Annals of DAAAM for 2010 & Proceedings of the 21st International DAAAM Symposium, Volume 21, No. 1, ISSN 1726-9679
ISBN 978-3-901509-73-5, Editor B. Katalinic, Published by DAAAM International, Vienna, Austria, EU, 2010
Make Harmony Between Technology and Nature, and Your Mind will Fly Free as a Bird
Annals of DAAAM International

A MODEL REPRESENTATION FOR IMAGE CONTENT

MOCANU, I[rina] G[eorgiana] & NEGREANU, L[orina] C[ristina]

Abstract: *This paper presents a model for formalization of the image representation. The formalization is based on the description logics language ALC(MS). In the proposed model, an image will be described based on both its low-level features: color and shape and its semantic content. The presented model will allow the determination of image similarities based on the visual content (low level features: color and shape) and on the semantic content*
Key words: *ALC(MS), semantic content, image content, image similarity*

1. INTRODUCTION

While the number of images on the www keeps increasing, it becomes necessary to find new methods for obtaining similar images or image retrieval by content. There are large collections of images and we want to find images using a query. The query will be an image too. In this case "images like "query image" means images which are similar with the query image. It is a new problem of how to describe images for finding similarities between them. The description of an image is realized based on the low level features: color, shape and texture or combination of these. The images from the queried collections and the query images are described in the same way (using the same low level feature). First a metric distance will be applied on these descriptions. The similar images will be obtained by sorting ascending the distance values obtained in the first step. But sometimes is hard to specify a query. For example if we want to find images that contain sky, the query may be specified by color distribution. However the sky is blue in a sunny day, orange at sunset or grey in a cloudy day. Thus it is hard to specify such a query using only the low level feature color. For doing this the images are described by their semantic contents using annotations. These associations are a translation from image instances (object from images) to keywords, based on machine learning techniques, as described in (Duygulu et al., 2001). The reason for this difficulty is the existence of the gap to bridge between the low-level features of images' objects and their semantic contents.

This paper describes a formal model for image representation using some kind of description logics language. The presented model is intended to make more usefull the process of image retrieval using both the visual contents and the semantic contents. The image representation model is based on the visual and the semantic contents. The paper is organised as follows: in section 2 we briefly review related work. Section 3 presents the description logics language ALC(MS) used for the proposed formal model and section 4 draws the conclusions and the future approaches.

2. RELATED WORK

There are some models for image retrieval. A first approach is the logical model presented in (Sciascio et al., 1999). The model is built for exact recognition of two-dimensional objects in images. The recognition of objects is made by exactly matching their shape, considering all possible transformations between the compared shapes. This model is extended in (Sciascio et al., 2002) for similar matching by shape. Another approach of image representation is described in (Aiello, 2002) in which is presented a model using the semantic content together with the spatial relations between objects.

In this paper we present a model for describing images on visual content: color and shape and also on their semantics which will make more usefull the image retrieval process.

3. IMAGE CONTENT FORMALIZATION

3.1 The ALC(MS) formalization language

For the image content formalization we will use the ALC(MS) language. ALC(MS), a joining of the languages ALC and MS, is put together as described in (Kutz et al., 2001).

The description of the ALC(MS), as it is described in (Kutz et al., 2001), consists of: (a) The primitive symbols of ALC: (i) a list of concept names A_0, A_1, … (ii) a list of role names R_0, R_1, … (iii) a list of object names c_0, c_1, …, (iv) the Booleans $\sqcap$, $\neg$, $\top$ and $\exists$; (b) The primitive symbols of MS: (i) a list of set variables X_0, X_1, …; (ii) a list of nominals n_o, n_1, …; (iii) the set term constructors $E_{\leq\alpha}$ and $E{>}\alpha$, for $\forall\alpha\in Q_+$ (the subset of positive rational numbers) and the Booleans $\cap$ and $\neg$. Concepts and set terms of ALC(MS) are defined inductively as follows: (i) All concept names A_i are concepts; (ii) All set variables X_i and nominals are set terms; (iii) If C and D are concepts, R a role name and t a set term, then $C \sqcap D$, $\neg C$, $\exists R.C$ are concepts; (iv) If s and t are set terms, C is a concept, c an object name and $\alpha\in Q_+$ then $t \cap s$, $\neg t$, $E_{\leq\alpha}t$, $E_{>\alpha}t$ are set terms;(e) Formulas of ALC(MS) are Boolean combinations of atomic formulas of the form $c : C$, cRd, $C \subseteq D$, $t \subseteq s$, where C, D are arbitrary concepts, c, d are object names and t, s are arbitrary set terms.

Each model of the ALC(MS) language consists of a standard ALC model, a metric space model for MS and a relation between their domains.

3.2 The formalization model for image content

An image is viewed as the set of its component regions $\{r_1, r_2, \dots, r_n\}$. The component regions are obtained using different segmentation algorithms. The description of an image is composed by the descriptions of the component regions. Thus we want to formalize the model for the description of a region. In (Mocanu & Negreanu, 2009) we propose a modality to formalize the image similarities based both on the images' low level features (color and shape) and its semantics using the ALC(MS) language. Based on that formalization, if we want to specify an image which contains a toy in it we have:

$$I \subseteq Image,\ r \in \forall Region.I$$
$$(toy, r) : Representation$$
$$(r, c) : ColorRepresentation$$

If we want to specify an image with a red toy in it, we have:

$$I \subseteq Image,\ r \in \forall Region.I$$
$$(toy, r) : Representation$$

(r, c) : ColorRepresentation

$(c \in E_{\leq 0} \{red\} \sqcap (\neg(E_{>0} \{red\})))$

And if we want an image with a circle, red toy in it we will describe it like:

$I \subseteq$ Image, $r \in \forall$Region.I

(toy, r) : Representation

(r,c) : ColorRepresentation, (r,s) : ShapeRepresentation

$(c \in E_{\leq 0} \{red\} \sqcap s \in E_{\leq 0} \{circle\})$

But when we want to describe an image we will use similarity matches and not exactly matches. Thus the above description will be transformed in the description of an image that contains a toy, which has a shape similar with a circle and its color is similar with red. This will be formalized like:

$I \subseteq$ Image, $r \in \forall$Region.I

(toy, r) : Representation

(r, c) : ColorRepresentation, (r, s) : ShapeRepresentation

$(c \in E_{\leq \tau}\{red\}) \sqcap (s \in E_{\leq \tau}\{circle\})$

(the distance between the region's shape and the circle shape must be under a specified threshold; also the distance between the region's color and the red color must be under a specified threshold, too; the only modification is $E_{\leq 0}$ become $E_{\leq \tau}$).

Thus an image region will be represented by: (i) its semantic content – a keyword which describes the region; and (ii) the content description based on the low level features: color and shape. Thus the model must describe a region of an image based on its visual and semantic content. Also it is necessary to specify how we decide if two regions are similar.

A region will be described as <k, c, s>, where k is a keyword describing the region (a concept name), c and s are the color and the shape associated with a region (in fact c and s are color and shape representations for that region). An image I is viewed as the union of its component regions: $I = r_1 \cup r_2 \cup \dots \cup r_n$ and the description of the image based on this model will be: $<k_1, c_1, s_1> \sqcap <k_1, c_1, s_1> \sqcap \dots \sqcap <k_1, c_1, s_1>$

We use the idea of the semantics used in (Kutz et al., 2001): the domain of interpretation is a set of images Δ and regions are interpreted as subsets of Δ. An interpretation is a pair $<I, \Delta>$, where Δ is a set of images and I is a mapping from regions to subsets of Δ. An image $I \in \Delta$ belongs to the interpretation of a component region $<R>^I$ if image I contains a region who matches region R:

$$<R>^I = \{I \in \Delta \mid \exists\, r \in I: r \cong R\} \quad (1)$$

where $\cong$ means region r is similar with region R.

A region of an image may be viewed similar with another region R if r and R have similar shapes and/or they have similar colors and the associated concepts are equivalent. Thus (1) may be rewritten as:

$$<R>^I = \{I \in \Delta \mid \exists\, r \in I: k_r = k_R \sqcap c_r \cong c_R \sqcap s_r \cong s_R \quad (2)$$

where c_r and s_r represents the color representation of region r, respectively the shape representation of region r; k_r is the keyword associated with region r,

Mapping these similarities in ALC(MS) we will obtain:

$$<R>^I = \{I \in \Delta \mid \exists r \in I: (c_r \in E_{\leq\tau}\{c_R\}) \sqcap (s_r \in E_{\leq\tau}\{s_R\}))\} \quad (3)$$

When we write it as a function we obtain:

$$R^I(I) = \max r \in I \{sim(r,R)\} \quad (4)$$

where sim(r, R) is the similarity degree between the two regions r and R:

$$sim(r, R) = sim(c_r, c_R) * \alpha + sim(s_r, s_R) * \beta \quad (5)$$

Regions can be decomposed such that a region ri may be viewed as a set of component sub-regions. Let RC be a composite region: $<R_1> \sqcap <R_2> \sqcap \dots \sqcap <R_n>$. In the exact matching, the interpretation is the intersection of the sets interpreting each component of the region:

$$<RC>^I = \{I \in \Delta \mid I \in \sqcap_{i=1,n} <R_i>^I\} \quad (6)$$

A region description RC is recognized in an image I if for every interpretation (I, Δ) with $I \in \Delta$, we have $I \in RC^I$. For similar matching, the composite region description RC^I is determined by the least similar region component. So equation (6) will be transformed in:

$$RC^I(I) = \min_{i=1,n} <R_i>^I \quad (7)$$

Based on the used semantic an image database is a domain of interpretation and a region R is a subset of such a domain, that means all the images to be retrieved when R is a viewed as a query.

4. CONCLUSION AND FUTURE WORK

The paper presents a formal model for image representation based both on its content and semantics. The description of the model is realized using the description logics language ALC(MS). In future work this model will be extended with image similarities based on spatial similarity. But the interpretation of an image is not complete without some mechanism for understanding their semantic content. Thus the model will be extended with semantic similarities between annotations based on concepts, as described in (Sebti & Barfroush, 2008).

5. REFERENCES

Aiello, M.. (2002). Spatial Reasoning: Theory and Practice, *Available from* http://www.cs.rug.nl/~aiellom/publications/thesis.pdf *Accessed*: 2010-09-15

Duygulu, P., Barnard, K., de Freitas, J. F. G. & Forsyth, D. A. (2003). Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary, *Proceedings of the 7th European Conference on Computer Vision-Part IV* (Ed.), pp. 97–112, ISBN 3-540-43748-7, May 2002, Copenhagen, Danmark, Springer-Verlag, London

Kutz, O. Wolter, F. & Zakharyaschev, M. (2001). A Note on Concepts and Distance, Vol. 49, *Proceedings of the International Workshop on Description Logics CEUR-WS*, Stanford, D. McGuiness, P. Patel-Schneider, C. Goble, R. Moeller (Ed.), pp. 113 – 121, ISSN 1613-0073, California, USA, August 2001, CEUR-WS, Aachen

Mocanu, I. & Negreanu, L. (2009), Image Retrieval Queries in Description Logic, *Annals of DAAAM for 2009 & Proceeedings of the 20th International DAAAM Symposium*, Katalinik, B. (Ed.), pp. 167-168, ISBN 978-3-901509-70-4, Vienna, Austria, October 2009, DAAAM International, Vienna

Sciascio, DI E., Donini, F. & Mongiello, M. (1999). A Description Logic for Image Retrieval, *Proceedings of the 6th Congress of the Italian Association for Artificial Intelligence* , Evelina Lamma Paola Mello (Ed.), pp. 13-24, ISBN 3-540-67350-4, Bologna Italy, September 1999, Springer-Verlag, Germany

Sciascio, DI E., Donini, F. & Mongiello, M. (2002). Structured Knowledge Representation for Image Retrieval, *Journal of Artifficial Intelligence Research*, Vol. 16, Issue 1, (January, 2002) pp. 209-257, ISSN 1076-9757

Sebti, A., Barfroush, A.A., A new sense similarity measure in WordNet (2008). *IMCSIT*, pp. 369-373, ISBN 978-83-60810-14-9/ISSN 1896-7094, Wisla Poland, October 2008