S., Dolcos, F., Beck, D., and Huang, T. (2016). Brain-inspired deep network for image aesthetic assesment. arXiv : 1601.04155v2.
- [Wascheck, 2000] Wascheck, M. (2000). *Le chef d'œuvre : un fait culturel*, chapter 2, pages 25–46. Qu'est-ce qu'un Chef d'œuvre? Gallimard (Paris).
- [Wertheimer, 1922] Wertheimer, M. (1922). Untersuchungen zur Lehre von der Gestalt - i - prinzipelle Bemerkungen. In *Psychologische Forschung*, volume 1, pages 47–58.
- [Wertheimer, 1923] Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt - ii - laws of organization in perceptual forms. In *Psychologische Forschung*, volume 4, pages 301–350.
- [Wertheimer, 1938] Wertheimer, M. (1938). *Laws of of organization in perceptual forms*. Ellis, W.D., A source book of Gestalt psychology, London.
- [Whitfield and Slatter, 1979] Whitfield, T. W. and Slatter, P. E. (1979). The effects of categorization and prototypicality on aesthetic choice in a furniture selection task. *British Journal of Psychology*, 70(1) :65–75.
- [Wilber et al., 2017] Wilber, M. J., Fang, C., Jin, H., Hertzmann, A., Collomosse, J., and Belongie, S. (2017). Bam! the Behance artistic media dataset for recognition beyond photography. In *Proc. Int. Conf. Comp. Vision*, volume 1, page 4.
- [Xun, 2015] Xun, J. (2015). *Contemplation sur l'art*. SDX Joint publishing company, (Taiwan).
- [Xun, 2016] Xun, J. (2016). *Aube réelle*. People's Literature Publishing Society.
- [Yao et al., 2012] Yao, L., Suryanarayan, P., Qiao, M., Wang, J. Z., and Li, J. (2012). Oscar : On-site composition and aesthetics feedback through exemplars for photographers. *International Journal of Computer Vision*, 96(3) :353–383.
- [Yarbus, 1967] Yarbus, A. (1967). *Eye movements and vision*. Plenum Press.
- [Zeki, 1999] Zeki, S. (1999). *Inner vision : An exploration of art and the brain*. Oxford University Press.
- [Zemach, 1991] Zemach, E. (1991). Real beauty. *Midwest studies in Philosophy*, XVI :249–265.
- [Zemach, 1987] Zemach, E. M. (1987). Aesthetic properties, aesthetic laws and aesthetic principles. *The journal of aesthetics and art criticism*, 46(1) :239 :251.

- [Zemach, 2005] Zemach, E. M. (2005). *La beauté réelle : une défense du réalisme esthétique*. Presses Universitaires de Rennes ; collection "Aesthetica".
- [Zhang et al., 2016] Zhang, J., Sclaroff, S., Lin, Z., Shen, X., Price, B., and Mech, R. (2016). Unconstrained salient object detection via proposal subset optimization. In *IEEE Conf. on Comp. Vision and Pattern Recognition (CVPR)*, pages 5733–5742.

Index

- A-Lamp, méthode d'évaluation, 111
- AADB, (base d'images), 85
- A&C, méthode d'évaluation, 109
- accutance, 59
- Acquine, méthode d'évaluation, 91
- affect circomplexe, 123
- affordance, 47
- aires visuelles, 38
- Alberti, L.B., 137
- Alberti,L.B., 19
- AlexNet, architecture de RNP, 101
- Alhazem, 136
- amazement*, 11
- appraisal*, 28
- apprentissage numérique, 80
- approche
 - objectiviste, 19
 - subjectiviste, 25
- Arnheim, R., 54, 57
- arousal*, 12
- arousalarousal*, 27
- art, 13
- art brut, 143
- artisanales (primitives), 89
- attrait, 83
- Augustin d'Hippone, 18, 19, 135
- AVA, (base d'images), 85
- AVA-Comments (base de données), 114
- AVA2, (base d'images), 85

- bag of visual words* (BOVW), 97
- BAM (base de données d'œuvres d'art), 86
- bases de données, 83
 - Photo.Net, 83
 - AADB, 85
 - AVA, 85
 - AVA2, 85
 - BAM, 86
 - BEAUTY, 85
 - CHUK-BD, 84
 - DPChallenge, 84
 - FACT, 87
 - Flickr, 83
 - ImageCLEF, 84
 - ImageNet, 83
 - JenAesthetics, 86
 - SUN, 84
 - Uni Tübingen, 86
- Baumgarten, A., 141
- beauté, 13
- BEAUTY, (base d'images), 85
- Bell, C., 24
- Benjamin, W., 22
- Bense, M., 75
- Berlyne, D. E., 27
- big data*, 80
- Birkhoff, D., 73
- Bourdieu, P., 14
- brain-inspired* RNP, méthode d'évaluation, 110
- Brown, S., 36, 47
- Brown, S. *et al.*, (modèle d'émotion esthétique), 43
- Burke, E., 13

- cadrage, 58
- CAFFE, environnement de RNP, 105
- canon (esthétique), 133
- centre hédonique, 35
- Changeux, J.-P., 28, 48
- Chatterjee, A., (modèle d'émotion esthétique), 42
- Chevreul, E., 64, 73
- Chine, (l'esthétique en), 145
- CHUK-DB, (base d'images), 84
- CIELab (métrique), 65
- classification, 89
- cognition, (aires de la), 38
- composition, 55, 57
- conditions d'observation standart, 24
- cortex
 - insulaire, 36
 - préfrontal, 35
- couleur, 61, 143
- Cuda-Convnet, 111
- Cupchnik, G., 38

- d'Ockam, G., 137
- Démocrite, 134
- Damasio, A., 46
- Danto, A.C., 29
- de Crousaz, J.-P., 139
- design, 29
- Dessalles, J.L. , 10

- Di Dio, C., 28, 38, 39
 diagonales, 58
 Diderot, D., 18, 26
 DMA-Net, méthode d'évaluation, 105
 DPChallenge, (base d'images), 84
 Du Bos, J.-B., 139
 Dubuffet, J., 143
- échantillonnage de Shannon, 59
 EEG, 40
embodiment, 39
 ERP, (*event-related potential*), 40
 esthétique, 13
 analytique, 52
 contextualisée ou historique, 29
 fonctionnelle, 29
event-related potential, (ERP), 40
 excitation, 12
 experts, 82
- FACD, (base de données), 87
 Fechner, G.T., 9, 15, 54
 Ferry, L., 18
 flou, 59
 format d'aspect, 101
 format d'image, 54
 forme signifiante, 24
 fuyantes, 58
- Gärdenfors, U., 9
 GAN, *generative adversarial network*, 107
 Gestalt, (théorie de la), 74
 Goethe, J.W., 26, 64, 143
 Gombrich, A.E., 29
 GoogLeNet, réseau de neurones, 105
- hand-crafted*, (*features*), 89
 hédonique (centre), 35
 Hegel, G.W.F., 23, 25, 141
 Heidegger, M., 28
 Henry, C., 73
high-keys, 60
 histogramme, 60
 Hogarth, W., 139
 Hume, D., 139
 Hutcheson, F., 28, 139
- IAPS, *International affective picture system*, 12, 41
 icône, 135
 ImageCLEF, (base d'images), 84
 incarnation (*embodiment*), 39
 Inception, réseau de neurones, 114
interestingness, 10
 IRMF, imagerie par résonance magnétique fonctionnelle, 34
- Ishizu, T., 28, 35
 Itten, J., 64
- James, W., 10
 JenAesthetics (base de peintures), 86
 Jullien, F., 145
- Kant, E., 13, 21, 25, 141
 Koelsch, S. *et al.*, (modèle d'émotion esthétique), 45
- l'art pour l'art, 22
 laid, 36
 Leder, H., (modèle d'émotion esthétique), 44
 Leonardo da Vinci, 138
 Levinson, J., 24
 Li-Hsiang Hsu, (modèle d'émotion esthétique), 46
 Locke, J., 26
low-keys, 60
- machine learning*, 80
 masses de données, 80
 Matsuda, Y., 67
 MEG, 40
 mémoire, (aires de la), 38
memorability, 10
 méthode d'évaluation
 A-Lamp, 111
 A&C, 109
 Acquine, 91
 brain-inspired RNP, 110
 DMA-Net, 105
 MNA, 105
 MTRL-CNN, 108
 NAIR, 114
 NIMA, 113
 OSCAR, 95
 RAPID, 103
 MNA, méthode d'évaluation, 105
 Mobile-Net, réseau de neurones, 114
 modèle d'émotion esthétique
 de A. Chatterjee, 42
 de C. Redies, 45
 de H. Leder, 44
 de Li-Hsiang Hsu, 46
 de S. Brown *et al.*, 43
 de S. Koelsch *et al.*, 45
 modèle neuro-géométrique de J. Petitot, 42
 Moles, A., 75
 Molnar, F., 27, 55
 montagne-rivière (*shan shui*), 150
 Moon, P. et Spencer, D.E., 65
 MTRL-CNN, méthode d'évaluation, 108
 Munsell, A., 65
 mutualisation (*pooling*), 101

- NAIR, méthode d'évaluation, 114
 neuro-
 biologie, 33
 esthétique, 28
 géométrie, 42
 neurones miroirs, 39
 Nietzsche, F., 27, 144
 NIMA, méthode d'évaluation, 113
 nombre d'or, 54

 objectiviste, (approche), 19
 OSCAR, méthode d'évaluation, 95

 palette chromatique, 64
 Panofsky, E., 134
 peinture chinoise, 145
 perceptibilité, 134
 personnification (*embodiment*), 39
 pertinence, 11
 Petitot, J., 42
 phénoménologie, 23
 Photo.net, (base d'images), 83
 piqué, 59
 Platon, 18, 19
popularity, 10
 Pouivet, R., 24
 prédicats esthétiques, 24
Prägnanz, 27
 primitives, 89
 artisanales, 89
 de haut niveau, 93
 psychologie expérimentale, 51

qualia, 30

 réalisme esthétique, 24
 Rand, A., 19
 RAPID, méthode d'évaluation, 103
 rapport d'aspect, 54
 recommandation (système de), 128
 Redies, C. (modèle d'émotion esthétique), 45
 règle du tiers, 55
 réseau de neurones
 adverse, 107
 Inception, 114
 Mobile-Net, 114
 profonds (RNP), 80, 101
 ResNet, 109
 siamois, 107
 VGG, 103, 105, 114
 ResNet, réseau de neurones, 109
 résonance magnétique fonctionnelle (IRMf), 34
reward, 35
 RNP : réseau de neurones profonds, 80, 101

 Saint-Augustin (Augustin d'Hippone), 18

 Schelling, F., 143
 Schmidhuber, J., 76
 Schopenhauer, A., 19, 23, 143
 sens commun, 141
shan shui (montagne-rivière = paysage), 150
 Shannon, (échantillonnage, théorème), 59
 Sheikh (ou Xie He), 150
 Shitao, 147
 Six lois de
 Vitruve, 133, 150
 Xie He, 150
 SOC, *standart observation conditions*, 24
 Socrate, 18, 19
 Solso, R.L., 27
 SPARK, *photographic contest*, 15, 120
 spectre de densité de puissance, 58
 Sperber, D. et Wilson, D., 10
 subjectiviste, (approche), 25
 sublime, 13, 143
 SUN, (base d'images), 84
 survéance, 24
 symétrie, 20, 134

 Tatarkiewicz, W., 133
 TensorFlow, bibliothèque logicielle, 114
 texture, 59
 théorème de Shannon, 59
 Tversky, A. et Kahneman, D., 9

unusualness, 10

 Vartanian, O., 39
 VGG, réseau de neurones, 103, 105, 114
 VISTA+, (système de recommandation), 129
 Vitruve, 133, 150

 Wilson, D. et Sperber, D., 10

 Xie He (ou Sheikh), 150

 Yarbus, A., 55

 Zeki, S., 28, 35
 Zemach, E., 24