Artificial Intelligence and Pattern Recognition, Vision, Learning

Isabelle Bloch, Régis Clouard, Marinette Revenu and Olivier Sigaud

Abstract This chapter describes a few problems and methods combining artificial intelligence, pattern recognition, computer vision and learning. The intersection between these domains is growing and gaining importance, as illustrated in this chapter with a few examples. The first one deals with knowledge guided image understanding and structural recognition of shapes and objects in images. The second example deals with code supervision, which allows designing specific applications by exploiting existing algorithms in image processing, focusing on the formulation of processing objectives. Finally, the third example shows how different theoretical frameworks and methods for learning can be associated with the constraints inherent to the domain of robotics.

1 Introduction

The intersection between the domains of artificial intelligence (AI), and of pattern recognition, computer vision and robotics is getting more and more important and visible. The overlap between these domains was significantly enlarged during the last years. The objective of this chapter is to show a few aspects of this overlap, in particular for high level visual scene understanding and for integrating knowledge in processing and interpretation methods.

I. Bloch (✉)
LTCI, Télécom ParisTech, Université Paris Saclay, Paris, France
e-mail: isabelle.bloch@telecom-paristech.fr

R. Clouard · M. Revenu
GREYC, Normandie Univ., ENSICAEN, CNRS, Caen, France
e-mail: regis.clouard@ensicaen.fr

M. Revenu
e-mail: marinette.revenu@ensicaen.fr

O. Sigaud
ISIR, CNRS, Sorbonne Université, Paris, France
e-mail: sigaud@isir.upmc.fr

P. Marquis et al. (eds.), *A Guided Tour of Artificial Intelligence Research*,
https://doi.org/10.1007/978-3-030-06170-8_10

Several topics addressed in other chapters and several of the therein described methods can also be associated with problems in pattern recognition, artificial vision or image understanding, and robotics. For instance, uncertainty theories are widely used for modelling imperfections of data, of objectives and of reasoning procedures, as for image fusion. Learning is at the core of many recent developments, such as for image mining or for robotics. Multi-agents systems have been exploited for developing cooperation between methods in image processing, as well as for developing interactions between or with robots. Finally, as a last example, structural representations (graphs, hypergraphs, Bayesian networks, ontologies, knowledge based systems…) are naturally used for modelling and interpreting image or video content. They allow associating low level information with higher level one and with knowledge, to guide the interpretation of the observed scene. This is for instance the case in spatial reasoning (see also chapter "Qualitative Reasoning About Time and Space" of Volume 1).

In this chapter, we describe a few examples of these multiple interactions. In Sect. 2, an overview of interactions between artificial intelligence and computer vision is proposed, in particular for recognizing objects in images, focusing on knowledge based systems. While ontologies are more and more developed to guide scene understanding, by describing and formalizing concepts related the scene contents, they are also exploited to describe the objective of image processing. In this perspective, Sect. 3 presents code supervision methods for automatically generating applications in image processing. Finally, in Sect. 4, the domain of robotics is presented under the light of related learning aspects.

2 AI for Computer Vision and Pattern or Object Recognition

In this section, an overview of interactions between AI and computer vision is proposed, focusing on knowledge based systems for image and visual scene understanding, pattern or shape recognition in images. The general objective of these approaches is to add semantics to the images, by associating visual information and features extracted from the images on the one hand, and knowledge or models on the other hand (Crevier and Lepage 1997; Le Ber and Napoli 2002).

One of the main difficulties, beyond knowledge representation and reasoning issues, is to establish a matching between perceptual and conceptual levels. The perceptual level includes features extracted from images, hence close to pixel (in 2D) or voxel (in 3D) information. The conceptual level is often given in textual form. This problem of matching is known as semantic gap, defined by Smeulders et al. (2000) as: "*the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation*". This problem is close to other problems in AI and robotics, such as symbol grounding or anchoring (Harnad 1990; Coradeschi and Saffiotti 1999).

420055_1_En_10_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:18/5/2019 Pages: 28 Layout: T1-Standard

2.1 Knowledge

The type of knowledge modelled in knowledge based systems is related to the scene and anything that can be useful for its interpretation. According to the classical categorization of Matsuyama and Hwang (1990), the following types are distinguished:

- generic knowledge on the type of scene, describing the objects it contains or may contain, the relationships between these objects, or the type of image;
- specific knowledge about the image, including the observation of the scene and its processing, which is required to extract useful information from images;
- knowledge bridging the semantic gap between a real scene and its observations as images.

2.2 Spatial Relations

Knowledge about space, in particular about spatial relations, is very important for image understanding (Bloch 2005; Kuipers and Levitt 1988). Indeed, human beings use intensively spatial relations for describing, detecting and recognizing objects. They allow solving ambiguities between objects of similar shape or appearance, based on their spatial arrangement, and are often more stable than characteristics of objects themselves. This is for instance the case of anatomical structures, as illustrated later in this chapter.

Spatial reasoning has raised a lot of attention in computer vision and pattern recognition, in artificial intelligence, in cognitive sciences, in mobile robotics, or in geographical information systems. According to the semantic hierarchy proposed by Kuipers and Levitt (1988), important spatial relations can be grouped into topological and metrical relations. Among the metrical relations, directional and distance relations can be distinguished, as well as more complex relations such as "between" or "along".

In the domain of qualitative spatial reasoning, most representation models are symbolic, often relying on logical formalisms, and mostly deal with topological (Vieu 1997) or cardinal (directional) (Ligozat 1998) relations (see chapter "Qualitative Reasoning about Time and Space" of Volume 1). To reason on real data such as images, quantitative or semi-quantitative formalisms are more expressive. For instance, fuzzy models of numerous spatial relations have been proposed (Bloch 2005). They are appropriate to address the issue of the semantic gap, for instance using the concept of linguistic variable, the semantic of each linguistic value being given by a fuzzy set in the concrete domain of the variable. As an example, the fuzzy representation of a concept such as "close to" allows representing the imprecision inherent to this concept, and instantiating its semantics according to the considered application domain (Hudelot et al. 2008). It also allows answering two main questions in structural image understanding:

- to which degree is a spatial relation satisfied between two given objects?
- what is the area of space in which a spatial relation to a reference object is satisfied (up to some degree)?

Among such fuzzy models of spatial relations, those relying on mathematical morphology offer a unified representation framework, able to handle purely quantitative, purely qualitative, as well as semi-quantitative or fuzzy representations (Bloch 2006).

2.3 *Knowledge Representation and Organization*

As in other domains, in vision and pattern recognition one may characterize knowledge representation by:

- the definition of a representation as a set of syntactic and semantic conventions for describing a knowledge element;
- logical representations, with a level of expressivity depending on the logic;
- compact representations, where only relevant properties and characteristics are explicitly represented;
- easy manipulation;
- explicit representation of what is useful for reasoning.

Since most data in the domain of computer vision and pattern recognition are numerical, using logical representations (which are often more compact than numerical ones) requires to convert such data in a symbolic form.

Requirements for symbolic representations are ontological, epistemic and computational. The first two levels impose constraints on the representation language, and the third level on the inference mechanisms.

Recent knowledge based systems can be seen as extensions of classical expert systems, by providing different ways for knowledge representation and reasoning. A few classical examples include:

- production rules, which are easy to adapt or extend, and their results can be explained; however expressivity highly depends on the involved logics;
- frames (Minsky 1974), which are declarative systems well adapted to describe objects classes based on their attributes and properties; hierarchical links allow handling different levels of granularity, with inheritance, specialization or generalization mechanisms; an example in image processing can be found in Clément and Thonnat (1993);
- semantic networks (Quillian 1967), which rely on a graphical representation of a knowledge base, in which vertices represent concepts and objects, and edges represent relations; inference rules exploit inheritance from a class of objects to a more specific class; their representation as attributed relational graphs is often used to model spatial information;

- conceptual graphs (Sowa 1984; Chein and Mugnier 2008), which represent concepts and relations as vertices, linked by edges; again graphical representations are computationally efficient;
- ontologies and description logics, which provide a shared, consistent conceptual formalization of knowledge in a given domain (Gruber 1993) (see also chapter "Reasoning with Ontologies" of Volume 1).

In computer vision and image processing, where the environment is only partially known, early applications of knowledge based systems have been developed for program supervision (Clément and Thonnat 1993; Nazif and Levine 1984) and for image understanding (Desachy 1990; Hanson and Rieseman 1978; Matsuyama 1986; McKeown et al. 1985). Specific problems related to focalization of attention, adaptation of procedures to revise, repair or maintain consistency, cooperation and fusion, coordination could also be added to knowledge based systems (Garbay 2001). A renewed interest led recently to several works in these areas.

For instance, recent works use ontologies to add a semantic level and to solve the semantic gap problem. For instance in Town (2006), the terms of a query language are anchored in the image domain using supervised learning, for application to keyword based image mining. A similar approach was used by Mezaris and Kompatsiaris (2004) and Hudelot (2005), who defined an ontology of visual concepts, anchored to descriptors extracted from the images. This type of approach allows both performing queries in a qualitative way based on the ontology concepts, and filtering or selecting relevant results according to their visual features.

Reasoning procedures associated with these different types of representations depend on the involved logic. One of the difficult problems to be solved is the matching between a knowledge model and information extracted from images, because of the semantic gap. This problem is simplified when information is directly linked to object representations (Saathoff and Staab 2008; Benz et al. 2004). Otherwise, for instance when only an over-segmentation of the image is available (i.e. several regions should be merged to be interpreted as an object), methods such as inexact graph matching, constraint satisfaction or spatial reasoning have to be developed (Perchant and Bloch 2002; Bengoetxea et al. 2002; Deruyver and Hodé 1997, 2009; Colliot et al. 2006; Fouquier et al. 2012; Nempont et al. 2013; Atif et al. 2013).

2.4 Uncertainty

In image understanding and computer vision, one has to deal with imperfect information. These imperfections are of different natures, and include ambiguity, bias, noise, incompleteness, imprecision, uncertainty, inconsistency, conflict…Additionally, when dealing with dynamic scenes, the information can be variable and evolves during time. These imperfections, found similarly in different problems in general information processing (Dubois and Prade 2001), may be due to the observed phenomenon itself, limitations of sensors, image reconstruction and processing methods and algorithms, noise, lack of fiability, representation models, knowledge and concepts that are handled.

It is of high importance to account for these imperfections in representation models and in reasoning methods.

The main numerical models used in image processing and understanding to model uncertainty rely on probability theory and statistics, belief functions, fuzzy sets and possibility theory. They were developed in particular in the domain of information fusion (Bloch 2008), where the combination of several sources of information aims at making better decision while coping with imperfections of information, but also to represent structural information such as spatial relations (Bloch 2005).

In probabilistic representations, the language is constituted by probability distributions on a given reference domain. They account rigorously for random and stochastic uncertainty, but not easily for other types of imperfections, from both semantic and formal point of view. Bayesian inference is often used in this framework.

Belief functions (or Dempster–Shafer theory (Shafer 1976)) rely on a language defining several functions (belief function, plausibility…) on the power set of the decision space. Such representations cope with both imprecision and uncertainty (including of subjective nature), with ignorance and incompleteness, and allow computing a degree of conflict between data or information sources. The well known Dempster orthogonal rule performs a conjunctive combination, while other rules propose different types of behaviour in the combination (Denœux 2008).

In fuzzy sets and possibility theory (Dubois and Prade 1980, 1988; Zadeh 1965), the language includes fuzzy sets defined on a domain, or possibility distributions. Qualitative, imprecise and vague information can be suitably represented. Inference relies on logical rules, and qualitative reasoning is available. The usefulness of fuzzy sets for information processing in image and vision can be found at several levels (Bloch 2003, 2006):

- the ability of fuzzy sets to represent spatial information in images along with its imprecision, at different levels (local, regional, global), and under different forms (ranging from purely quantitative to purely qualitative) and different levels of granularity;
- the possibility to represent heterogeneous information, either extracted from the images or derived from external knowledge (such as expert or generic knowledge about a domain or an applicative problem);
- the possibility to generalize to fuzzy sets many operations to handle spatial information;
- the flexibility of combination operators, useful to combine information of different natures in various situations.

More details about uncertainty representations can be found in chapters "Representations of Uncertainty in Artificial Intelligence: Probability and Possibility" and "Representations of Uncertainty in Artificial Intelligence: Beyond Probability and Possibility" of Volume 1.

These models have been integrated in the knowledge representation methods described above, including ontologies (Hudelot et al. 2008, 2010), for successful applications in image understanding.

2.5 Example: Recognition of Brain Structures in 3D MRI

The automatic interpretation of complex scenes such as the brain requires a model representing knowledge on the structures present in the scene. In the easiest situations, each object has a different appearance, and prior knowledge on it may be sufficient to detect and recognize the objects. However, this is not the case in magnetic resonance images (MRI) of the brain, since the appearance is not discriminative enough. Other properties such as the spatial arrangement of the structures is then very important and helpful.[1]

Brain anatomy is commonly described in a hierarchical fashion and can be formalized using ontologies, such as the Foundational Model of Anatomy (FMA) (Rosse and Mejino 2003). In addition, the spatial organization of the anatomical structures is a major component of linguistic descriptions of the brain anatomy (Hasboun 2005; Waxman 2000). The overall structure of the brain is quite stable, while the shapes and sizes of the individual structures are prone to substantial variability, and therefore it is relevant to include spatial relations in a model of the brain anatomy. This allows coping with anatomical variability and offering good generalization properties.

Graphs are often used to represent the structural information in image interpretation, where the vertices represent objects or image regions (and they may carry attributes such as their shapes, sizes, and colours or grey levels), and the edges carry the structural information, such as the spatial relations among objects, or radiometric contrasts between regions. Although this type of representation has become popular in the last 30 years (Conte et al. 2004), a number of open problems remain in its efficient implementation. In one type of approach, the graph is derived from the image itself, based on a preliminary segmentation into homogeneous regions, and the recognition problem is expressed as a graph matching problem between the image and model graphs, which is an annotation problem. However this scheme often requires solving complex combinatorial problems (Conte et al. 2004). These approaches assume a correct initial segmentation of the image. However, the segmentation problem is a known challenge in image processing, to which no universal solution exists. The segmentation is usually imperfect, and no isomorphism exists between the graphs being matched. An inexact matching must then be found, for instance by allowing several image regions to be assigned to one model vertex or by relaxing the notion of morphism to that of fuzzy morphism (Perchant and Bloch 2002; Cesar et al. 2005). For example, previous studies (Deruyver and Hodé 1997, 2009) employ an over-segmentation of the image, which is easier to obtain. A model

[1]This section is to a large part adapted from Nempont et al. (2013).

structure (i.e. a graph vertex) is then explicitly associated with a set of regions, and the recognition problem is expressed as a constraint satisfaction problem.

To deal with the difficulty of obtaining a relevant segmentation, the segmentation and recognition can also be performed simultaneously. For instance, in Bloch et al. (2003), Colliot et al. (2006), the structures of interest are segmented and recognized sequentially, in a pre-calculated order (Fouquier et al. 2008, 2012). The structures that are easier to segment are considered first and adopted as reference objects. The spatial relations to these structures are encoded in the structural model and are used as constraints to guide the segmentation and recognition of other structures. This approach benefits from an ontological representation of anatomical knowledge and of fuzzy models of spatial relations, which establish the links between concepts and image space, thus addressing the semantic gap issue (Hudelot et al. 2008). Due to the sequential nature of the process, the errors are potentially propagated. Backtracking may then be needed, as proposed by Fouquier et al. (2012).

To overcome the problems raised by sequential approaches while avoiding the need for an initial segmentation, an original method was proposed by Nempont et al. (2013). It still employs a structural model, but solves the problem in a global fashion. A solution is the assignment of a spatial region to a model object, in a way that satisfies the constraints expressed in the model. A progressive reduction of the solution domain for all objects is achieved by excluding assignments that are inconsistent with the structural model. Constraint networks constitute an appropriate framework for both the formalization of the problem and the optimization (see chapter "Constraint Reasoning" of Volume 2 for constraint reasoning methods). An original feature of this approach is that the regions are not predetermined, but are instead constructed during the reduction process. The image segmentation and recognition algorithm therefore differs from an annotation procedure, and no prior segmentation of the image into meaningful or homogeneous regions is required. More precisely, a constraint network is constructed from the structural model, and a propagation algorithm is then designed to reduce the search space. Finally, an approximate solution is extracted from the reduced search space. This procedure is illustrated in Fig. 1, using the interpretation of a brain MRI as an example. The solution space for the left caudate nucleus CNl is derived from the constraint "CNl is exterior to the left lateral ventricle LVl". Once the propagation process terminates, the solution space is typically reduced substantially for all of the model structures. The final segmentation and recognition results can then be obtained using any segmentation method that is constrained by this solution space. An example of result in a pathological case is illustrated on one slice in Fig. 2.

This approach has been extended by Vanegas et al. (2016) to deal with complex relations, involving groups of objects, unknown numbers of instances of concepts in the images and fuzzy constraints, for applications in remote sensing image understanding.

A concluding message is that model based understanding is a growing research topic, at the cross-road of image processing, computer vision and pattern or object recognition on the one hand, and of artificial intelligence on the other hand. The association between generic structural models and specific information related to the

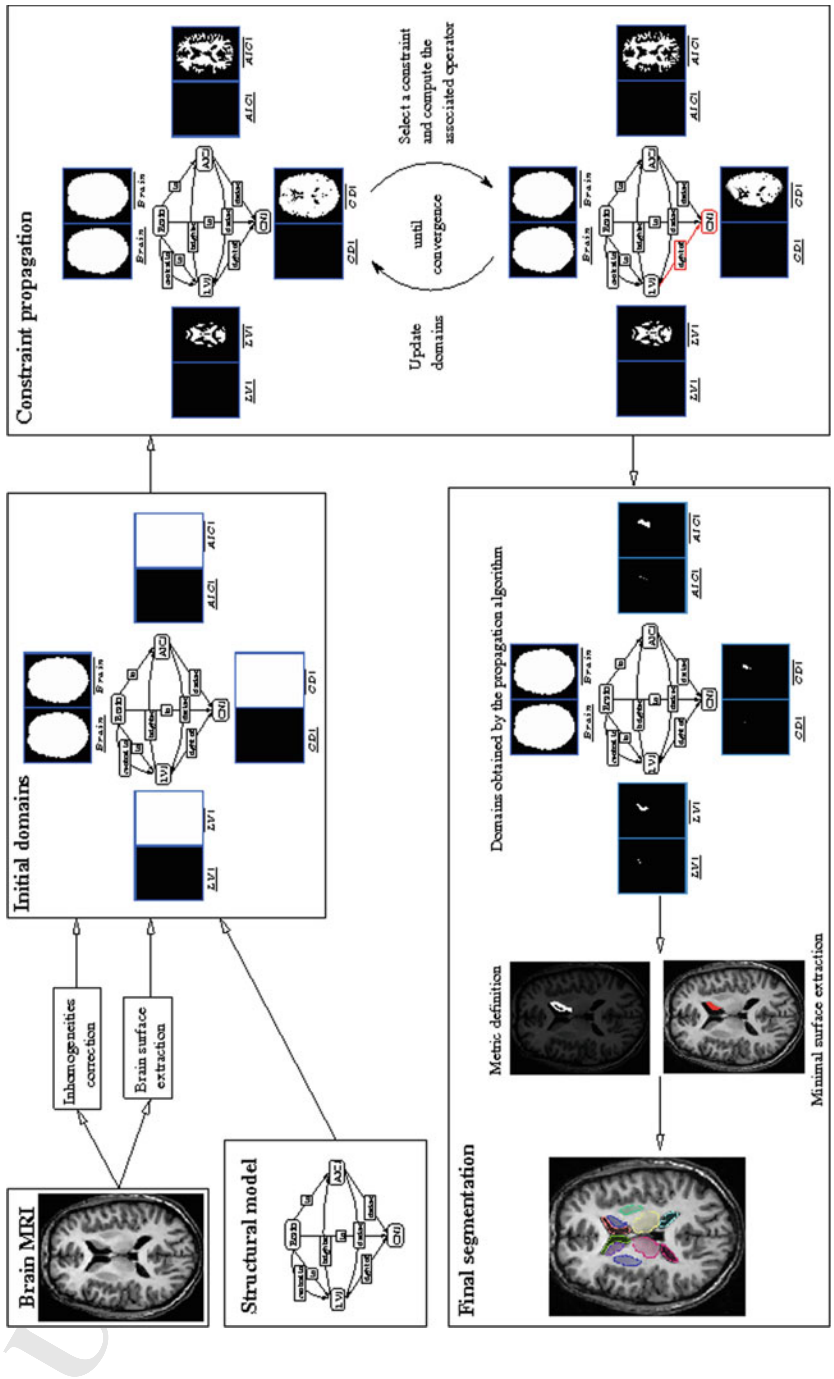


Fig. 1 Overview of the constraint propagation method for segmenting and recognizing brain structures in MRI (Nempont et al. 2013)

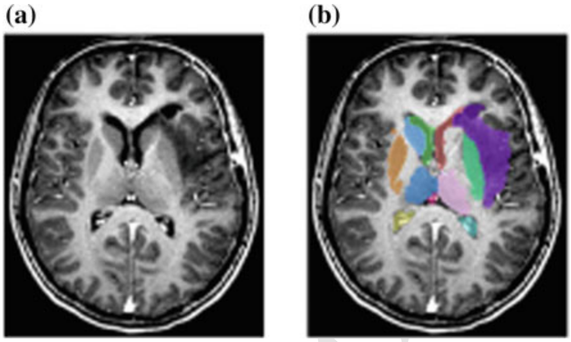


Fig. 2 **a** Axial slice of a 3D MRI of a patient with a brain tumour. **b** Segmentation and recognition results for several internal brain structures (Nempont et al. 2013)

context, accounting for uncertainty and variability, allows one to cope with the semantic gap problem and to propose computationally efficient methods to solve it. These approaches are currently further developed for image and video annotation, segmentation and recognition of structures, spatial reasoning for image exploration, or the derivation of high level descriptions of the content of images or image sequences.

3 Code Supervision for Automatic Image Processing

The need for automatic image analysis software is becoming increasingly pressing as digital image emerges as a privileged source of information. Acquisition devices now provide access to previously unknown or inaccessible data that are of strategic importance in many fields such as medicine, security, quality control, astronomy, environmental protection. However, the multiplicity of these devices leads to the production of an ever-expanding volume of data that is impossible to exploit manually.

Image processing is a preliminary stage that aims to prepare the images for subsequent analysis by humans or interpretation systems. It covers all objectives of image-to-image transformation that are intended to reduce, refine or organize the initial data. Five image processing objectives are usually distinguished: data *compression*, *enhancement* of visual rendering, *restoration* of missing information, *reconstruction* of spatio-temporal information (3D or motion), *segmentation* into more abstract primitives (regions or contours) and *detection* of known objects. Image processing has no decision-making power, but its role is crucial since it must ensure that changes on images are made without loss or alteration of the relevant information.

Image processing research traditionally provides its expertise in the form of image processing algorithms. Many algorithms covering a wide range of operations have been developed. Each algorithm is developed on a presupposed model of information to be processed, which determines its domain of applicability and effectiveness. Therefore, there is no universal algorithm. A concrete application should combine several of these algorithms according to a top-down, bottom-up or mixed processing strategy. Thus, the development consists in selecting, tuning and linking appropriate algorithms. However, appropriate use of image processing algorithm libraries requires highly specialized expertise to know when and how to utilize the algorithms.

Code supervision systems are designed to provide users with a tool to build their own applications by exploiting a library of precoded algorithms. Users no longer need to be experts in image processing. Their role is focused on the formulation of application objectives. It is the system responsibility to control the code library for building programs suited to the application objectives.

3.1 Formulation of Application Objectives

The formulation of application objectives is of paramount importance because it is used by the system to guide selection, tuning and chaining of codes. Two categories of information should be given by users for an exhaustive formulation:

1. The *definition of the image class* is required to bridge the sensory and semantic gaps (Smeulders et al. 2000) (see Fig. 3). Part of the definition should describe the image acquisition process in order to restore information about the observed scene that were lost, altered or hidden during the image production. Another part should assign a semantics to the scene content in order to specify information that has to be considered as relevant for that precise application.
2. The *specification of the processing goals* is required to clarify the role of the application in the complete analysis system.

Image Class Definition

Various models of image class definition have been proposed in the literature whether the definition is done by extension or by intension.

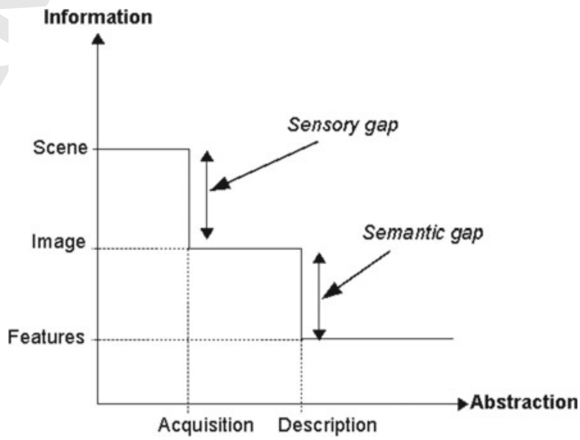


Fig. 3 The sensory gap results from the loss of information between the reality of a scene and its representation as an image. The semantic gap separates the interpretation of a scene that anyone can make from an image-based representation and from a feature-based description (Smeulders et al. 2000)

420055_1_En_10_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:18/5/2019 Pages: 28 Layout: T1-Standard

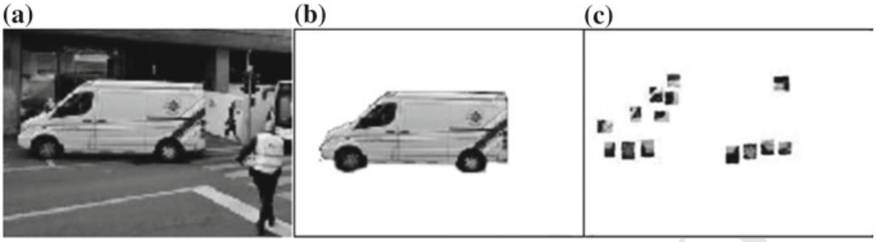


Fig. 4 Two ways to extensionally describe the vehicle in figure **a**: **b** by a mask that specifies the object pixels, **c** by a list of patches around points of interest

An extensional definition represents information using an iconic dictionary built with image parts. These parts can be specified either by *masks* or by *patches*. A mask delineates an object of interest or a specific image area (see example in Fig. 4b.). They are used by the system to automatically extract a set of feature values of the specified object (colour, shape, size, etc.) or a set of image characteristics of the specified area (type of acquisition noise, illumination distribution, etc.). A patch is a thumbnail extracted from a sample image that isolates one salient part of an object of interest (often localized around a point of interest as shown in Fig. 4c). They are used by the system to detect instances of these objects in images from their characteristic parts (Agarwal et al. 2004; Leibe et al. 2008). The benefit of extensional definition is to limit the cognitive load of users since no representation language is required. The drawback is that the same feature extraction or patch selection algorithms are used for all applications. Thus, a part of the application definition is assigned by the system and cannot be adapted to each application.

An intensional definition represents information about images using a linguistic description. It provides a language to represent the acquisition effect and the scene content semantics. Ontologies are widely used for this purpose (Hunter 2001; Bloehdorn et al. 2005; Town 2006; Renouf et al. 2007; Anouncia and Saravanan 2007; Maillot and Thonnat 2008; Neumann and Möller 2008; Gurevich et al. 2009). The description language is usually constructed from an *ontology domain* that provides the language primitives. The description of a particular image class is an *application ontology* that is obtained by selection and reification of domain ontology primitives (Cãmara 2001). For example, Maillot and Thonnat (2008) propose the "Ontology of Visual Concepts" which defines the concepts of texture, colour, geometry and topological relations. Figure 5 gives a textual representation of the definition of a pollen grain with this ontology. To better reflect the variability of the visual manifestations of the objects in the scene, the language accepts qualitative values for the features such as ("*pink*", "*very circular*", "*strongly oblong*") and for the spatial relations such as ("*in front of*", "*close to*"). The advantage of this definition is to take greater advantage of the user's expertise about scene content and thus better capture application variability. However, the construction of the solution requires quantitative values. Therefore, intensional definition must address the problem of symbol grounding in order to connect linguistic symbols to image data values. Sym-

420055_1_En_10_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:18/5/2019 Pages: 28 Layout: T1-Standard

Fig. 5 Textual representation of the definition of a pollen grain of type "poaceae" from the "Ontology of Visual Concepts" proposed by Maillot and Thonnat (2008)

bol grounding can be based on dictionaries such as the "Colour Naming System" (Berk et al. 1982) where the HSL space is divided into 627 distinct colours, each of them labelled with a name, or the "Texture Naming System dictionary" (Rao and Lohse 1993). However, most often symbol grounding is seen as a learning problem from a set of masks. Therefore, usually mixed approaches are preferred. Intensional definition is completed with extensional definition that allows anchoring ontology concepts into data (Maillot and Thonnat 2008; Hudelot et al. 2008; Clouard et al. 2010).

Goal Specification

The specification of application goals can be made either by examples of the expected results or by tasks to perform.

According to specification by example, a goal is formulated through reference images containing the representation of the results to be obtained on test images. Three different representations of the expected results have been proposed in the literature:

- *Sketches* are lines drawn by the user on test images that give examples of the expected contours or regions boundaries (Draper et al. 1999), as in the Fig. 6a.
- *Manual segmentations* give the region areas to be obtained on test images (Martin et al. 2006), as in the example in Fig. 6b.
- *Scribbles* are markers that indicate regions of interest without completely delineate them (Protire and Sapiro 2007). Generally, scribbles are lines drawn directly inside the regions of interest and inside the background region, as in Fig. 6c.

The advantage of the specification by example paradigm is its quantitative nature since it takes values directly into the image data. In addition, it reduces the cognitive load of users because no specialized vocabulary is required. The drawback is that a reference image is not sufficient to formulate all kinds of goals. Only segmentation, detection and possibly enhancement goals are really addressed. Compression, restoration and reconstruction goals are not straightforward. Moreover, it does not cover all image classes. In particular, it is tedious to implement for 3D images and image sequences. Finally, there is no means for varying constraints attached to goals, such as "*prefer false detection to misdetection*" or "*prefer no result to imperfect result*".

420055_1_En_10_Chapter ✓ TYPESET ☐ DISK ☐ LE ✓ CP Disp.:18/5/2019 Pages: 28 Layout: T1-Standard

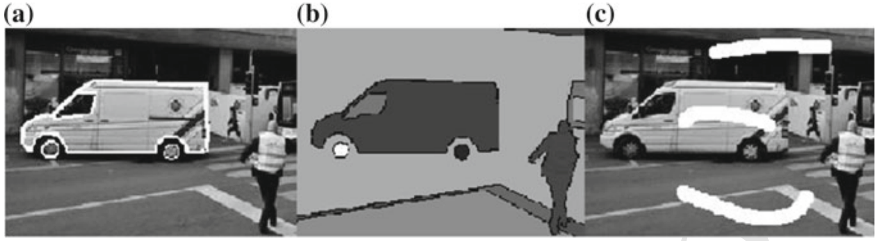


Fig. 6 Three different approaches to specify a goal by example: **a** by sketch, **b** by manual segmentation, **c** by scribbles

The specification by task paradigm requires a language. A task describes a system functionality by means of a sentence, such as "detect object vehicle" or "segment the image". The advantage of this approach is that it is possible to associate constraints to the task in order to restrict its scope. Moreover, all image processing objectives can be covered: it is sufficient to name a task and related constraints. The drawback is that the formulation is qualitative with no real link to the image data. This has two important consequences: first, specification by task is not strongly grounded into data, and secondly, there is only a finite number of possible objective formulations. That is why recent approaches use mixed approaches that combine specification by task and specification by example paradigms. Figure 7 presents an ontology (Clouard et al. 2010) that covers the definition of the image class by mixing intensional and extensional approaches and specifying goals by mixing approaches by task and by example.

3.2 Code Supervision

The formulation of application objectives is the prerequisite for the development of a solution as a processing chain. In the paradigm of code supervision (Thonnat and Moisan 2000), image processing techniques are implemented as independent executable codes and stored in a library. An image processing program is represented in canonical form as a directed graph of codes. Links between codes describe network of images and parameter values exchanged between codes. For example, Fig. 8 shows a processing chain that performs edge detection by difference of two Gaussians.

The problem of code supervision was addressed in several ways in the literature of which the most advanced are:

- competitive strategy;
- plan skeleton instantiation;
- case-based reasoning;
- chain planning;
- incremental result construction.

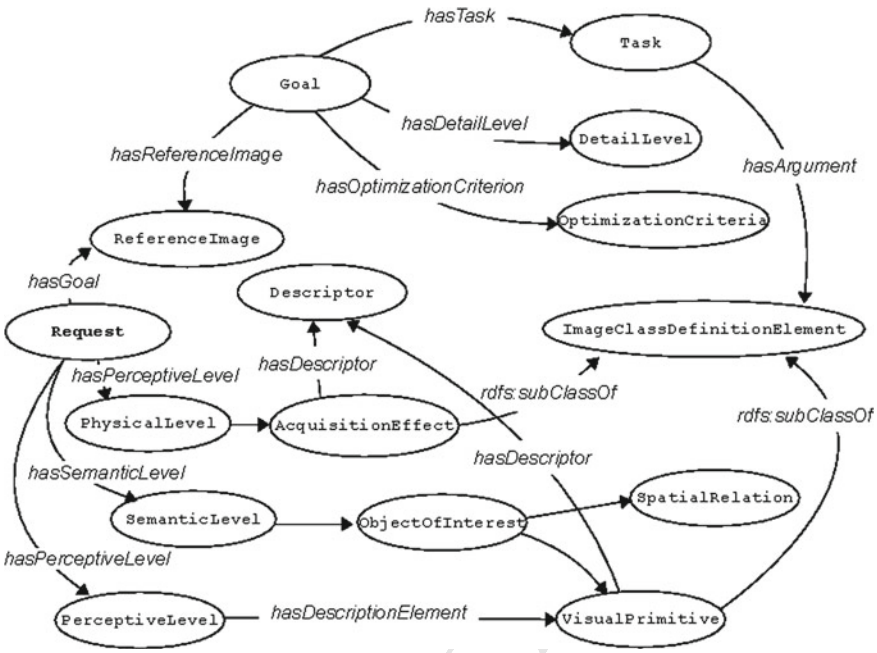
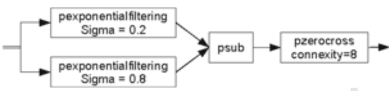


Fig. 7 The concepts of an ontology for formulating image processing goals (Clouard et al. 2010)




```
pexponentialfiltering 0.2 image1 image2
pexponentialfiltering 0.8 image1 image3
psub image2 image3 image4
pzerocross 8 image4 result
```

Fig. 8 A program is a graph of parametrized executable codes. On the left is given the representation of an edge detection algorithm using the DOG (Difference of Gaussian) in the form of a code graph. On the right, the same algorithm is represented as a script of executable codes

Competitive Strategy

The main idea behind this approach is to exploit the competition between several predefined processing strategies. For example, Charroux and Philipp (1995) execute several image segmentation chains in parallel, and then build the final result with the best segmented regions yielded by each of these chains. The quality of a region is measured by its degree of membership to domain object classes, calculated by a classifier trained to recognize the domain object classes from masks made on sample images.

Martin et al. (2006) create competition between multiple image segmentation chains and then select the best chain with the best settings. The selection is made off-line through supervised learning where a set of sample images with related handmade reference segmentation is used to train the classifier. The resulting chain, with its setting, is the one that minimizes the distance between the segmentation obtained on test images and the reference segmentation made for these images.

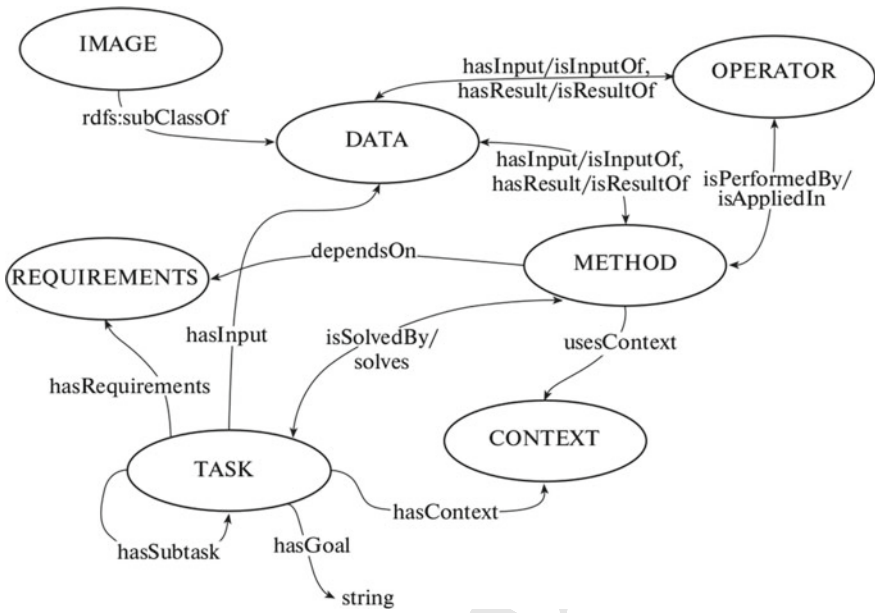


Fig. 9 Concepts and basic elements of an image processing ontology which specifies how to solve a task using operators with regard to a specific context (Gurevich et al. 2009)

The advantage of this approach is that it requires no explicit expertise. Only reference object masks or reference images must be provided. The drawback is that it relies on processing chains that are fixed and in finite number. Parameter tuning is the only possible adaptation.

Plan Skeleton Instantiation

This is certainly the approach that has generated the higher number of systems, with pioneering work such as: OCAPI (Clément and Thonnat 1993), VSDE (Bodington 1995), CONNY (Liedtke and Blömer 1992), COLLAGE (Lansky et al. 1995) or MVP (Chien and Mortensen 1996).

The processing expertise is encoded in hierarchical plan skeletons that combine along several decomposition levels a task corresponding to a problem with a set of codes that constitute elements of a possible chain of processing. Plan skeletons are encoded as AND/OR trees that indicate how a task can be decomposed into subtasks. Production rules are attached to each node. They are used to select the most appropriate branch of the decomposition and parameter values with regard to formulation elements.

Figure 9 presents an ontology that models the way to solve a task with a sequence of operators with regard to a specific context.

Compared to competitive strategy, this approach allows chain adjustment to the specifications given in the formulation of objectives. However, it requires knowing how to identify and represent the expertise for each possible problem type.

Case-Based Reasoning

Case-based reasoning exploits processing chains built successfully for past applications to process a new "similar" one.

In image processing, this approach has been used to build processing plans (Charlebois 1997; Ficet-Cauchard et al. 1999) or to find out convenient set of parameters to configure a general processing chain (Perner et al. 2005; Frucci et al. 2008). The reasoning is based on the analysis of the problem formulation to try to find a similar case. The retrieved case is then adapted to the context of the current problem. If there is no similar case, then a new case has to be learned and stored in the database.

Case-based reasoning does not require explicit representation of processing expertise. However, the critical point of this approach lies in the adaptation of cases to the particular context of the application that is of considerable importance in image processing regarding the high variability of images in a class.

Chain Planning

Unlike previous approaches which explicitly encode a set of processing chains, in chain planning the processing chains are built dynamically.

Systems using linear planning are based on modelling a type of expression that can be propagated along the processing chains. The reasoning is focused on the operations to be applied to the initial expression to build the expected final expression. The initial expression is the formulation provided by users in intensional or extensional form. In the latter form, expression is constructed by automatic extraction of features in sample images. The generation of chains can be combinatorial. In this case, each operator in the chain is modelled by a list of preconditions and a list of effects on the expression, as in the system EXTI (Dejean and Dalle 1996). But, the generation of chains can also be achieved by production rules attached to nodes that select the next operators according to the current expression, as in systems LLVE (Matsuyama 1989) and SOLUTION (Rost and Mnkel 1998).

The planning approach creates chains from scratch for each application. However, it faces the difficulty to model the problem as an expression that can be propagated along processing chains and especially the difficulty of having to a priori estimate the impact of operations on the expression. To improve planning efficiency, this problem has also been addressed using a hierarchical planning. The BORG system (Clouard et al. 1999) used a blackboard to build plans using multiple levels of abstraction. The initial goal formulated by the user is gradually divided into more and more precise subtasks until they correspond to executable codes. Knowledge sources encode various decomposition alternatives of a task to lower level subtasks. Figure 10 presents an example of construction of such a plan.

In all cases, the final application is the processing chain built operator by operator, which produce a standalone program. To limit the impact of choices made during the construction of chains, Draper et al. (1999), with the ADORE system, propose to keep all the alternative chains in the program. This system then uses a Markov decision process to dynamically choose the best path in these chains during the execution of the solution, from features automatically extracted from the processed image.

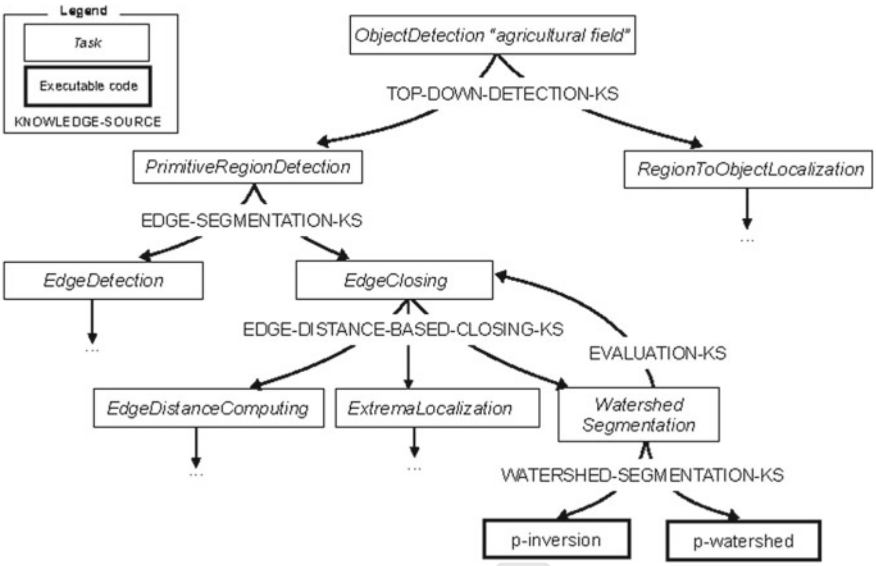


Fig. 10 Excerpt from a hierarchical processing plan for the detection of agricultural fields in aerial images (Clouard et al. 2010)

Incremental Result Construction

Incremental construction of results proceeds by gradual and controlled evolution of the input image to the desired output image. This approach can be seen as dual of the previous approaches in the sense that the reasoning is focused on the analysis of data produced after application of processing. The image processing algorithms are completely split into a set of production rules (Nazif and Levine 1984) or independent rational agents (Boucher et al. 1998; Bovemkamp et al. 2004). In such an approach, there is no explicit strategy of generation of processing chains. The reasoning remains focused on the analysis of the current state of the image after application of the first processing in order to determine the next processing to be applied in the case of production rules or resolve data access conflicts in the case of multi-agent.

The design of such systems requires a knowledge acquisition phase. Nevertheless, the decentralized control makes the acquisition of such knowledge easier, since it is not necessary for the knowledge engineer to explain the resolution strategies. However, the overall resolution process remains complex to master because convergence towards a solution is only guaranteed by the action of rules or agents that have only a local vision of their effects. Each rule or agent is responsible for estimating the value of its contribution compared to the current state of the resolution. This limit often requires adding abstraction levels in the hierarchy of rules or rational agents to have a more global vision of the resolution.

3.3 Conclusion

The challenge of the research on code supervision for automatic image processing and image analysis is to develop solutions that allow image consumers unskilled in image processing (e.g., geographers, biologists, librarians, special effect technicians) to design their own software alone. In shorter term, the goal is to build configurable systems that help vision engineers rapidly deploy dedicated applications without any programming activity.

Today, the results of these works are exploited in recent research in semantic image indexing, content-based image search and video analysis. These problems are also addressed using statistical methods with spectacular results for face detection or object recognition for example. They operate from the extensional definition of image classes using comparison with learned sample images. But these statistical methods are insufficient in cases of complex scenes or problems other than detection. In such situations, artificial intelligence methods and techniques are an undeniable asset. They cover a wider variety of applications and moreover they better take into account of the user needs. In this context, statistical methods are integrated as regular codes that can be used in specific cases.

However, the design of systems covering a wide range of applications with high efficiency remains a challenge. With this in mind, some current research work is directed towards the development of solutions based on human machine interaction, which emphasize collaboration to jointly converge towards building a suitable processing chain, each bringing its skills, the user's knowledge of the problem and the system knowledge of image processing.

4 Machine Learning for Robotics

Most industrial robots of the last century were used in highly structured and controlled environments such as assembly lines. All day long, they were realizing highly repetitive and specialized tasks without any room for uncertainty and away from human workers, mostly for security issues.

In the early 21st century, a new generation of robots is now emerging, whose employment context is fundamentally different (see also chapter "Robotics and Artificial Intelligence" in this volume). These so-called "personal" robots, whether they be food processors, playful companions or patient support, will have to perform extremely varied tasks in unknown changing environments, where uncertainty is omnipresent, and in direct contact with their users, who will not be experts in robotics. In this context, specifying in advance the behaviour of robots for any possible situation and for any possible task is no longer possible. The only reasonable alternative is to equip these versatile robots with abilities to learn and adapt to their environment.

While machine learning methods have been extensively developed in the last two decades, the robotic framework confronts these methods to specific constraints such as the limited duration of the experiments, the often prohibitive cost of failures, the need to operate in real time or the large number of high-dimensional problems to be solved.

Therefore, no unifying theoretical framework has yet imposed itself to formalize the corresponding robot learning problems, and there are many attempts of varied natures to equip robots with learning abilities.

Part of the work is based on different theoretical machine learning frameworks (see chapters “Statistical Computational Learning” and “Reinforcement Learning” of Volume 1, and “Designing Algorithms for Machine Learning and Data Mining” of Volume 2): supervised learning, reinforcement learning, inductive learning, etc. to build tools specifically adapted to the robotic constraints.

Another part, which intersects significantly with the former, relies on understanding learning processes in biological systems to develop new methods inspired from these processes. This is the case of imitation learning, developmental robotics, evolutionary robotics, or various neuro-mimetic approaches to learning, for example. The intersection arises because these methods will eventually use machine learning tools designed within the first approach.

4.1 Machine Learning Methods and Robotics

Of all the approaches mentioned above, the one that provides the most obvious alternative for replacing direct programming of behaviour is imitation learning, also called learning by demonstration. This approach is relatively well developed and produced many significant results in recent years, through quite different methodological approaches. Some researchers use motion capture tools to record the movement of humans trying to perform a task in a particular context, then make sure that the robot performs the same movement in the same context. This last point requires to solve a problem known as the “correspondence problem” when the geometry, kinematics and dynamics of the human and the robot are significantly different, which is usually the case, except for a few humanoid robots. To avoid solving this correspondence problem, another approach consists in driving the robot through a remote operation system to make it realize the required movement once, and then to build on the recorded movement to perform it again and again. However, those two approaches pose a widespread problem: the circumstances being never exactly the same, the recorded movement is never perfectly adequate and the robot must adapt to these variations. For various syntheses or particularly outstanding work in the context of learning by imitation, we refer the reader to the work by Atkeson et al. (1997), Schaal (1999), Ijspeert et al. (2002), Calinon (2009), Coates et al. (2008), Ratliff et al. (2009).

Another approach that directly takes into consideration the need to generalize is to solve an “inverse reinforcement learning” (or inverse optimal control) problem. The idea is to consider a set of trajectories made by experts as optimal and extract the cost function that experts seem to have followed. Given the cost function, an optimization algorithm can be used to generate the new robot movements that optimize the same cost function (Abbeel 2008).

Learning by imitation is not enough to solve all the problems posed by the need for robots that adapt to their environment. Indeed, in the general context of use described above, it is not possible to show the robot what should be its behaviour in all situations it would be likely to encounter. To go further, it is necessary that the robot is able to adapt its behaviour to unexpected situations. For this, one must still provide the robot with a capacity to assess the quality of its behaviour in a given situation, which can be done through a cost function. Learning how to improve one's behaviour by seeking to minimize a cost function (or maximize a performance function) is a problem that is formalized within the framework of reinforcement learning (Sutton and Barto 1998). The difficulty encountered in robotics to use reinforcement learning methods arises because these methods were originally developed in the problem solving context in which situations and actions are finite and limited, while in robotics problems are often continuous or very large. However, many recent algorithmic advances helped obtain increasingly significant results in this area (Stulp and Sigaud 2012).

Moreover, the command used for complex robots often uses kinematics, velocity kinematics and dynamics models of these robots, mainly for planning by determining the immediate response of the robot to a particular command. Identification is the activity of determining these models using a set of simple experiments that extract all relevant variables. The supervised learning methods that approximate functions from elementary data provide an interesting alternative to traditional parametric identification, to the extent that a robotic model is a function that can be estimated from the sensors of the robot. On the one hand, these methods require no a priori assumption on the shape of the models (Stulp and Sigaud 2015). Moreover, model learning can be performed during the robot operation, thus avoiding a tedious preliminary phase and, above all, allowing to immediately adapt the model in case of alteration of the robot or variation of the mechanical conditions of use. Though these supervised learning methods are still largely confined to learning robot models themselves (D'Souza et al. 2001; Salaun et al. 2010), they begin to tackle more original questions related to the interaction with a priori unknown objects (Vijayakumar et al. 2005), which falls within the more ambitious context of use that we described in the introduction to this section.

Robot learning finds its most compelling application context in the interaction between a robot and a human (Najar et al. 2015). Indeed, this context prominently requires rapid adaptation to a changing context from the robot and provides the framework within which imitation learning comes most naturally. Imitation is also a kind of human-robot interaction, allowing to consider the latter area as more general than the former. There are also research works that do not fit in previous frameworks, such as research on the social acceptability of behaviour of robots (Kruse 2010) or human-robot verbal interaction in a cooperation framework (Dominey 2007).

The human-robot interaction can be physical, when either of the protagonists exerts a force on the other. This is the case for example in the context of robotic assistance and rehabilitation, when it comes to helping patients with motor disorders (Saint-Bauzel et al. 2009). The implementation of learning technologies in this context is a new trend (Pasqui et al. 2010). The interaction may also be simply communicative, whether through the spoken word or through other nonverbal methods (Dominey and Warneken 2009). The interaction may finally be fully implicit, when the human and the robot adapt their behaviour to each other without any communication, just by adjusting their behaviour to the behaviour observed in the other.

4.2 Bio-inspired Learning and Robotics

A second approach to learning in robotics is to attempt to replicate the learning mechanisms found in living beings. The goal is to endow robots with adaptive properties similar to those of animals or humans, which is far from the case today. Such an approach is likely to improve the development of adaptive mechanisms for robots. Furthermore, and vice versa, this approach is likely to contribute to progress in understanding the adaptation mechanisms of living beings, through validation or invalidation by robotics experiments (Guillot and Meyer 2008).

These bio-inspired approaches can take very different forms depending on the level at which the adaptation mechanisms are integrated. Indeed, living systems are characterized by a complex hierarchy of physiological and psychological processes at different scales, and adaptive mechanisms can be found at most of these levels, if not all.

Broadly speaking, there are two main research lines:

- the first finds its inspiration in psychological research about child development and is called "developmental robotics". It is mainly concerned with works modelling the cognitive learning abilities of babies and young children (Lungarella et al. 2003; Oudeyer et al. 2007; Quinton et al. 2008) and is particularly interested in solving the so-called "symbol grounding problem" that any artificial intelligence system is facing (Harnad 1990);
- the second is rather inspired from neuroscience research and proposes "neuromimetic" approaches, which can be clustered into two main families. The first is interested in decomposing the brain into distinct functional areas and proposes models whose components mimic the functions of these different areas. For example, one model the learning capabilities of rodents by building a neuro-mimetic model of the rat basal ganglia, which are deep nuclei of the brain which are believed to play a role in the evaluation of our behaviour (Doya 2000; Lesaint et al. 2014). The second focuses instead on the elementary computational properties of neurons, again at different levels, depending on whether one looks at the average activity of the neuron over time or at its propensity to issue elementary pulses according to a specific dynamics.

The central challenge that faces this general bio-inspired approach is due to the complex stack of integration levels. For a given adaptive phenomenon, it is sometimes difficult to determine whether a unique level of integration can account for the phenomenon, or whether the mechanisms from several levels should systematically be combined. In this context, the robot proves an invaluable tool for the advance of knowledge in living sciences by providing a demanding experimental validation framework in which different theories can be analysed or compared.

4.3 Current Challenges

The desire to endow robots with learning ability is doubtlessly not new, but the corresponding research has substantially grown in recent years, with the emergence of many workshops dedicated to this topic in the main robotics conferences, the publication of numerous special issues in journals, or the growing number of dedicated summer schools. The result of this rapid growth is a burgeoning development in which many approaches are being developed in parallel in sometimes very different directions, often attacking very different problems. It seems that in the more or less close future, all of these searches should be structured and that new models combining different mechanisms should emerge from this abundance. A very recent and major evolution in robot learning results from the emergence of deep learning techniques (LeCun et al. 2015). The outstanding pattern recognition capabilities of these techniques and their focus on learning flexible representations from data opens new perspective on solving the symbol grounding problem in a developmental robotics perspective. But the methodological constraints of developmental robotics differ from those of standard pattern recognition challenges, thus the emergence of dedicated deep learning techniques is required with a potentially huge impact on robot learning (Sigaud and Droniou 2016).

5 Conclusion

In this chapter, far from being exhaustive, illustrations have shown convergence areas between artificial intelligence, computer vision, pattern recognition, learning and robotics. These convergences can be found in other domains, such as speech recognition and automatic natural language processing. Associating theories and methods from different domains is an ever growing approach, and leads to important developments and original research works. In image understanding, high level approaches use more and more intensively knowledge representation methods and reasoning services. For instance, abduction and revision, integrating learning and uncertainty models, can be used for image or video understanding (Atif et al. 2013) (see also chapters “Knowledge Representation: Modalities, Conditionals and Nonmonotonic Reasoning”, “Reasoning with Ontologies”, “Belief Revision, Belief Merging and Information Fusion”, “Multicriteria Decision Making” and “Decision Under Uncertainty” of Volume 1). In parallel to these model and knowledge based methods, a large field of research is now based on learning (and in particular deep learning) methods, with impressive results based on large training data sets (and without exploiting knowledge) (LeCun et al. 2015; Vinyals et al. 2015) (see also chapters "Statistical Computational Learning" and "Reinforcement Learning" of Volume 1, and "Designing Algorithms for Machine Learning and Data Mining" of Volume 2). Man-machine interactions can also support new solutions, as mentioned for code supervision, but also for other domains, such as robotics. Finally, the multiplication of methods and models incite researchers to combine their advantages.

References

Abbeel P (2008) Apprenticeship learning and reinforcement learning with application to robotic control. PhD thesis, Stanford University, Computer Science

Agarwal S, Awan A, Roth D (2004) Learning to detect objects in images via a sparse, part-based representation. IEEE Trans Pattern Anal Mach Intell 26(11):1475–1490

Anouncia SM, Saravanan R (2007) Ontology-based process plan generation for image processing. Int J Metadata Semant Ontol 2(3):211–222

Atif J, Hudelot C, Bloch I (2013) Explanatory reasoning for image understanding using formal concept analysis and description logics. IEEE Trans Syst Man Cybern 44:552–570

Atkeson CG, Moore AW, Schaal S (1997) Locally weighted learning. Artif Intell Rev 11(1):11–73

Bengoetxea E, Larranaga P, Bloch I, Perchant A, Boeres C (2002) Inexact graph matching by means of estimation of distribution algorithms. Pattern Recognit 35(12):2867–2880

Benz U, Hofmann P, Willhauck G, Lingenfelder I, Heynen M (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS J Photogramm Remote Sens 58(3–4):239–258

Berk T, Brownston L, Kaufman A (1982) A new color-naming system for graphics languages. IEEE Comput Graph Appl 2(3):37–44

Bloch I (2003) Traitement d'images. In: Bouchon-Meunier B, Marsala C (eds) Traitement de données complexes et commande en logique floue (Chap. 3). Hermes, Paris, pp 95–152

Bloch I (2005) Fuzzy spatial relationships for image processing and interpretation: a review. Image Vis Comput 23(2):89–110

Bloch I (2006) Spatial reasoning under imprecision using fuzzy set theory, formal logics and mathematical morphology. Int J Approx Reason 41:77–95

Bloch I (2008) Information fusion in signal and image processing. ISTE-Wiley, London

Bloch I, Géraud T, Maître H (2003) Representation and fusion of heterogeneous fuzzy information in the 3D space for model-based structural recognition - application to 3D brain imaging. Artif Intell 148(1–2):141–175

Bloehdorn S, Petridis K, Saathoff C, Simou N, Tzouvaras V, Avrithis Y, Handschuh S, Kompatsiaris Y, Staab S, Strintzis M (2005) Semantic annotation of images and videos for multimedia analysis. Second European Semantic Web Conference (ESWC). Crete, Greece, pp 592–607

Bodington R (1995) A software environment for the automatic configuration of inspection systems. International Workshop on Knowledge Based Systems for the reUse of Program Libraries (KBUP). Sophia Antipolis, France, pp 100–108

Boucher A, Doisy A, Ronot X, Garbay C (1998) A society of goal-oriented agents for the analysis of living cells. Artif Intell Med 14(1–2):183–199

Bovemkamp E, Dijkstra J, Bosch J, Reiber J (2004) Multi-agent segmentation of IVUS images. Pattern Recognit 37:647–63

Calinon S (2009) Robot programming by demonstration: a probabilistic approach. EPFL/CRC Press, Lausanne

Cãmara G, Engenhofer M, Fonseca F, Monteiro A (2001) What's in an image? International Conference on Spatial Information Theory: Foundations of Geographic Information Science, Morro Bay, CA 2205:474–488

Cesar R, Bengoetxea E, Bloch I, Larranaga P (2005) Inexact graph matching for model-based recognition: evaluation and comparison of optimization algorithms. Pattern Recognit 38(11):2099–2113

Charlebois D (1997) A planning system based on plan re-use and its application to geographical information systems and remote sensing. PhD thesis, University of Ottawa, Canada

Charroux B, Philipp S (1995) Interpretation of aerial images based on potential functions. 9th Scandinavian Conference on Image Analysis. Uppsala, Sweden, pp 671–678

Chien S, Mortensen H (1996) Automating image processing for scientific data analysis of a large image database. IEEE Trans Pattern Anal Mach Intell 18(8):854–859

Chein M, Mugnier M (2008) Graph-based knowledge representation: computational foundations of conceptual graphs. Springer, New York

Clément V, Thonnat M (1993) A knowledge-based approach to integration of image procedures processing. CVGIP Image Underst 57(2):166–184

Clouard R, Porquet C, Elmoataz A, Revenu M (1999) Borg: a knowledge-based system for automatic generation of image processing programs. IEEE Trans Pattern Anal Mach Intell 21(2):128–144

Clouard R, Renouf A, Revenu M (2010) An ontology-based model for representing image processing objectives. Int J Pattern Recognit Artif Intell 24(8):1181–1208

Coates A, Abbeel P, Ng AY (2008) Learning for control from multiple demonstrations. In: 25th International Conference on Machine Learning, pp 144–151

Colliot O, Camara O, Bloch I (2006) Integration of fuzzy spatial relations in deformable models - application to brain MRI segmentation. Pattern Recognit 39:1401–1414

Conte D, Foggia P, Sansone C, Vento M (2004) Thirty years of graph matching in pattern recognition. Int J Pattern Recognit Artif Intell 18(3):265–298

Coradeschi S, Saffiotti A (1999) Anchoring symbols to vision data by fuzzy logic. In: Hunter A, Parsons S (eds) ECSQARU'99. LNCS, vol 1638. Springer, London, pp 104–115

Crevier D, Lepage R (1997) Knowledge-based image understanding systems: a survey. Comput Vis Image Underst 67(2):161–185

Dejean P, Dalle P (1996) Image analysis operators as concept constructors. IEEE Southwest Symposium on Image Analysis and Interpretation. San Antonio, USA, pp 66–70

Denœux T (2008) Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. Artif Intell 172(2–3):234–264

Deruyver A, Hodé Y (1997) Constraint satisfaction problem with bilevel constraint: application to interpretation of over-segmented images. Artif Intell 93(1–2):321–335

Deruyver A, Hodé Y (2009) Qualitative spatial relationships for image interpretation by using a conceptual graph. Image Vis Comput 27(7):876–886

Desachy J (1990) ICARE: an expert system for automatic mapping from satellite imagery. In: Pau LF (ed) Mapping and spatial modelling for navigation, vol F65. NATO-ASI. Springer, Berlin

Dominey PF (2007) Sharing intentional plans for imitation and cooperation: integrating clues from child developments and neurophysiology into robotics. In: Proceedings of the AISB 2007 Workshop on Imitation

Dominey PF, Warneken F (2009) The basis of shared intentions in human and robot cognition. New Ideas Psychol

Doya K (2000) Complementary roles of basal ganglia and cerebellum in learning and motor control. Curr Opin Neurobiol 10:732–739

Draper B, Bins J, Baek K (1999) ADORE: adaptive object recognition. International Conference on Vision Systems (ICVS). Las Palmas de Gran Canaria, Spain, pp 522–537

D'Souza A, Vijayakumar S, Schaal S (2001) Learning inverse kinematics. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 1:298–303. https://doi.org/10.1109/IROS.2001.973374

Dubois D, Prade H (1980) Fuzzy sets and systems: theory and applications. Academic, New-York

Dubois D, Prade H (1988) Possibility theory. Plenum, New-York

Dubois D, Prade H (2001) La problématique scientifique du traitement de l'information. Inf-Interact-Intell 1(2):79–98

Ficet-Cauchard V, Porquet C, Revenu M (1999) CBR for the management and reuse of image-processing expertise: a conversational system. Eng Appl Artif Intell 12(6):733–747

Fouquier G, Atif J, Bloch I (2008) Sequential spatial reasoning in images based on pre-attention mechanisms and fuzzy attribute graphs. European Conference on Artificial Intelligence, ECAI, Patras, Greece 178:611–615

Fouquier G, Atif J, Bloch I (2012) Sequential model-based segmentation and recognition of image structures driven by visual features and spatial relations. Comput Vis Image Underst 116(1):146–165

Frucci M, Perner P, Sanniti di Baja G (2008) Case-based-reasoning for image segmentation. Int J Pattern Recognit Artif Intell 22(5):829–842

Garbay C (2001) Architectures logicielles et contrôle dans les systèmes de vision. In: Jolion JM (ed) Les systèmes de Vision (Chap. 7). Hermès, Paris, pp 197–252

Gruber TR (1993) Towards principles for the design of ontologies used for knowledge sharing. In: Guarino N, Poli R (eds) Formal ontology in conceptual analysis and knowledge representation. Kluwer Academic Publishers, Deventer. http://citeseer.ist.psu.edu/gruber93toward.html

Guillot A, Meyer JA (2008) La Bionique: quand la science s'inspire de la nature. Dunod, collection UniverSciences

Gurevich IB, Salvetti O, Trusova YO (2009) Fundamental concepts and elements of image analysis ontology. Pattern Recognit Image Anal 19(4):603–611

Hanson AR, Rieseman EM (1978) Visions: a computer system for interpreting scenes. Academic, New York, pp 303–333

Harnad S (1990) The symbol grounding problem. Physica 42:335–346

Hasboun D (2005) Neuranat. http://www.chups.jussieu.fr/ext/neuranat/index.html

Hudelot C (2005) Towards a cognitive vision platform for semantic image interpretation; application to the recognition of biological organisms. PhD in computer science, Université de Nice Sophia Antipolis

Hudelot C, Atif J, Bloch I (2008) Fuzzy spatial relation ontology for image interpretation. Fuzzy Sets Syst 159(15):1929–1951

Hudelot C, Atif J, Bloch I (2010) Integrating bipolar fuzzy mathematical morphology in description logics for spatial reasoning. In: ECAI 2010, Lisbon, Portugal, pp 497–502

Hunter J (2001) Adding multimedia to the semantic web - building an MPEG-7 ontology. International Semantic Web Working Symposium (SWWS), Stanford, CA, pp 261–281

Ijspeert JA, Nakanishi J, Schaal S (2002) Movement imitation with nonlinear dynamical systems in humanoid robots. In: International Conference on Robotics and Automation (ICRA)

Kruse T, Kirsch A, Sisbot EA, Alami R (2010) Exploiting human cooperation in human-centered robot navigation. In: International Symposium in Robot and Human Interactive Communication (IEEE ROMAN), pp 212–217

Kuipers BJ, Levitt TS (1988) Navigation and mapping in large-scale space. AI Mag 9(2):25–43

Lansky A, Friedman M, Getoor L, Schmidler S, Short N Jr (1995) The collage/khoros links: planning for image processing tasks. AAAI Spring Symposium: Integrated Planning Applications. Menlo Park, CA, pp 67–76

Le Ber F, Napoli A (2002) The design of an object-based system for representing and classifying spatial structures and relations. J Univers Comput Sci 8(8):751–773

LeCun Y, Bengio J, Hinton GE (2015) Deep learning. Nature 521(7553):436–444

Leibe B, Ettlin A, Schiele B (2008) Learning semantic object parts for object categorization. Image Vis Comput 26(1):15–26

Lesaint F, Sigaud O, Flagel SB, Robinson TE, Khamassi M (2014) Modelling individual differences in the form of pavlovian conditioned approach responses: a dual learning systems approach with factored representations. PLoS Comput Biol 10(2):e1003466

Liedtke C, Blömer A (1992) Architecture of the knowledge based configuration system for image analysis “Conny”. In: IEEE International Conference on Pattern Recognition (ICPR), The Hague, Netherlands, pp 375–378

Ligozat G (1998) Reasoning about cardinal directions. J Vis Lang Comput 9:23–44

Lungarella M, Metta G, Pfeifer R, Sandini G (2003) Developmental robotics: a survey. Connect Sci 15:151–190

Maillot N, Thonnat M (2008) Ontology-based complex object recognition. Image Vis Comput 26(1):102–113

Martin V, Maillot N, Thonnat M (2006) A Learning approach for adaptive image segmentation. In: 4th IEEE International Conference on Computer Vision Systems (ICVS), New York, pp 40–47

Matsuyama T (1986) Knowledge-based aerial image understanding systems and expert systems for image processing. In: International Geoscience and Remote Sensing Symposium (Zurich), pp 1026–1038

Matsuyama T (1989) Expert systems for image processing: knowledge-based composition of image analysis processes. Comput Vis Graph Image Process 48(1):22–49

Matsuyama T, Hwang VSS (1990) SIGMA: a knowledge-based aerial image understanding system. Plenum, New York

McKeown DM, Harvey WA, McDermott J (1985) Rule-based interpretation of aerial imagery. IEEE Trans Pattern Anal Mach Intell PAMI 7(5):570–585

Mezaris V, Kompatsiaris I (2004) Strintzis MG (2004) Region-based image retrieval using an object ontology and relevance feedback. Eurasip J Appl Signal Process 6:886–901

Minsky M (1974) A framework for representing knowledge. In: Winston P (ed) The psychology of computer vision. McGraw Hill, New York

Najar A, Sigaud O, Chetouani M (2015) Social-task learning for HRI. Social robotics. Springer, Cham, pp 472–481

Nazif A, Levine M (1984) Low level image segmentation: an expert system. IEEE Trans Pattern Anal Mach Intell 6(5):555–577

Nempont O, Atif J, Bloch I (2013) A constraint propagation approach to structural model based image segmentation and recognition. Inf Sci 246:1–27

Neumann B, Möller R (2008) On scene interpretation with description logics. Image Vis Comput 26(1):82–110

Oudeyer PY, Kaplan F, Hafner V (2007) Intrinsic motivation systems for autonomous mental development. IEEE Trans Evol Comput 11(2):265–286

Pasqui V, Saint-Bauzel L, Sigaud O (2010) Characterization of a least effort user-centered trajectory for sit-to-stand assistance user-centered trajectory for sit-to-stand assistance. In: IUTAM Symposium on Dynamics Modeling and Interaction Control in Virtual and Real Environments

Perchant A, Bloch I (2002) Fuzzy morphisms between graphs. Fuzzy Sets Syst 128(2):149–168

Perner P, Holt A, Richter M (2005) Image processing in case-based reasoning. Knowl Eng Rev 20(3):311–314

Protire A, Sapiro G (2007) Interactive image segmentation via adaptive weighted distances. IEEE Trans Image Process 16(4):1046–1057

Quillian M (1967) Word concepts: a theory and simulation of some basic semantic capabilities. Behav Sci 12(5):410–430

Quinton JC, Buisson JC, Perotto F (2008) Anticipative coordinated cognitive processes for interactivist and Piagetian theories. In: Wang P, Goertzel B, Franklin S (eds) 1st Conference on Artificial General Intelligence. Frontiers in Artificial Intelligence and Applications vol 171. IOS Press, Memphis, pp 287–298

Rao A, Lohse G (1993) Towards a texture naming system: identifying relevant dimensions of texture. 4th IEEE Conference of Visualization, San Jose, CA, pp 220–227

Ratliff N, Silver D, Bagnell J (2009) Learning to search: functional gradient techniques for imitation learning. Auton Robot 27(1):25–53

Renouf A, Clouard R, Revenu M (2007) How to formulate image processing applications? International Conference on Computer Vision Systems (ICVS). Bielefeld, Germany, pp 1–10

Rosse C, Mejino J (2003) A reference ontology for bioinformatics: the foundational model of anatomy. J Biomed Inform 36(6):478–500
Rost U, Mnkel H (1998) Knowledge-based configuration of image processing algorithms. International Conference on Computational Intelligence and Multimedia Applications (ICCIMA). Monash, Australia, pp 9–11
Saathoff C, Staab S (2008) Exploiting spatial context in image region labelling using fuzzy constraint reasoning. In: Ninth International Workshop on Image Analysis for Multimedia Interactive Services WIAMIS'08, pp 16–19
Saint-Bauzel L, Pasqui V, Monteil I (2009) A reactive robotized interface for lower limb rehabilitation: clinical results. IEEE Trans Robot Spec Issue Rehabil Robot 25:583–592
Salaun C, Padois V, Sigaud O (2010) Learning forward models for the operational space control of redundant robots. In: Peters J, Sigaud O (eds) From Motor Learning to Interaction Learning in Robots, vol 264. Springer, Berlin, pp 169–192. https://doi.org/10.1007/978-3-642-05181-4_8
Schaal S (1999) Is imitation learning the route to humanoid robots? Trends Cogn Sci 6:233–242
Shafer G (1976) A mathematical theory of evidence. Princeton University Press, Princeton
Sigaud O, Droniou A (2016) Towards deep developmental learning. IEEE Trans Cogn Dev Syst 8(2):99–114. https://doi.org/10.1109/TAMD.2015.2496248
Smeulders A, Worring M, Santini S, Gupta A, Jain R (2000) Content-based image retrieval at the end of the early years. IEEE Trans Pattern Anal Mach Intell 22(12):1349–1380
Sowa J (1984) Conceptual graphs: information processing in mind and machine. Addison Wesley, Reading, p 234
Stulp F, Sigaud O (2012) Path integral policy improvement with covariance matrix adaptation. 29th International Conference on Machine Learning (ICML). Edinburgh, Scotland, pp 1–8
Stulp F, Sigaud O (2015) Many regression algorithms, one unified model: a review. Neural Netw 69:60–79
Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. MIT Press, Cambridge
Thonnat M, Moisan S (2000) What can program supervision do for program reuse? IEEE Proc Softw 147(5):179–185
Town C (2006) Ontological inference for image and video analysis. Mach Vis Appl 17(2):94–115, www.cl.cam.ac.uk/~cpt23/papers/TownMVA2006.pdf
Vanegas MC, Bloch I, Inglada J (2016) Fuzzy constraint satisfaction problem for model-based image interpretation. Fuzzy Sets Syst 286:1–29
Vieu L (1997) Spatial representation and reasoning in artificial intelligence. In: Stock O (ed) Spatial and temporal reasoning. Kluwer, Dordrecht, pp 5–41
Vijayakumar S, D'Souza A, Schaal S (2005) Incremental online learning in high dimensions. Neural Comput 12:2602–2634
Vinyals O, Toshev A, Bengio S, Erhan D (2015) Show and tell: a neural image caption generator. In: IEEE Conference on Computer Vision and Pattern Recognition, pp 3156–3164
Waxman S (2000) Correlative neuroanatomy, 24th edn. McGraw-Hill, New York
Zadeh LA (1965) Fuzzy sets. Inf Control 8:338–353