Towards Ontology Based Cognitive Vision

Nicolas Maillot, Monique Thonnat, and Alain Boucher

INRIA Sophia Antipolis - Orion Team
2004 Route des lucioles - B.P. 93
06902 Sophia Antipolis, France
{nicolas.maillot, monique.thonnat, alain.boucher}@sophia.inria.fr
http://www.inria.fr/orion

Abstract. Building knowledge bases for knowledge-based vision systems is a difficult task. This paper aims at showing how an ontology composed of visual concepts can be used as a guide for describing objects from a specific domain of interest. One of the most important benefits of our approach is that the knowledge acquisition process guided by the ontology leads to a knowledge base closer to low-level vision. A visual concept ontology and a dedicated knowledge acquisition tool have been developed and are also presented. We propose a generic methodology that is not linked to any application domain. Nevertheless, an example shows how the knowledge acquisition model can be applied to the description of pollen grain images. The use of an ontology for image description is the first step towards a complete cognitive vision system that will involve a learning layer.
keywords : cognitive vision, ontological engineering, knowledge-based vision systems.

1 Introduction

Many knowledge-based vision systems have been suggested in the past (SCHEMA [1], SIGMA [2], SYGAL and PROGAL [3]). They all need knowledge bases specifically designed for the application domain. As explained in [1], designing such bases is very time consuming. This task also needs multidisciplinary skills. Indeed, both domain knowledge and image processing techniques are involved in this process. Our goal is to acquire domain knowledge without requiring image processing skills. We propose to use a visual concept ontology to hide the low-level vision layer complexity and to guide the expert in the description of the objects of his/her domain.

We are planning to build a complete classification system that will use the resulting knowledge base. Please note that our approach is different from certain kinds of ontology-based image retrieval techniques [4] where a domain-dependent ontology is used to annotate images : retrieval is based on attached annotations and not on image analysis techniques.

This paper first details our proposed approach in section 2. Section 3 introduces the reader to ontological engineering. Section 4 is dedicated to the knowledge acquisition process we propose. Section 5 presents an ontology composed of three different types of visual concepts : spatio-temporal, texture and colour related concepts. Section 6 shows how the proposed ontology can be used for describing objects from a specific domain. Finally, section 7 details the features of the specific knowledge acquisition tool we have developed and used for the description of several pollen grains types.

2 Proposed Approach

As described in [2], several types of knowledge can be identified in knowledge-based vision systems (Fig. 1) : (1) domain knowledge, (2) knowledge about the mapping between domain knowledge and image processing knowledge, (3) image processing knowledge.

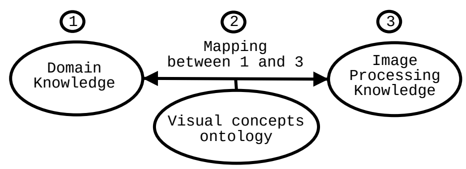

Fig. 1. Knowledge Overview

Our work is focused on the mapping betwen domain knowledge and image processing knowledge. Extracting domain knowledge means producing a hierarchical structure of domain concepts associated with their subparts (Fig. 2). This knowledge belongs to the domain of interest and is shared by the specialists of the domain. It is important to note that domain knowledge is independant of any vision layer and can be reused for other purposes. In our approach, the mapping between the scene and the image is done during a separate step and leans on a visual concept ontology. This ontology is used as a guide which provides a simple vocabulary used to give the visual description of domain concepts.

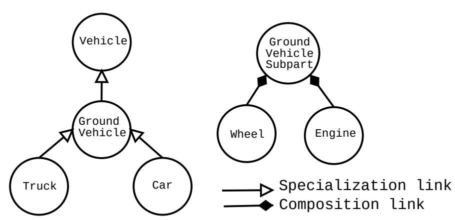

Fig. 2. Domain Knowledge Structure

3 Ontological Engineering

Tom Gruber gives a **definition** of the notion of ontology in [5] : "An ontology is an explicit specification of a conceptualization".

As explained by B. Bachimont (see [6]), the aim of ontologies is to define which primitives, provided with their associated semantics, are necessary for knowledge representation in a given context.

An ontology is composed of the set of objects that can be represented and also of relations between objects. It is important to notice that a semantic must be given to the ontology. This can be achieved by specifying axioms. An exhaustive list of axioms is given in [6]. They can be related to reflexivity, symmetry or transitivity of relations.

To be efficient, communication between people and software systems must rely on a shared understanding. As explained in [6], lack of shared understanding leads to difficulties in identifying requirements and to limited inter-operability or reusability. These problems are often met when building or interacting with computer vision systems. Ontologies are a common base to build on and a shared reference to align with [6]. That is why ontological engineering can be useful for our community.

A relevant example of ambiguity was given by Gómez-Pérez during a talk at ECAI98 [6]. What should be answered to the question "What is a pipe?". There are several possible answers: a short narrow tube with a small container at one end, used for smoking tobacco; a long tube made of plastic or metal that is used to carry water or oil or gas; a temporary section of computer memory that can link two different computer processes.

Ontological engineering is based on a consensus that avoids ambiguous situations. As explained in [7], ontology development process has to be done in four distinct phases. The first one is called specification and states why the ontology is built and who are the end-users. The next phase is conceptualization and leads to a structured domain knowledge. Then comes the formalization phase that transforms the conceptual model into a formal model. Finally, implementation transforms the formal model into a computational model. This methodology has been used to obtain the visual concept ontology presented in section 5.

4 Overview of the Knowledge Acquisition Process

As described in Fig. 3, the proposed knowledge acquisition process leans on the visual concept ontology. The expert starts by producing domain knowledge [1]. Then comes the visual concept ontology-driven description phase. This means that the expert uses the vocabulary provided by the ontology to describe the objects of the domain. This task is performed in a user-friendly way with a graphical user interface. The result of the description phase is a knowledge base composed of the visual concepts provided by the ontology associated with domain

[1] Structured as a hierarchy of domain concepts with their subparts

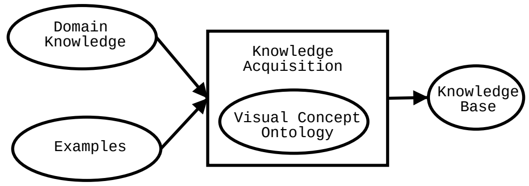

Fig. 3. Knowledge Acquisition Process

concepts. For example, the visual concept "Circular surface" provided by the ontology can be used to describe the shape of a domain object.

Our final goal is to build a classification system that will make a relevant use of examples associated with domain concepts. That is why domain objects examples are also provided. Once again, a user interface is used to provide these examples.

5 An Ontology for Computer Vision

5.1 Motivations

We believe that the creation of a standardized ontology for computer vision would be a great step forward and would ease interoperability between vision systems. In this section, we propose a prototype of an ontology for computer vision. We did not intend to build an exhaustive ontology. Our goal is to give an overview of a methodology based on a visual concept ontology. Due to the limited number of pages of this paper, some parts of the ontology are not presented. It is important to note that this ontology is not application-dependent.

We have structured this ontology in three main parts. The first one contains spatio-temporal related concepts, the second one contains texture related concepts and the last one is made of colorimetric concepts. Each part of this ontology is going to be detailed in the next sections.

5.2 Spatio-temporal Concepts

This part of the ontology is used for describing domain objects from a spatio-temporal point of view. For example, a part of the hierarchy is composed of geometric concepts that can be used to describe the shape of domain objects (Fig. 4). A justification of an approach based on geometric shapes can be found in [8]. The size of an object can also be described and quantified with a set of quantifiers. Note that quantification can be done in an absolute way or relatively to another concept. This means that size of object A can be described as being important relatively to object B. The notion of elongation is also present and can be quantified. We have also added a set of spatial relations based on the RCC-8 model [9] that can be used to define relations between objects and their subparts. Temporal relations have not yet been introduced in the ontology.

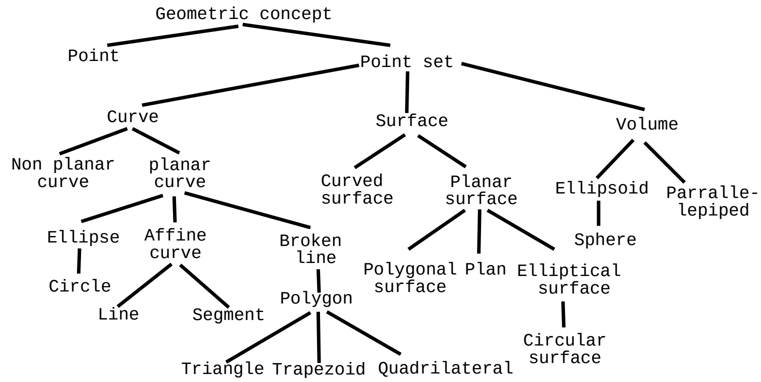

Fig. 4. Geometric Concepts

5.3 Texture Concepts

This part of the ontology has been inspired by [10] which is the result of two experiments. The first one is on the categorization of texture words in order to identify the underlying dimensions used to categorize texture words. The second part of the experiment measures the strength of association between words and texture images. The resulting hierarchy is given in Fig. 5. A very interesting aspect of this study is that hundreds of synonyms have been gathered in association with texture notions. This rich terminology gives expressivness to the knowledge acquisition system we propose. Once again, several texture concepts can be quantified : the quantifier "important" can be used in association with the "granulated texture" concept.

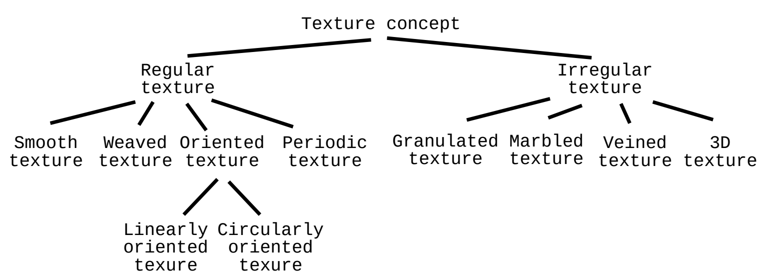

Fig. 5. Texture Concepts

5.4 Colour Concepts

This part of the ontology allows the description of objects from the following points of view : luminosity, hue, transparency. These concepts can be quantified to allow a more precise description. For example, the quantifier "important" can be used to quantify colorimetric concept "dark".

5.5 Link with the Low-Level Vision Layer

The previous subsections have quickly introduced the reader to the structure of a proposed visual concept ontology. The knowledge base resulting from our knowledge acquisition process is for classification purposes. During the classification of a given object, numerical descriptors are computed. To be interpreted as visual concepts, a link must be established between computed numerical descriptors and symbolic visual concepts. Currently the link between symbolic visual concepts and numerical descriptors is statically defined. For example, the ratio $length/height$ computed for a region is used to characterize the "elongation" visual concept.

There is not a unique way to associate a visual concept with a set of numerical descriptors. That is why we are planning to introduce a descriptor selection layer that chooses the most discriminative descriptors for a given application domain. Examples of numerical descriptors for texture discrimination are statistical moments or descriptors obtained with gabor filtering. Colour characterization can be done with numerical entities like histograms, colour coherence vectors or auto-correlograms.

5.6 Context Description

Experts often observe the objects of their domain in precise observation conditions. For example, when using a microscope, magnification or lighting conditions are controled. This kind of information must be given because it is linked to the way objects are described. Context knowledge is necessary to build coherent sets of examples. Context depends on the application domain. That is why context hierarchy given in Fig. 6 can be extended for a particular domain.

6 Knowledge Representation

In the previous section, we have made a conceptualization of a specific domain knowledge. The result of the conceptualization process is an abstract object. Before obtaining an operational entity, a formalism has to be chosen. This is called the formalization process which implies the choice of a representation formalism. Different kinds of representation formalisms are enumerated in [11]. Commonly used techniques are formal logics, fuzzy logics, frames or semantic nets. We wanted an expressive and powerful formalism. Several reasons exposed

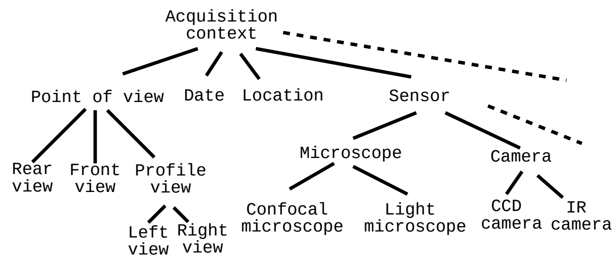

Fig. 6. Context Concepts

in [12] lead us to description logics (DL). DL is the family name of object-based knowledge representation formalisms which provide formal semantics for semantic nets and the logical fundations for inferences.

A concept of the domain is described through four relations : *hasForSpatioTemporalDescription*, *hasForTexturalDecription*, *hasForColorimetricDescription* and *hasForDescriptionContext*.

Description logics are used to structure the description information:

$$C_i \equiv C_j \sqcap \ (\exists \ "hasForSpatioTemporalDesc".C_{SpatioTemporal_i})$$
$$\sqcap \ (\exists \ "hasForTextureDesc".C_{Texture_i})$$
$$\sqcap \ (\exists \ "hasForColorimetricDesc".C_{Colour_i})$$
$$\sqcap \ (\exists \ "hasForDescContext".C_{Context_i}) \sqcap \ (\exists \ "isASubpartOf".C_k)$$

This means that C_i is a subconcept of C_j and a subpart of C_k. Relations *hasForSpatioTemporalDescription*, *hasForTexturalDescription*, *hasForColorimetricDescription*, *hasForDescriptionContext* are respectively restricted to concepts $C_{SpatioTemporal_i}$, $C_{Texture_i}$, C_{Colour_i}, $C_{Context_i}$. The powerful expressivness of description logics allows to define $C_{SpatioTemporal_i}$, $C_{Texture_i}$, C_{Colour_i}, $C_{Context_i}$ as unions or intersections of concepts provided by the visual concept ontology.

7 A Knowledge Acquisition Tool for Image Description

7.1 Overview

Section 5 contains details about the structure of a visual concept ontology. To be used as a guide for the description of domain concepts, a dedicated graphical tool has been developed. This tool is currently able to carry out three distinct tasks : domain knowledge definition; visual concept ontology-driven symbolic description of concepts and their subparts; examples management.

The output result of the acquisition process is a knowledge base composed of domain concepts described by some visual concepts provided by the ontology.

The formalism used to structure the resulting knowledge is based on description logics. The Java platform has been used to create this tool. The knowledge layer leans on DAML+OIL [2], an ontology language based on the formal rigor of description logics.

7.2 Tool Characteristics

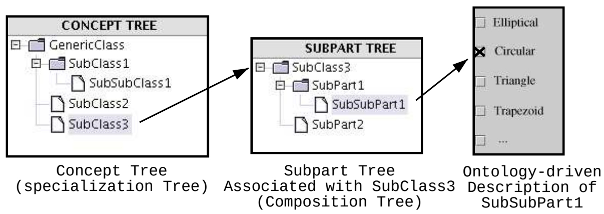

Fig. 7. Description of subpart "SubSubPart1" of Concept "SubClass3"

Our tool allows domain knowledge creation. As can be seen in Fig. 7, domain knowledge is organized as a taxonomy of domain concepts in a specialization tree. This approach is natural for people who are familiar with a taxonomic approach (Ex.: biologists). Whenever a concept is added to the tree, the visual concept ontology is displayed on the screen. The user is then able to describe a new concept with the terminology contained in the ontology. As it was previously explained, a concept can be composed of subparts (*isASubpartOf* relation).

Subparts description is performed in the same way as the description of domain concepts. Note that the subpart tree is a composition tree and not a specialization tree. Every domain concept has an associated subpart tree.

Another important characteristic of our tool is the example management module which allows to provide examples of the visual concepts used during the description phase. First of all, a set of examples has to be chosen. Then, a region of interest is selected (with a drawing tool) and used to compute the required numerical descriptors. Fig. 8 describes how an example of the subpart Pori [3] is provided. Since the Pori subpart is described as being circular, the computation of a form factor on this particular image gives an example of what is the notion of circularity in this particular application domain.

7.3 Results

Automatic pollen grains classification is useful for clinicians so as to provide near real time accurate information on aeroallergens and air quality to the sensitive

[2] DAML+OIL belongs to the family of XML languages
[3] A subpart of certain pollen grains

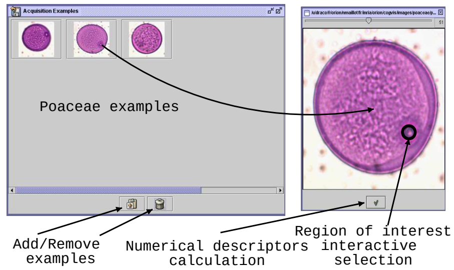

Fig. 8. An example of Poaceae's Subpart "Pori" Provided by Interactive Selection of a Region of Interest

users. Our tool is currently used as a help for communicating with experts in palynology. This tool is useful to guide the description of pollen grains in a user-friendly and efficient manner. The visual concepts contained in the ontology can be seen as a communication language between us and the experts. Although sets of numerical descriptors are associated with visual concepts, they are hidden to the expert : when choosing a visual concept, the expert implicitly chooses a set of numerical descriptors. This why the generated knowledge base is closer to low-level vision. 350 images of 30 different types of pollen grains have been acquired during the A.S.T.H.M.A [4] european project. This database is useful to give examples of the visual concepts used to describe these pollen grains.

A possible example could be the symbolic description of the specific pollen grain type *Poaceae*[5] (Fig. 8). This concept can be described as a subconcept of *Pollen with apertures*. It is described as an **elliptical or circular surface** [6]. The **size** of this pollen type is quantified with the following quantifiers : **small**, **average** or **important**. Palynogists use the concept **granulated texture** to describe Poaceae's texture. The colorimetric concept **dark** is used to describe the subpart *Pori*. Finally comes acquisition context : In this case, a **light microscope** is used to acquire pollen images.

8 Conclusion and Future Work

We have proposed an original approach to the creation of knowledge-based vision systems. The notion of visual concept ontology has been introduced. Its structure is based on three distinct notions : spatio-temporal, texture, colour concepts. The description is contextualized by a set of context concepts. This ontology can be used as a guide for describing the objects from a specific domain. We are planning to extend the ontology : our efforts are now focused on creating

[4] http://www-sop.inria.fr/orion/ASTHMA/
[5] Italic terms are provided by the expert
[6] Bold terms are provided by the ontology

an ontology that will allow the definition of improved spatio-temporal relations. We aim at applying this future ontology to the description of video content.

Another important aspect of the model we propose is the examples database. In the future, this database will feed a supervised learning process of the symbolic visual concepts. Different kinds of texture in the provided examples will allow to learn the difference between a granulated texture and a smooth texture. The remaining step is the generation of a complete classification system which should make heavy use of the resulting learned concepts.

A difficult remaining problem is the segmentation process. Indeed, in order to be classified, images have to be segmented to allow descriptor computation. The symbolic description made by the expert may help finding the image processing tasks required for extracting the pertinent information from the provided images. As an example, an object described with the "granulated texture" concept may be segmented with a texture based segmentation algorithm. The regions of interest selected by the expert (see Fig. 8) should be used to validate the resulting segmentation.

References

1. Draper, B., Hanson, A., Riseman, E.: Knowledge-directed vision: control, learning and integration. In: Proc. of IEEE. Volume 84. (1996) 1625–1681
2. Matsuyama, T., Hwang, V.S.: SIGMA - A Knowledge-Based Aerial Image Understanding System. Plenum Press New York USA (1990)
3. Thonnat, M., Bijaoui., A.: Knowledge-based galaxy classification systems. Knowledge-based systems in astronomy, Lecture Notes in Physics. **329** (1989)
4. Soo, V.W., Lee, C.Y., Yeh, J.J., chih Chen, C.: Using sharable ontology to retrieve historical images. In: Proceeding of the second ACM/IEEE-CS joint conference on Digital libraries, ACM Press (2002) 197–198
5. Gruber, T.R.: 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, Deventer, The Netherlands, Kluwer Academic Publishers (1993)
6. Gandon, F.: Ontology engineering: A survey and a return on experience. Technical Report 4396, INRIA (2002) http://www.inria.fr/rrrt/rr-4396.html.
7. Blazquez, M., Fernandez, M., Garcia-Pinar, J., Gómez-Pérez, A.: Building ontologies at the knowledge level using the ontology design environment. In: KAW98. (1998)
8. Sciascio, E., M.Donini, F., Mongiello., M.: Structured knowledge representation for image retrieval. Journal of Artificial Intelligence Research **16** (2002) 209–257
9. Cohn, A.G., Hazarika, S.M.: Qualitative spatial representation and reasoning: An overview. Fundamenta Informaticae **46** (2001) 1–29
10. Bhushan, N., Rao, A., Lohse, G.: The texture lexicon: Understanding the categorization of visual texture terms and their relationship to texture images. Cognitive Science **21** (1997) 219–246
11. Crevier, D., R.Lepage: Knowledge-based image understanding systems: A survey. Computer Vision and Image Understanding **67** (1997) 161—185
12. Moller, R., Neumann, B., Wessel, M.: Towards computer vision with description logics: some recent progress. In: Proc. Integration of Speech and Image Understanding. (1999)