How to formulate image processing applications ?

Arnaud Renouf, Régis Clouard, and Marinette Revenu

GREYC Laboratory (CNRS UMR 6072)
6, Boulevard Maréchal Juin
14050 CAEN cedex FRANCE
Arnaud.Renouf@greyc.ensicaen.fr

Abstract. This paper presents an ontology-based system for helping inexperienced users formulate image processing applications. We propose models and their formalization through ontologies that identify and organize the necessary and sufficient information to design such applications. We also explain the interaction means that lead the user to give a good formulation of his/her application.

1 Introduction

In the last fifty years, a lot of image processing applications have been developed in many fields (medicine, geography, robotic, industrial vision, ...). Nevertheless, the problem of formulating such applications in image processing terms has been little studied. Image processing specialists develop their applications by trial-error cycles and do not define a real formulation of the problem they consider. Thus, applications are merely formulated through their solution and not by a formal definition of the problem they solve.

However a complete and rigorous formulation of the application is essential towards the goal of designing more robust and more reliable vision systems, and in order to fix the limits of the application, to favor reusability and to enhance the evaluation. Such a formulation has to clearly specify the objective and to identify the range of the considered input data. Unfortunately, formulating an image processing application is a problem of qualitative nature that relies on subjective choices. Hence, an exhaustive or exact definition of the problem does not exist. Only an approximative characterization of the desired application behavior can be defined [1].

Our approach tends to acquire the user's business knowledge and map it to an image processing one in order to find a solution. Hence, we propose to study the formulation of image processing applications through the design of a system that allows inexperienced users to build their applications by interaction. In this paper, we present a brief state of the art on image processing application formulation (Section 2) and describe the global system we are developing (Section 3). Then we propose a model for the formulation of image processing applications and its formalization through an ontology (Section 4). Finally, the interaction part of the system is explained (Section 5).

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8

2 State of the Art

Attempts to formulate image processing applications were conducted around knowledge based systems such as LLVE [2], CONNY [3], OCAPI [4], MVP [5] or BORG [6]. A priori knowledge on the application context (sensor effects, noise type, lighting conditions, ...) and on the goal to achieve were more or less implicitly encoded in the knowledge base. The formulation of the application is reduced to the choice of a transformation objective in a list of predefined image processing tasks which are well known. Therefore they make the saving of a complete and explicit formulation of the image processing application.

Among these works, an original approach was proposed by J. Hasegawa et al. [7]. The formulation is given by the user through a sample figure as a sketch. This kind of formulation is interesting because it is objective and its nature is quantitative. Nevertheless its scope is limited to a restrictive subset of segmentation or detection goals and it does not really take into account the input images variability.

More recent approaches bring more explicit modelling [8][9][10][11] but they are all restricted to the modelling of business objects description for tasks such as detection, segmentation, image retrieval, image annotation or recognition applications. They use ontologies that provide the concepts needed for this description (a visual concept ontology for object recognition in [9], a visual descriptor ontology for semantic annotation of images and videos in [12] or image processing primitives in [8]) or they capture the business knowledge through meetings with the specialists (e.g. use of the NIAM/ORM method in [10] to collect and map the business knowledge to the vision knowledge). But they do not completely tackle the problem of the application context description (or briefly as in [9]) and the effect of this context on the images (environment, lighting, sensor, image format). Moreover they do not define the means to describe the image content when objects are a priori unknown or unusable (e.g. in robotic, image retrieval or restoration applications). They also suppose that the objective is unique and known beforehand (to detect, extract or recognize an object with a restrictive set of constraints) and therefore they do not address its formulation.

3 A System to Formulate Image Processing Applications

Our work is part of a global project which aims at developing a system that automatically generates image processing softwares from user-defined problem formulations. This system is composed of two sub-systems (Fig. 1): a formulation system for image processing applications which is the focus of our work, and a planning system for image processing tasks [6]. At the moment, the system handles only one image processing task at a time. If a real application needs several tasks, the system is used as many times as there are tasks. It is also restricted to image processing tasks. It considers only image-to-image transformations without any interpretation of the image content. The user defines the problem with the terms of his/her domain by interaction with the user layer of the formulation system. This part of the system is a human-machine interface which uses

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

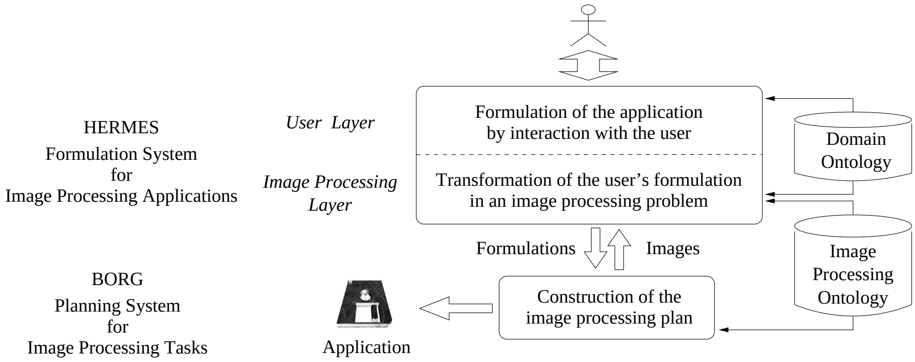

Fig. 1. The system architecture of the project.

a *domain ontology* to handle the information dedicated to the user. It groups concepts that allow the users to formulate their processing intentions and define the image class. Then the formulation system translates this user formulation into image processing terms taken from an *image processing ontology*. It uses a set of rules to ground the symbols given by the user which are terms of his/her domain and concepts of the *domain ontology* to image data which are concepts of the *image processing ontology* (see Section 5.3). This translation realizes the mapping between the phenomenological domain knowledge of the user and the image processing knowledge.

The result of this translation is an image processing request which is sent to the planning system to generate the program that responds to this request. This cooperation needs the two sub-systems to share the *image processing ontology*. Then the formulation system executes the generated program on test images and presents the results to the user for evaluation purposes. The user is responsible for modifying the request. This process is repeated until the results are validated.

4 Formulation Model for Image Processing Applications

We propose a complete model for the formulation of image processing applications [13]. This model allows to identify and organize the relevant information which are necessary and sufficient for the planning system (or an image processing specialist) to develop a convenient solution. It covers all image processing tasks and is independent from any application domain. In this section, we present a brief review of this model formalization through an *image processing ontology* (notice that this ontology tackles the problem of the formulation from the image processing experts point of view). This ontology is composed of two parts: the objectives specification and the image class definition. It is written in OWL DL which is based on description logics, defines 300 concepts, 23 roles and 183 restrictions. During the translation step, the formulation system instantiates this

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

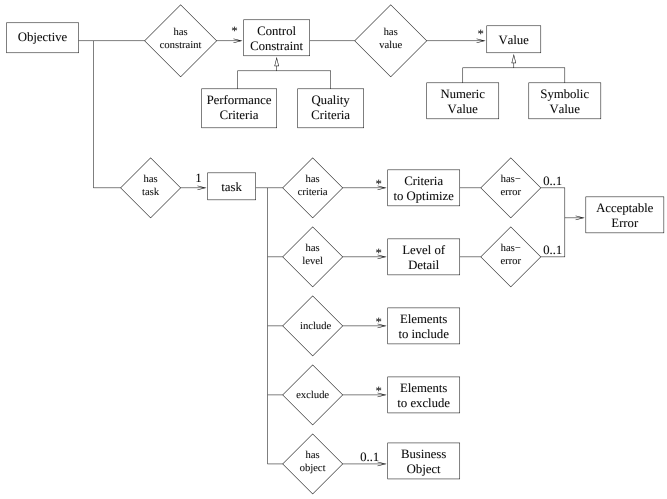

Fig. 2. CML representation of the objective specification model.

ontology and sets their properties using an OWL API in order to represent the formulation of the user at the image processing level.

4.1 The Objective Specification

This specification relies on the teleological hypothesis [13]. This hypothesis leads us to define an image processing application through its finalities. Hence, an application objective is represented by a list of tasks with related constraints (criteria to optimize with acceptable errors, levels of detail with acceptable errors, elements to be excluded and included in the result) and a list of control constraints. We present on Fig. 2 the formalization of this model using the CML formalism of CommonKADS which covers the ontologies specification [14]. This figure represents the highest abstraction level of the objective specification part of the image processing ontology and does not show all its concepts. A concept like *Task* is the root of a hierarchic tree where nodes are processing categories and leaves are effective tasks. Processing categories belong to reconstruction, segmentation, detection, restoration, enhancement and compression objectives. Effective tasks are extract object, detect object, enhance illumination, ...

To illustrate the use of this model, we take the example of a cytological application [15] where the user wants to extract serous cell nuclei. In Table 1, we present the formulation of the application objective which is obtained from the instances (and their properties) created by the formulation system in the *image processing ontology*. The terms (task, extract object, criteria to optimize, ...) are concepts of the ontology and terms in italic are instances created by the system (object1 is an instance of the concept *business object*, num1 is an instance of the concept *interval*).

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

Table 1. Formulation of the application objective for the cytologic application.

<table>
<tr><td rowspan="4">Task: Extract object</td><td colspan="2">Criteria to optimize: Boundaries localization
Acceptable error: Prefer boundaries inside</td></tr>
<tr><td colspan="2">Level of detail: Separate just-touching
Acceptable error: Prefer not to separate</td></tr>
<tr><td colspan="2">Business object: object1 ("Serous Cell Nucleus")</td></tr>
<tr><td colspan="2">Elements to be excluded: Touching borders</td></tr>
<tr><td rowspan="2">Control constraints</td><td>Quality criteria</td><td>Processing time: num1 ([0,2] secs)</td></tr>
<tr><td>Performance criteria</td><td>Application quality: Reliability</td></tr>
</table>

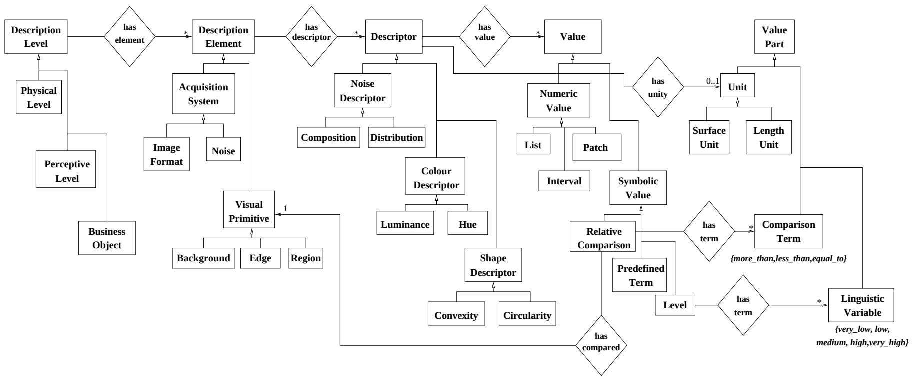

Fig. 3. CML representation of the image class definition model.

4.2 The Image Class Definition

This definition relies on two hypotheses: the semiotic and the phenomenological hypotheses [13]. The semiotic hypothesis leads us to define the image class considering three levels of description:

1. the physical level (i.e. sign level in semiotic) focuses on the characterization of the acquisition system effects on the images.
2. the perceptive level (i.e. signifiant) focuses on the description of the visual primitives (regions, lines, ...) without any reference to business objects.
3. the semantic level (i.e. signified) focuses on the identification of the business objects visualized in the images.

The phenomenological hypothesis states that a visual characterization of the images and the business objects are sufficient to design image processing applications. Hence, we do not need a complete description of the objects given by the user but only a description of its visual manifestation. The system asks the user to describe how an object appears in the images but not what it is. For example, in the case of aerial images, cars are simply defined as homogeneous rectangular regions. We present on Fig. 3 the formalization of this model using the CML representation. Here again, this representation is only a part of the *image processing ontology* and does not show all the concepts: the descriptors

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8

Table 2. A part of the definition of the image class at the physical level.

Element	Descriptors	Values	Range of values
Lighting	Lighting Conditions	{stable, non-uniform}	{under-exposed, over-exposed stable, varying, ...}
	Lighting Model	*lighting.bmp*	image file
Image Format	Width	*512* pixels	$\mathbb{Z}$
	Height	*512* pixels	$\mathbb{Z}$
	Color Space	RGB	{binary, grey level, RGB, ...}
	Quantification	*24* bits per pixel	$\mathbb{Z}$
Noise	Composition	additive	{additive, multiplicative}
	Distribution	Gauss	{Gauss, Poisson, Rayleigh, ...}
	Mean	*0*	$\mathbb{R}$
	Standard Deviation	*[0 , 4]*	$\mathbb{R}^+$

Table 3. A part of the definition of the image class at the semantic level for the business concept *serous cell nucleus*.

Visual Primitives	Descriptors	Values	Range of values
Region	Luminance	*[0 , 0.8]*	[0 , 1]
	Hue	*[9π/8 , 3π/2]*	[0 , 2π]
	Convexity	high,very high	{very low,low,average, high,very high}
Edge	Type	step	{step,roof}
	Contrast	very low,low	{very low,low,average, high,very high}

(color, texture, shape, geometric, photometric descriptors), the visual primitives (region, edge, background, point of interest, points cloud) and the types of values (symbolic and numeric). This part of the ontology is also instantiated by the formulation system to construct the image class definition. In Table 2, we show a representation of this definition at the physical level, and in Table 3 the definition at the semantic level for the business object *Serous Cell Nucleus*. Only the terms in italic result from an instantiation of the system, the others are concepts or individuals that are part of the *image processing ontology*.

5 The System

In this section, we present how the system can lead the user to specify the objective and define the image class definition. We propose an interaction model which is formalized through a *domain ontology*. This ontology is very close to the image processing one: its structure is the same, some concepts are shared (e.g. width, color, size, ...), some are just transformed in other words (the visual primitive "line" in the *domain ontology* is an "edge" in the *image processing ontology*) and some are added with new properties to suit the interaction model requirements.

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8

5.1 The interaction model

The role of the system is to help the user to give a formulation which is:
- compact: it is reduced to essential information;
- representative: it keeps a strong correspondence to the real objects;
- discriminative: it contains enough information to discriminate the objects;
- precise: it allows a good identification of objects;
- dense: all foreseen image content variability of the scene is modelled.

We consider that the user is inexperienced in the image processing field. Hence the system has to propose interfaces that lead this user to construct the formulation using terms and representations easily understandable (these terms are part of the *domain ontology*). Nevertheless, we also consider the user as a specialist of his/her domain. Hence he/she is able to specify the set of tasks to achieve, describe the effects of the acquisition system on the produced images, give a relevant description of the images and objects visualized, and evaluate the results of the processing plan produced. Therefore, he/she is responsible for giving a formulation which is both representative and discriminative.

Consequently, the role of the formulation system is to help users to produce a formulation that is compact, dense and precise. We know that the formulation is essentially of qualitative nature because it results from subjective user's choices. To make these choices as most objective as possible, we propose an interaction model which leans on different principles:
- the specification of the objective is initiated from the post-processing tasks;
- the definition of the image class at the physical level is guided by the characterization of the acquisition system;
- the definition of the image class at the semantic level is based on the construction of a business objects tree.

5.2 The Interaction Means

The Objective Specification: The system firstly proposes to specify the image processing objectives from the specification of the post-processing tasks. These tasks are not part of the formulation of the image processing application but they are used as a mean of interaction (part of the *domain ontology*). They give information on the way the results of the application are going to be used (size or shape, position, orientation, radiometric, topological measurements, artistic visualization, visualization of details, counting) and thus, on its desired behavior. For example, if one wants to perform size measurements on the regions extracted by a segmentation task, one can deduce that a criteria to optimize will be the boundaries localization. Moreover, if one wishes to achieve other measurements like radiometric ones on the pixels color for example, one can deduce that an acceptable error on the boundaries localization will be to place the border inside the region (to avoid taking background pixels in the measurements).

With such a specification, in a first step the user chooses the post-processing tasks; then he/she specifies the task to be performed to reach the objective. The interface proposes a list of processing tasks which are part of the *domain*

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8

ontology. Thanks to the post-processing specification, the system is now able to propose default values for the constraints associated to the task (by realizing inferences using first order production rules). However, in fine, it is always the user that defines the values of these constraints, either by validating the default values or by setting them directly.

The Image Class Definition: The system helps the user in the definition of the image class. At the physical level, a database of classical acquisition systems (CCD camera, MRI, PET, Infrared, ...) allows to propose default values for noise or defects description. We know for example that images produced by ultrasound sensors are perturbed by speckle noise. Such a database can easily be updated by image processing specialists. Nevertheless, we consider that the user, as a specialist of his/her domain, can also directly give the effects produced on the images (often the case in fields like astronomy or robotics).

At the semantic level, the description is initiated by the construction of a business objects tree. This choice leans on the hypothesis that domain specialists share a taxonomy of their objects. Thus, they are able to identify and organize them hierarchically using inheritance relations. This first step aims at representing their differences like in the differential paradigm proposed by B. Bachimont for the design of ontologies [16]. Then, they specify the composition relations that links these objects and describe their spatial organization. According to the phenomenological hypotheses, a new object has to be created for each way it can appear in the images. These information obtained by interaction with the user give a part of objectivity in the user's formulation and leads him/her to give a dense representation of the objects. Based on the tree, the system asks the user to give a phenomenological characterization of the objects from the visual primitives that compose them.

When the task is a global one (e.g. enhance illumination) or when the objects are a priori unknown, the description is completed at the perceptive level. Hence the user describes "syntaxically" the images content with visual primitives chosen in the *domain ontology* (contour, region, line, ...).

5.3 The Anchoring Problem

Once a user formulation has been done, the system has to translate it into an image processing formulation. The terms of the user's domain (which he/she uses to qualify the objects) have to be associated to image data or to terms of the image processing field (concepts of the *image processing ontology*). This association is a problem as such and known as the symbol grounding or anchoring problem [17][18]. We explain here how the system handles this problem for the different modalities of description given to the users.

The Anchoring of the User's Domain Terms: One of the problem to be tackled with visual descriptions is the users' subjectivity. A color for example is not perceived in the same way by different users: for example a same color can be described as rather blue or rather green by different persons. Hence we decided that the users have to name themselves the colors. Then they have to give values to the associated descriptors either by setting the values of the descriptors (e.g.

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8
This document and other contributions archived and available at: http://biecoll.ub.uni-bielefeld.de

hue, saturation, luminance) or by giving patches that surround image areas as examples and counter-examples of the given description. While other systems calculates a set of values from these patches [17], our system sends directly the patches to the planning system and let it extract by itself the needed values. Actually, we do not aim at solving completely the symbol grounding problem with the formulation system but we give the planning system the means to do it.

The Linguistic Variables: Linguistic variables allow the user to keep imprecision on descriptors that characterize imprecise nature properties (e.g. circularity, compactness, ...) [17]. This is also very useful when the user wants to introduce an important variability on features of the image class. The user chooses a term in a predefined list (e.g. null, very low, low, average, high, very high) to give a value to a descriptor. These values are kept in the formulation at the image processing level because resolution can be done from these linguistic variables. They can also be associated with patches as described previously.

The Comparison Values: The user can also choose to valuate a descriptor by comparison. This is very useful in the image processing field to express that some features are used directly to discriminate two elements (object or visual primitives already described). Two types of comparison are possible: relative ones (e.g. the size of A is greater than the size of B) or superlative ones (e.g. the size of the background is the greatest).

We need the two types of values (the symbolic and numerical values) for the formulation of the image processing application. The symbolic values (linguistic variables, ontology terms, comparison values) are necessary in the first step of the resolution which is the construction of the processing plan. Then numerical values (image patches, image file, descriptor values) are used to tune the parameters of the operators selected in the previous step of the resolution.

6 Conclusion

Our formulation model for image processing applications identifies and organizes the necessary and sufficient information for a specialist, or an automatic resolving system, to develop a solution. The formalization of this model through an ontology specifies the grammar and the vocabulary of the formulation.

We defined interaction means that lead the user to give a compact, dense and precise formulation by trying, on one hand, to reduce the semantic gap between business knowledge and image processing knowledge and, on the other hand, to objectify as much as possible his/her formulation.

This work is a contribution to the image processing field because the formalization of the problems allows to give prominence to the knowledge used in such applications development. It defines a guideline on the way to tackle such applications and defines the formulation elements of image processing problems. The explicitness of these elements is very useful to acquire the image processing knowledge used by the planning system: they are used to justify the choice of the algorithms regarding the application context and therefore to define the conditions of applicability of the image processing techniques. It also favors the reusability of solution parts.

References

1. Clouard, R.: Une méthode de développement d'applications de traitement d'images. Traitement du signal **21**(4) (2004) 277–293
2. Matsuyama, T.: Expert systems for image processing: knowledge-based composition of image analysis processes. CVGIP **48**(1) (1989) 22–49
3. Liedtke, C., Blömer, A.: Architecture of the Knowledge Based Configuration System for Image Analysis "Conny". In: ICPR'92. (1992) 375–378
4. Clément, V., Thonnat, M.: A knowledge-based approach to integration of image procedures processing. CVGIP: Image Understanding **57**(2) (1993) 166–184
5. Chien, S., Mortensen, H.: Automating image processing for scientific data analysis of a large image database. IEEE PAMI **18**(8) (1996) 854–859
6. Clouard, R., Elmoataz, A., Porquet, C., Revenu, M.: Borg : A knowledge-based system for automatic generation of image processing programs. IEEE PAMI **21**(2) (1999) 128–144
7. Hasegawa, J., Kubota, H., Toriwaki, J.: Automated construction of image processing procedures by sample-figure presentation. ICPR **86** (1986) 586–588
8. Nouvel, A., Dalle, P.: An interactive approach for image ontology definition. In: 13ème Congrès de Reconnaissance des Formes et Intelligence Artificielle, Angers, France (2002) 1023–1031
9. Maillot, N., Thonnat, M., Boucher, A.: Towards ontology based cognitive vision. In: ICVS. Volume 2626 of LNCS., Springer (2003) 44–53
10. Bombardier, V., Lhoste, P., Mazaud, C.: Modélisation et intégration de connaissances métier pour l'identification de défauts par règles linguistiques floues. Traitement du Signal **21**(3) (2004) 227–247
11. Town, C.: Ontological inference for image and video analysis. Mach. Vision Appl. **17**(2) (2006) 94–115
12. Bloehdorn, S., Petridis, K., Saathoff, C., Simou, N., Tzouvaras, V., Avrithis, Y., Handschuh, S., Kompatsiaris, Y., Staab, S., Strintzis, M.G.: Semantic annotation of images and videos for multimedia analysis. In: ESWC. Volume 3532 of LNCS., Springer (2005) 592–607
13. Renouf, A., Clouard, R., Revenu, M.: Un modèle de formulation d'applications de traitement d'images. In: ORASIS'05, Fournols, France (2005)
14. Schreiber, G., Wielinga, B., Akkermans, H., Van de Velde, W., Anjewierden, A.: CML: The CommonKADS Conceptual Modelling Language. In: EKAW 94. Volume 867 of LNCS., Hoegaarden, Belgium, Springer Verlag (1994) 1–25
15. Lezoray, O., Cardot, H.: Cooperation of color pixel classification schemes and color watershed : a study for microscopic images. IEEE Trans. on Image Processing **11**(7) (2000) 783–789
16. Bachimont, B. In: Engagement sémantique et engagement ontologique : conception et réalisation d'ontologies en ingénierie des connaissances. Eyrolles (2000) 305–323
17. Hudelot, C., Maillot, N., Thonnat, M.: Symbol grounding for semantic image interpretation : from image data to semantics. In: Proceedings of the Workshop on Semantic Knowledge in Computer Vision, ICCV, Beijing, China (2005)
18. Coradeschi, S., Saffiotti, A.: An introduction to the anchoring problem. Robotics and Autonomous Systems **43**(2-3) (2003) 85–96

Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)
Published in 2007 by Applied Computer Science Group, Bielefeld University, Germany, ISBN 978-3-00-020933-8