Information, data and knowledge in image understanding

Isabelle Bloch

LTCI, Télécom Paris

isabelle.bloch@telecom-paris.fr

What is image understanding?

From the 1960's to today:

- Miller and Shaw (1968): survey of linguistic methods for picture processing, defined as analysis and generation of pictures by computers, with or without human interaction.
- Clowes (1971): linguistic approach for picture interpretation (pattern description language).
- Reiter (1989): interpretation = logical model of sets of axioms.
- Ralescu (1995): image understanding = verbal description of the image contents.
- Bateman (2010): needs for a semantic layer for spatial language.
- Xu et al. (2014): image interpretation = assigning labels or semantics representations to regions of a scene.

Introduction

What is image understanding?

Here:

- Beyond individual object recognition.
- Objects in their context, spatial arrangement.
- Global scene interpretation.
- Semantics extraction.
- Providing verbal descriptions of image content.
- Dynamic scenes: recognition and description of actions, gestures, emotions..
- Inference, higher level reasoning.

Important role of Artificial Intelligence.

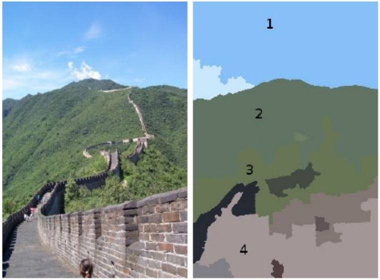
A few examples

A lot of work on image annotation: object $\rightarrow$ several objects $\rightarrow$ scene.

Magritte, 1928

A few examples

A lot of work on image annotation: object → several objects → scene.



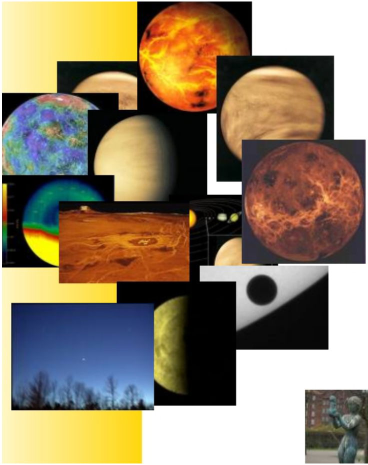
Millet et al., 2005

(rules, spatial reasoning...)

Region	without spatial relations	with spatial relations
1	sky	sky
2	grass	tree
3	tree	tree
4	building	building

A few examples

A lot of work on image annotation: object $\rightarrow$ several objects $\rightarrow$ scene.

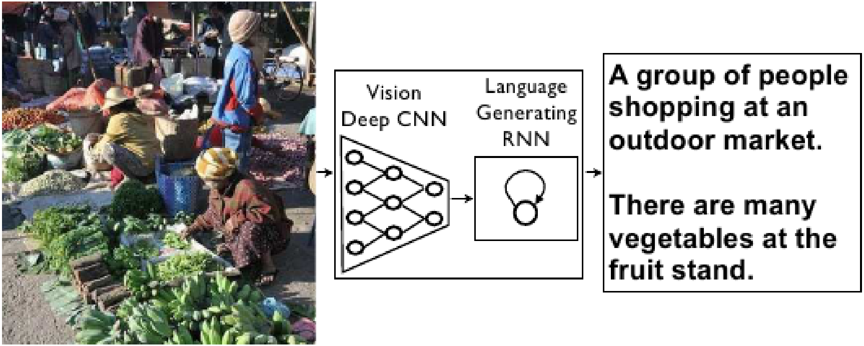


Venus?

A few examples

A lot of work on image annotation: object → several objects → scene.

"Show and tell":



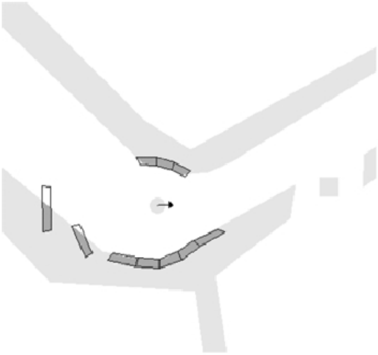
Vinyals et al., 2015
(convolutional neural networks, deep learning)

A few examples

A lot of work on image annotation: object → several objects → scene.

Fig. 1. Our system automatically generates the following descriptive text for this example image: *"This picture shows one person, one grass, one chair, and one potted plant. The person is near the green grass, and in the chair. The green grass is by the chair, and near the potted plant."*

Kulkarni et al., 2013



COARSE DESCRIPTIONS:

There are objects behind the robot.
An object is on the left of the robot.
An object is on the right of the robot.

DETAILED DESCRIPTIONS:

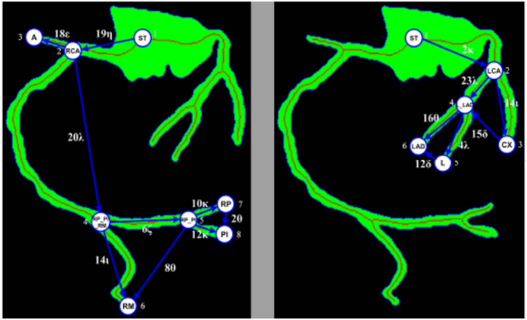
An object is on the left of the robot, but extends forward relative to the robot (the description is satisfactory). The object is very close to the robot.

An object is behind the robot (the description is satisfactory). The object is close to the robot.

An object is mostly behind the robot, but somewhat to the right (the description is satisfactory). The object is close to the robot.

An object is on the right of the robot, but extends forward relative to the robot (the description is satisfactory). The object is very close to the robot.

Skubic et al., 2003
(fuzzy modeling of spatial relations)



Ogiela et al. 2002, Trzupek et al. 2010
(graphs and grammars)

The patient 489478 presents a "Thoraciclumbar" spine curvature pattern with a matching degree of 1. The spine includes the curves:

- a "Cervical" curve LEFT oriented with 15,6 degrees between C2 and T2 and with apex in C6.

- a "Thoracic" curve RIGHT oriented with 21,5 degrees between T7 and L3 and with apex in T12

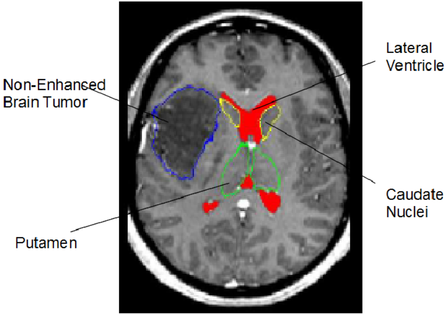
The patient 526257 presents a "Double Thoracic" spine curvature pattern with a matching degree of 1. The spine includes the curves:

- a "Cervical Thoracic" curve RIGHT oriented with 28.5 degrees between T2 and T7 and with apex in T5.

- a "Thoracic" curve LEFT oriented with 39.6 degrees between T7 and L1 and with apex in T10

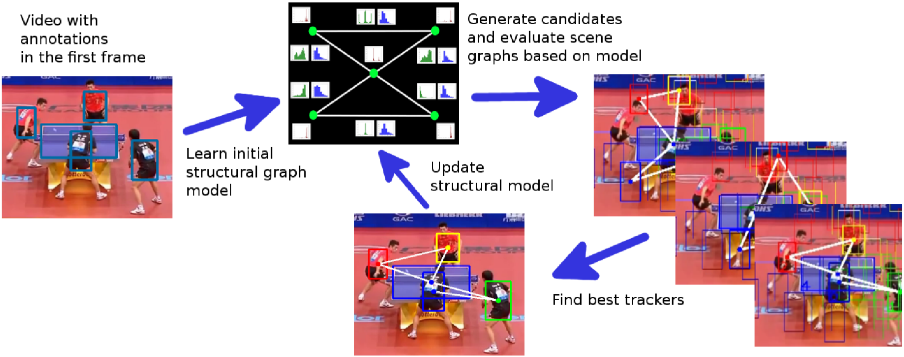
- a "Thoracic Lumbar" curve RIGHT oriented with 28.8 degrees between L1 and L5 with apex in L3

Trivino et al., 2010
(fuzzy rules)



- An abnormal structure is present in the brain.
- A peripheral non-enhanced tumor is present in the left hemisphere.

Atif et al., 2014
(spatial reasoning, abduction)



Morimitsu et al., 2015
(graphs, Bayesian tracking, hidden Markov models)

Data vs. knowledge

Is everything in the data?

- Powerful methods and impressive results.
- Accessibility of data.
- Size and number of data.
- Cost of learning.

Importance of knowledge.

Information processing

Information
=
element that can be encoded in order to be stored, processed or communicated

Real or virtual worlds (ex: preferences)

Generic notion

- Knowledge (classes of objects)
- Data (cases, facts, particular objects)

Types of d'information:

- particular, determined situation (data, facts)
  Lea is 25 years old
- undetermined situation whose existence or some properties can be claimed
  There exist persons older than 100 years
- several particular situations (ex: statistical data, prototypes)
- classes of situations (constraints, generic rules, knowledge)

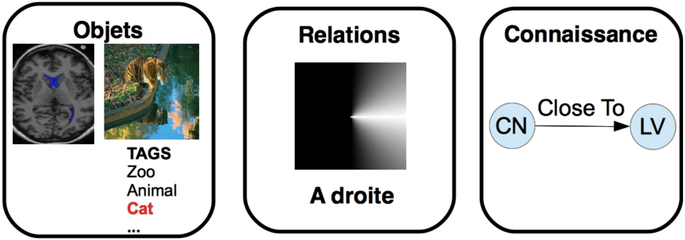
Types of imperfection:

- ambiguity
- bias, noise
- incompleteness
- imprecision, uncertainty
- inconsistency and conflict

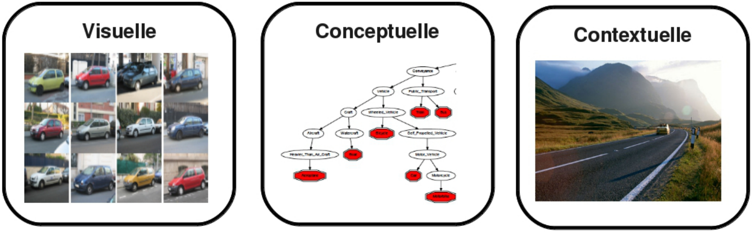
Objectives of information processing:

- representing information
  - preparing, improving
  - analysis, highlighting important elements
  - synthesis (approximate, simplified or structured description)
- storing, retrieving, eliciting information
- exploit information for deciding and acting
  - fusion, multi-criteria decision making, preferences modeling
  - constraint satisfaction, planning
  - dynamic system control, robotics
  - scene understanding
- communicating information

Imperfect information, multiple nature of information



Voiture



Uncertainty and imprecision in image processing and understanding

due to:

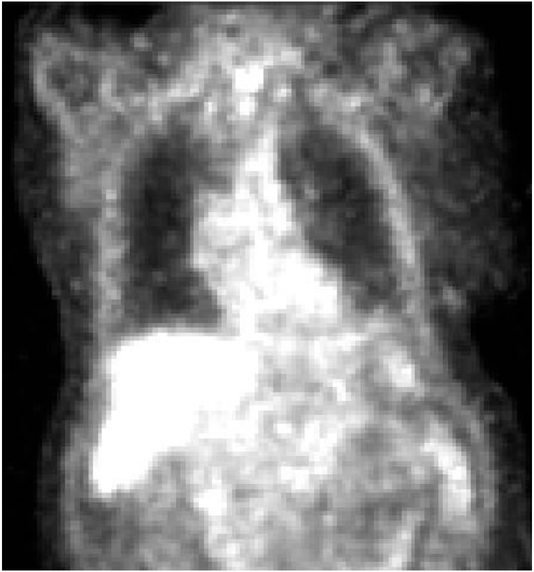
- observed phenomenon
- sensor limitations
- reconstruction algorithms
- noise
- limited reliability
- representations
- processing
- knowledge and concepts

$\Rightarrow$ uncertain, imprecise, partial, ambiguous, biased, conflictual information

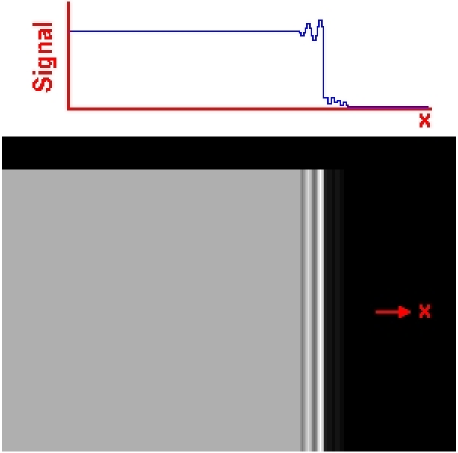
$+$ evolves (dynamic world)

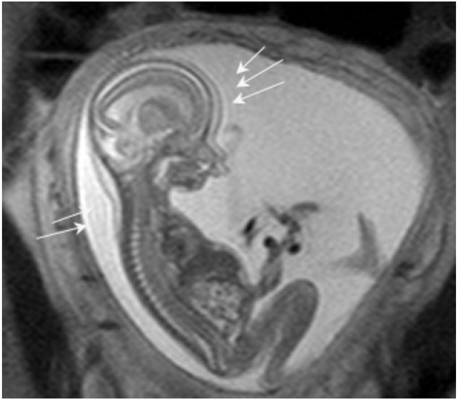
Uncertainty and imprecision in image processing and understanding: examples

Uncertainty and imprecision in image processing and understanding: examples

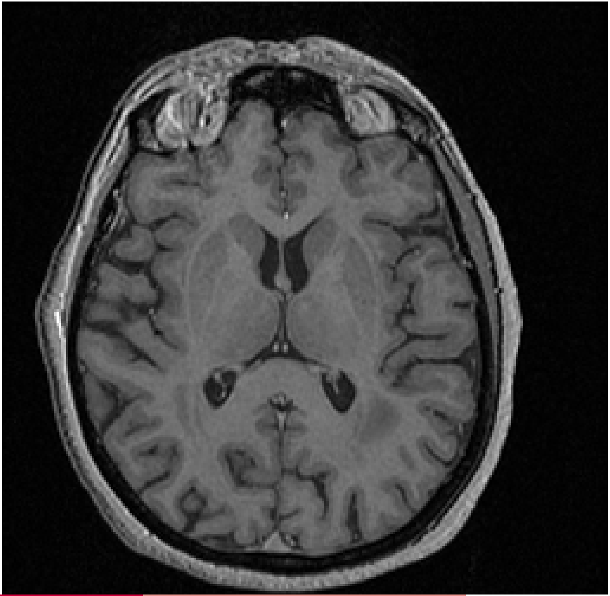


Uncertainty and imprecision in image processing and understanding: examples

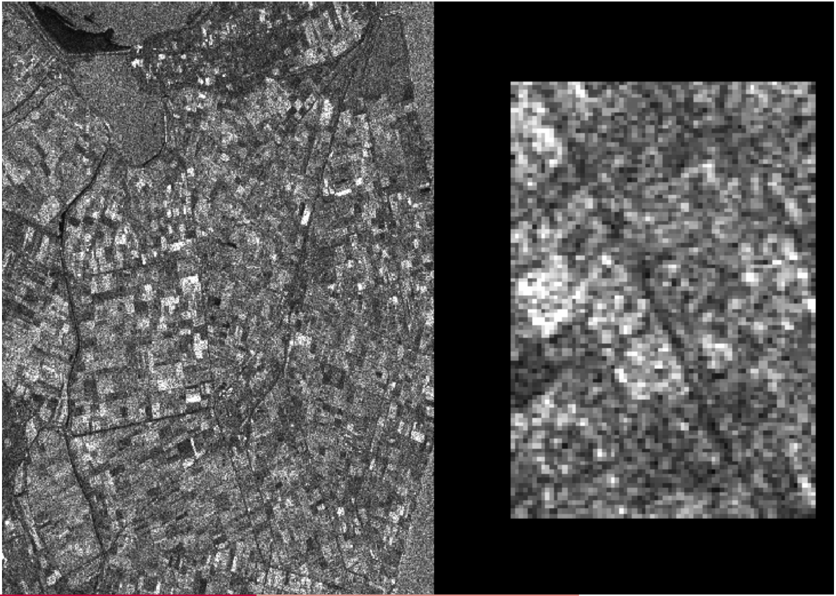




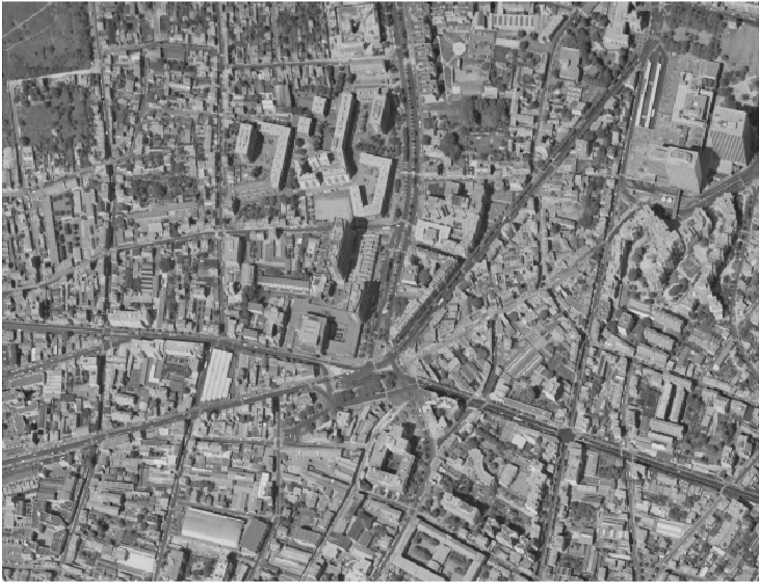
Uncertainty and imprecision in image processing and understanding: examples



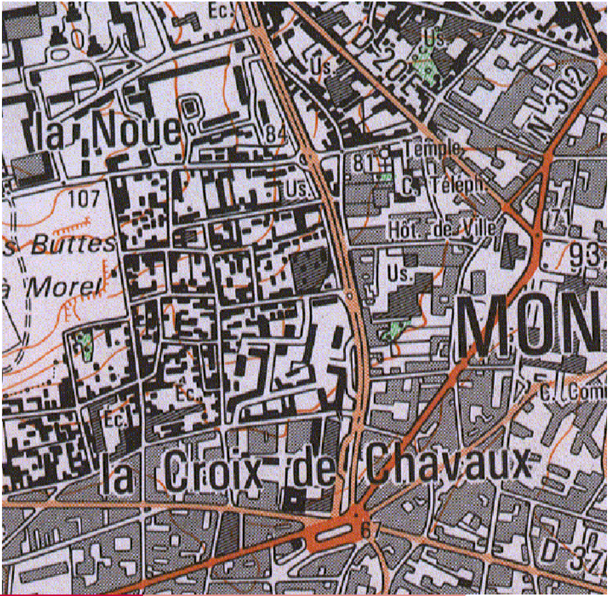
Uncertainty and imprecision in image processing and understanding: examples



Uncertainty and imprecision in image processing and understanding: examples



Uncertainty and imprecision in image processing and understanding: examples



Uncertainty and imprecision in image processing and understanding: examples

Uncertainty and imprecision in image processing and understanding: examples

[morphologie]

morphologie fonctionnel atlas animations vidéo reconstruction 3D liens

embryologie
histologie
crâne
méninges
moelle épinière
tronc cérébral
nerfs crâniens
cervelet
cortex cérébral
noyaux gris
substance blanche
v3
v4
ventricules latéraux

auteur : D Hasboun
illustrations : H Fournié

Noyau caudé

Le noyau caudé est un noyaux gris central, en forme de **fer à cheval** ouvert en avant.
Il s'enroule **autour du thalamus**, repose sur sa face supérieure, puis descend en arrière du thalamus.

Il se continue en avant dans le lobe temporal. Il présente d'avant en arrière :

–une **tête** volumineuse située en dehors de la corne frontale du ventricule latéral

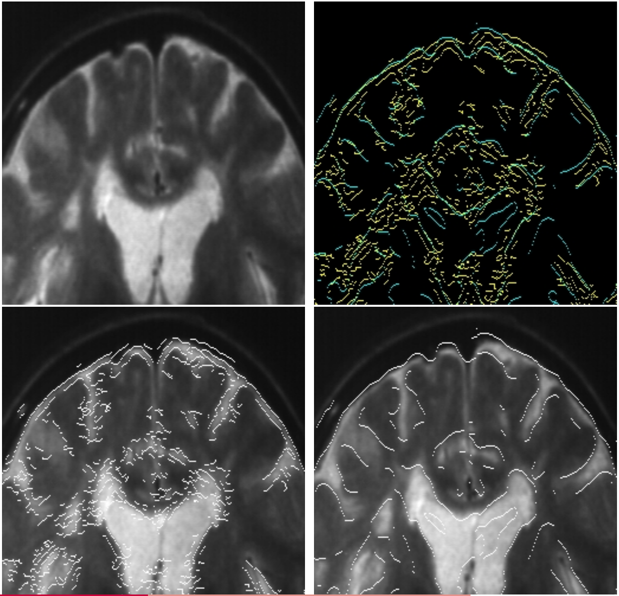
–un **corps** dont le volume diminue d'avant en arrière. Il repose sur le thalamus puis descend en arrière du pulvinar.

–la **queue** du noyau cadé chemine dans le lobe temporal, au–dessus de la corne temporale du ventricule latéral, dans la région sous lenticulaire.

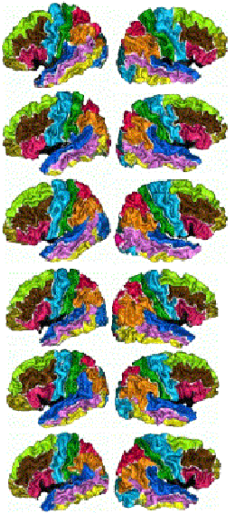
Le noyau caudé constitue le **néostriatum** avec le putamen.

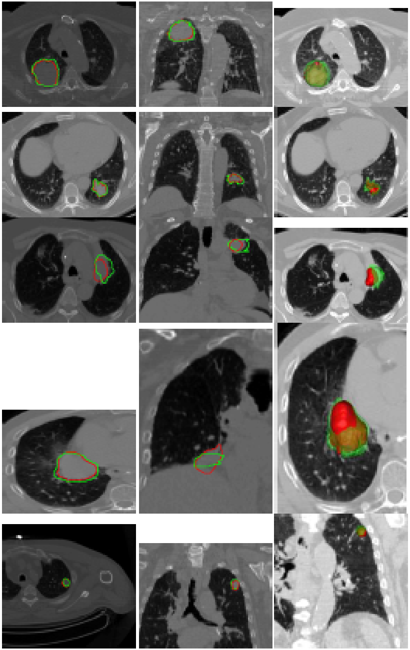
accueil - morphologie - fonctionnel - animations - vidéo - atlas - 3D - liens - préface - bibliographie - aide

Uncertainty and imprecision in image processing and understanding: examples



Variability





Which solutions?

- Eliminating imperfection

  ex: improving sensors

- Tolerating imperfection

  ex: robust programs, able to repare

- Reasoning under imperfection (considered as part of information)
  - modeling
  - approximate reasoning
  - meta-knowledge

Numerical representations of imperfect information

Main theories:

- probabilities and statistics
- belief functions
- fuzzy sets and possibility theory

But!

- do not model the same types of imperfection
- different semantics
- different representation power
- different reasoning power

Symbolic representations

▷ AI-based approaches:

- logics and spatial logics
- knowledge-based systems
- ontologies
- semantic networks
- ...

Basic types of reasoning

- deduction: consequences from facts

$$\frac{A \rightarrow B,\ A}{B}$$

- contraposition: non-observations

$$\frac{A \rightarrow B,\ \neg B}{\neg A}$$

- abduction: causes explaining observations

$$A \rightarrow B,\ B \quad \text{infer } A$$

- induction: rules from regular observations

$$\frac{B \text{ whenever } A}{A \rightarrow B}$$

- projection: consequences from actions

$$\frac{A \rightarrow B,\ \text{ do } A}{\text{expect } B}$$

- planning: actions from goals

$$\frac{A \rightarrow B,\ \text{ want } B}{\text{do} A}$$

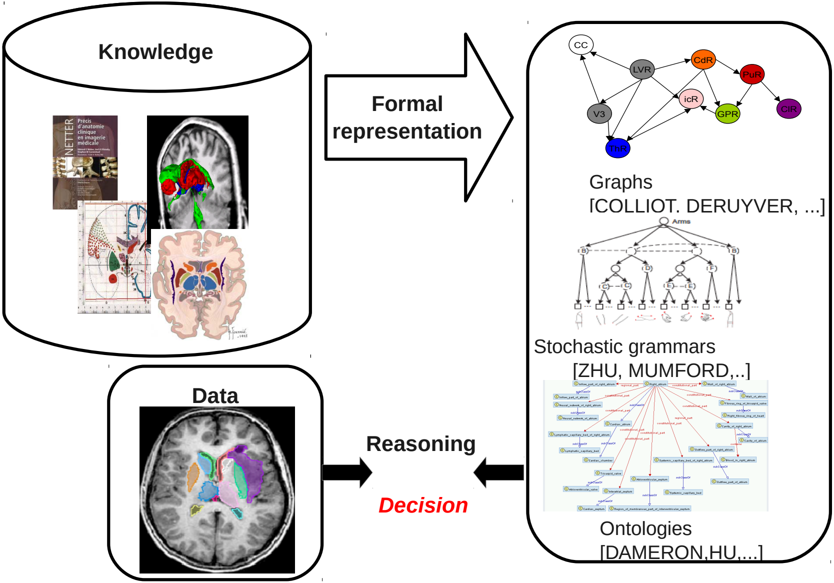
monotonic / non-monotonic logics

Models for image understanding

- Develop mathematical models to represent
  - knowledge (context, expert, spatial organization...),
  - information contained in images (geometry, statistics, shape, appearance...),
  - and to combine them (fusion process),

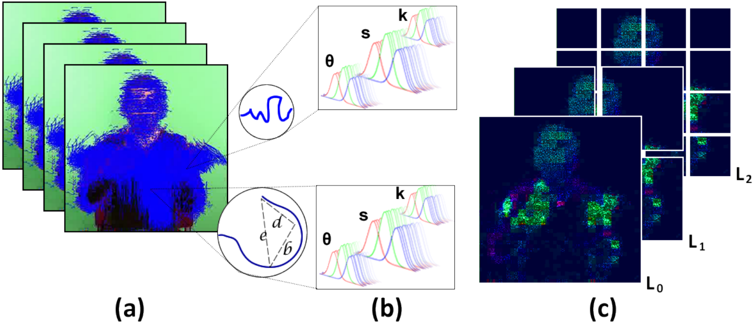
$\Rightarrow$ operational and efficient algorithms for image understanding

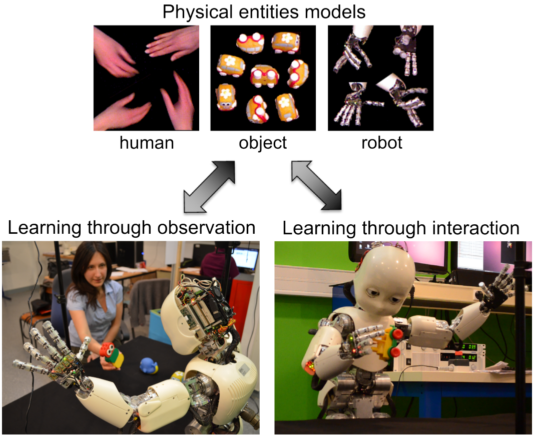
- Semantic gap.
- Pathological or unexpected cases.

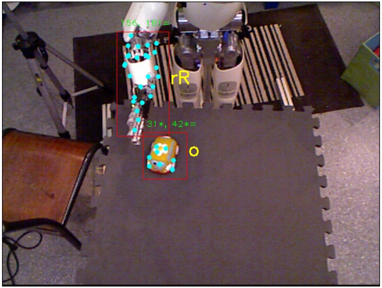

Knowledge
Formal representation
CC
LVR
CdR
PuR
icR
V3
GPR
CIR
ThR
Graphs
[COLLIOT, DERUYVER, ...]
Arms
Stochastic grammars
[ZHU, MUMFORD,..]
Ontologies
[DAMERON,HU,...]
Data
Reasoning
Decision

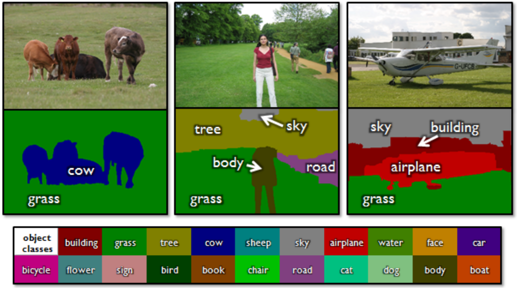
Semantic gap

- Symbol grounding = "How is symbol meaning to be grounded in something other than just more meaningless symbols?" (Harnad).
- Anchoring = "creating and maintaining the correspondence between symbols and sensor data that refer to the same physical object" (Saffiotti & Coradeschi).
- Semantic gap = "lack of coincidence between the information that one can extract from the visual data and the interpretation of these data by a user in a given situation" (Smeulders).

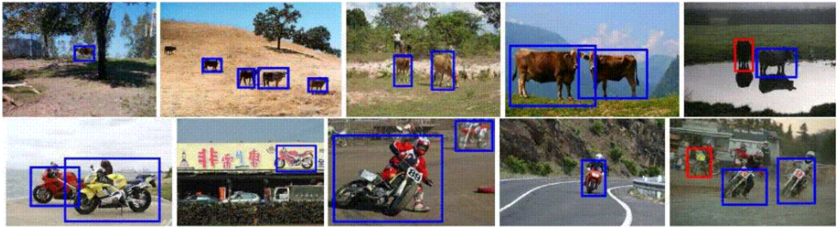

θ
s
k
d
e
b
L2
L1
L0
(a)
(b)
(c)


Physical entities models
human
object
robot
Learning through observation
Learning through interaction


rR
31*, 42*=
O



[Shotton 09]



[INRIA-Willow 08]

Zones résidentielles

Avions

Cuves industrielles

Gare de triage

Échangeurs routiers