CONTENT BASED VISUAL INFORMATION RETRIEVAL FOR MANAGEMENT INFORMATION SYSTEMS

Oleg Starostenko, Jose Alfredo Sánchez H. , Roberto Rosas R.

Universidad de las Américas-Puebla

Alberto Chavez-Aragon

Autonomous University of Tlaxcala

Abstract

This paper presents the results of our research devoted to development of efficient methods for retrieval and indexing of documents with multimedia information that can help a business work smarter and gain an important advantage in whatever that business does. Particularly, a novel hybrid method for visual information retrieval (VIR) is proposed. It combines shape analysis of objects in image with their automatic indexing by textual descriptions applying semantic Web approaches. A decision about similarity between a retrieved image and user queries is taken by computing the shape star field or two-segment turning functions combining them with matching of ontological annotations of objects in image providing in this way the machine-understandable semantics. For analysis of this method the image retrieval IRONS (Image Retrieval by Ontological Description of Shapes) system has been implemented and evaluated in some specific image domains.

Keywords: *Management IS, CBIR, Semantic Web, Shape analysis.*

1 INTRODUCTION

Recently, a quick increasing the amount and complexity of digital collections on Web requires the development of new searching engines and novel methods for multimedia information retrieval because the most of Web documents contain any kind of visual data. However, the most businesses don't know how to develop a successful management information system (MIS) that often stems from management's inability to understand both the power and limitations of information technology [Emery, 1987], [Kumar, 2005]. Integration and using the recent Web-based technologies and particularly, development of MIS for multimedia information search, access, retrieval, and processing permits to business generate new revenue from online products querying and sales, reduce costs through online customer support, attract new customers via web marketing and advertising, develop new Web-based markets and distribution channels for existing and new information-based products accessible on the Web [Gupta, 2005].

The most common way to classify images is generation of textual queries, however, the maintenance and the management of queries for image retrieval is a very time-consuming activity. The typical approach of automatic indexing the images as a principal part of retrieval process is based on the analysis of low-level image characteristics such as colour, texture or shape [Starostenko, 2005], [Chavez, 2006]. Unfortunately, this type of systems does not provide the semantics associated with content of an image. Nowadays, a novel approaches for development of VIR systems are those which provide the content based image retrieval (CBIR) to support management and treatment of multimedia documents. In practice, it is easier to execute queries based on low-level features such as shape, image dimensions, pixel values, colour, grey level, and histogram. Therefore, the problem resides in how to express a query, for instance *(recover red cars)* in terms of shape, colour, histogram, or pixels values. In other words, it is necessary to represent the content of an image query in a convenient way in order to search for the corresponding images from the collection and recover images with the best matching.

There are several reports about recent researches in the visual information retrieval area. CBIR systems can be classified into two main groups: Commercial Image Retrieval Systems such as Excalibur Visual RetrievalWare, IMatch, QBIC, Virage [Excalibur, 2004], [Virage, 2004], and Prototype Research Systems: AMORE, Photobook, VisualSeek, Black Box, IRONS, and Keyblock are examples of Prototype Research Systems [MIT, 2003], [Westphal, 2004], [Chavez, 2006]. The more perspective researches are those which use a semantic space to formulate queries. Table 1 summarizes the features of the well-known CBIR systems.

System	Colour	Shape	Texture	Relevance Feedback	Semantic Representation
Excalibur	YES	YES	YES	NO	NO
IMatch	YES	YES	NO	YES	NO
QBIC	YES	NO	YES	NO	NO
Virage	YES	YES	YES	-	NO
AMORE	YES	YES	NO	-	NO
Photobook	YES	YES	YES	YES	NO
VisualSEEK	YES	-	YES	YES	NO
Black Box	YES	YES	YES	YES	-
Keyblock	YES	-	-	YES	YES
IRONS	YES	YES	NO	YES	YES

Table 1. Visual features used in well-known Image Retrieval System

In this research we propose to apply the machine-understandable semantics for search, access, and retrieval of multimedia data using ontology. Grubber's famous description defines ontology as an approach which describes semantics, establishes a common and shared understanding of domain, and facilitates the implementation of user-oriented vocabulary of terms and their relationship with objects in image [Fensel, 2004], [Grubber, 1993]. In such a way, the meaning of an image may be obtained in textual form as set of descriptions for image related to a particular ontology. Proposed CBIR system may be used in business, medicine, education, architecture, entertainment, GIS, journalism, remote sensing, Internet shopping catalogue, tourist information support, etc. [Jensen, 2006], [Hiroshi, 2005].

2 PROPOSED SHAPE ANALYSIS APPROACHES

Perhaps the most obvious requirement of VIR systems is to retrieve images by shape because it represents significance of an object better than colour, pixel value, and texture.

Unfortunately, shape matching is considered one of the most difficult aspects of CBIR. The common way to describe a shape is the polygonal representation but it is not a convenient form to calculate how similar is that shape to another. In order to overcome this problem we propose to apply curve evolution process in order to reduce the set of elements which define the object without the lost of its significance. Fig.1 shows the results obtained after applying the proposed curve evolution algorithm to a shape. It is clear that curve evolution algorithm keeps the main visual parts of the original polygonal curve and the amount of information is reduced significantly. In order to make easier and more effective the matching process we propose an alternative to polygonal representation way, which we have called two-segment turning function (2STF) [Chavez, 2006]. Using 2STF a polygonal curve *P* is represented by the graph of a step function, the step on x - axis represents the normalized segment length and the y - axis represents the turn angle between two consecutive segments. Additionally, the normalization process to an angle between two consecutive segments is applied to make the representation in 2STF domain invariant to rotation. [Starostenko, 2007]. In Fig.1 2STF is presented for the same polygonal curve. The polygon on the right has been reflected and scaled by a factor of *0.8*. It is clear that both 2STFs have the similar characteristics. That allows getting the same representation for a set of shapes even though they are placed in different positions or has been scaled or rotated.

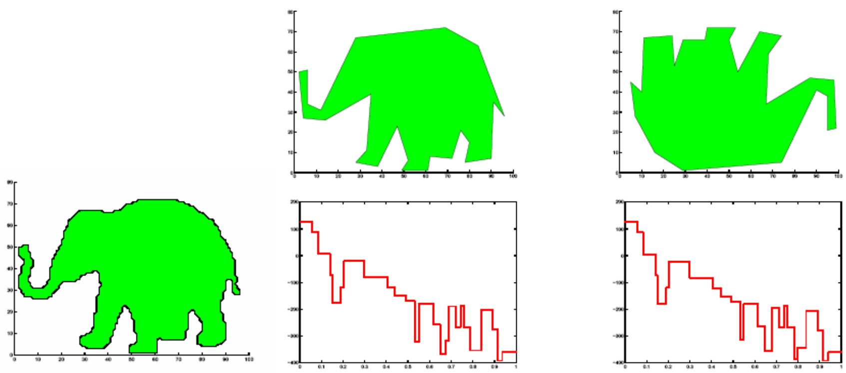

Figure 1. Evolution of the original shape (left) to the polygon of 30 segments and their 2STFs

The computing a similarity between two shapes in 2STF domain may be provided by three steps. The first step consists in decomposition of the whole curve into a set of maximal convex and concave arcs. Then, consecutive maximal arcs are grouped into sub-arcs. Finally, the best correspondence of sub-arcs from two different curves is calculated. The experiments for computing the time of valid combinations using a PC with processor of 2GHz and RAM of 1GB show that for polygons with 8 maximal arcs the time is acceptable (about 0.35s), for 9 arcs - about 4.3s, and for 10 arcs the time is more than 20s. Therefore, disadvantage of 2STFs comparison is significant time that it takes to find the best correspondence between two curves but the advantage of this approach is independence from scale, reflection, translation, and rotation. The problem of 2STF may be solved by another technique proposed by authors and called as Star Field. The Star Field (SF) is an alternative representation for shapes that allows obtaining more precise comparison. It means that we able to compare polygons with more than ten arcs (max value acceptable for 2STF). As a result, SF gives easier and faster matching process. Our Star Field method combines approach for computing the similarity among shapes proposed by [Mokhtarian, 1992] and its combination with 2STF. Mokhtarian proposed to use the maxima of curvature zero-crossing contours of Curvature Scale Space

(CSS) as a feature vector to represent shapes. However, computing CSS is an expensive process and we propose to use 2STF which is easier, faster and more effective [Medina, 2004].

Formally, a SF representation is a set of marks or stars $M_1;M_2:::M_{nm}$, where nm is the number of vertices of polygonal curve and it this number is equivalent to the number of steps in its 2STF. M_n is defined by means of two coordinates: x - coordinate indicates the normalized distance from the starting point to the corresponding vertex, making sure that in the middle of the SF plane there is a star that corresponds to the most important vertex of the polygon. The y - coordinate is the normalized angle between two consecutive segments which share the corresponding point. In other words, y - coordinate of stars correspond to the height of each step in its equivalent 2STF in the range [0,1] where a value of zero represents a $-\pi$ angle and one corresponds to $+\pi$. A SF diagram as 2D plane is divided horizontally into two sections. The upper section holds the stars that represent vertices of concave arcs. The lower part holds vertices of convex arcs. Actually, SF can be seen as a 3D surface and 2D SF is obtained by cutting and bending the torus presented in Fig.2 in the right.

For shape matching process we introduced the following strategy. Given two polygonal curves P_1 and P_2 and their star field SF_1 and SF_2, the graph G that allows us to compute their similarity is defined as $G = (V,E)$ where V and E are disjoint finite sets. We call V as the vertex set and E as the edge set of G. Our particular graph G has a set V that consists of two smaller subsets of vertices υ_1 and υ_2.

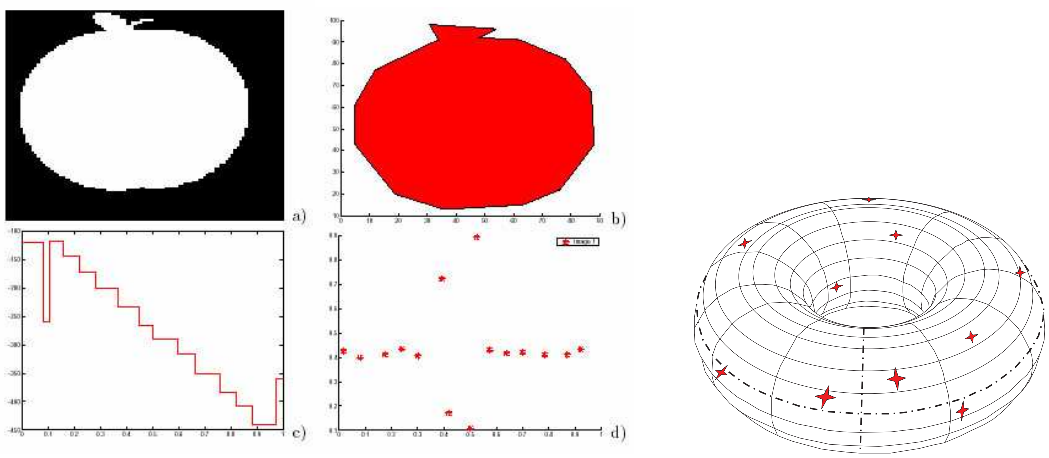

Figure 2. Original image, its polygon, 2STF, and SFs. Star Fields as a bending surface (torus)

Set $V = \upsilon_1 \cup \upsilon_2$., where υ_1 is the set of point of SF_1 and υ_2 is the set of points of SF_2. E is the set of pairs (r, s), where $r \in \upsilon_1$ and $s \in \upsilon_2$. Now we propose to use the adjacency matrix for representing the graph, where each cell of that matrix contains the cost for traveling from one column to each row and vice versa. The main idea behind the construction of the matching graph consists in building a connected weighted graph so that an algorithm to find the minimum spanning tree is applied. The minimum spanning tree is a subset of edges that forms a tree of vertexes, where the total weight of all edges in the tree is minimized. Thus, for the more similar shapes we obtain the lower value of corresponding total weight [Chavez, 2006]. The Fig.3 shows a 3D representation of the obtained graph. All star-like marks are connected to a single mark with weight equal zero. Additionally, they are connected with those crosslike marks which Euclidian distance has a lowest value. The searching the minimum span tree may be obtained by applying the Prim's algorithm [Medina, 2004].

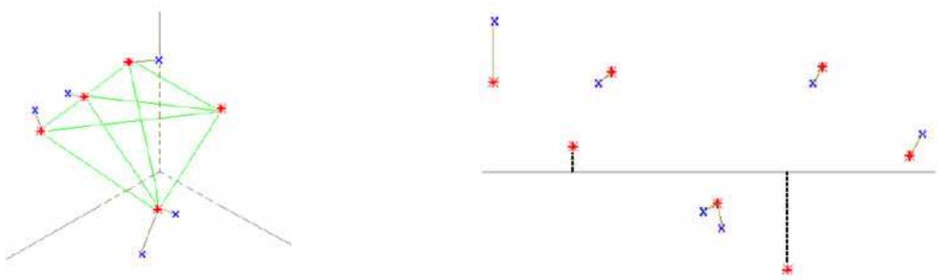

Figure 3. Matching graph (in the left) and additional cost of non-connected stars (in the right)

Finally, we can define how to calculate the similarity among shapes. The most important part of this calculation is the value of the cumulative weight of the edges that make up the spanning tree. However, the similarity is also affected by so-called penalty value. It is possible that some stars of the first shape never connected with stars from the second one, a penalty value is added to the final similarity measure. The additional cost is shown in Fig.3 in the right as dotted lines representing the distances that are added to the cumulative length. After experiments we conclude that new SF approach maintains advantages over 2STF providing the computing the connected weighted graph which is more fast and accurate approach.

3 IRONS IMAGE RETRIEVAL SYSTEM

Using the proposed method Image Retrieval by Ontological Description of Shapes (IRONS) system has been designed and implemented. Its block diagram is shown in Fig.4. The input for the system may be an image, its shape, or a keyword, which describes an object. The retrieved images will be those with more similarity to low-level features of a query and will have a high degree of matching with the ontological annotations defining content of an image. The IRONS system consists of 4 principal modules: query pre-processing, indexing module, feature vector comparison, and feedback GUI. The query pre-processing module provides the selection of sub-region containing the relevant objects. Once the sub-region is extracted, the object within that sub-region is found by the Regions Pre-processing Algorithm in Images (CORPAI) shape extraction algorithm [Medina, 2004]. The Smallest Unvalued Segment Assimilating Nucleus (SUSAN) method [Smith, 1996] for corner detection and CORPAI algorithm have been used for pre-processing of queries by image. The result of applying those algorithms is a convex polygon. The pre-processed polygon is represented by SF using 2STF invariant to scaling, rotation, reflection, and translation. The indexing module generates a feature vector which contains the polygon description and content-based ontological annotations which establish the relationship between the object and its formal explicit definition. This permits to speed up the matching process as well as reduce the number of iterations with nonsense results. The ontological annotation tool is used for searching matches in the ontology name space. In such a way, the meaning of an image may be obtained in textual form as a set of descriptions for each object related to a particular ontology. The Resource Description Framework (RDF) language supporting the ontology management has been used in this approach that defines a syntactic convention and a simple data model to implement machine-readable semantics. Using RDF it is possible to describe each Web resource with relations to its object-attributes-value based on metadata standard developed by the World Wide Web Consortium.

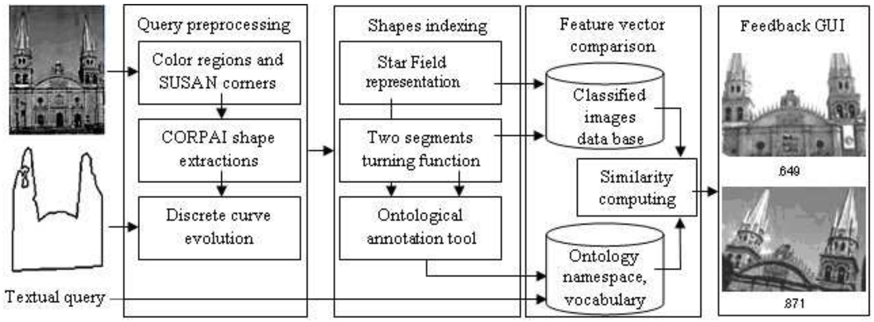

Figure 4. Block diagram of the proposed IRONS system

The ontology is described by a directed acyclic graph; each node has a feature vector that represents the concept associated with that node. Concept inclusion is represented by the IS-A inter-relationship. The feature vectors of each node in the ontology name space consist of keywords linking the previously classified images to the characteristics of a new shape extracted by SF or 2STF. The query interfaces of the IRONS system are presented in Fig. 5 where the images with high degree of matching are shown in downward order.

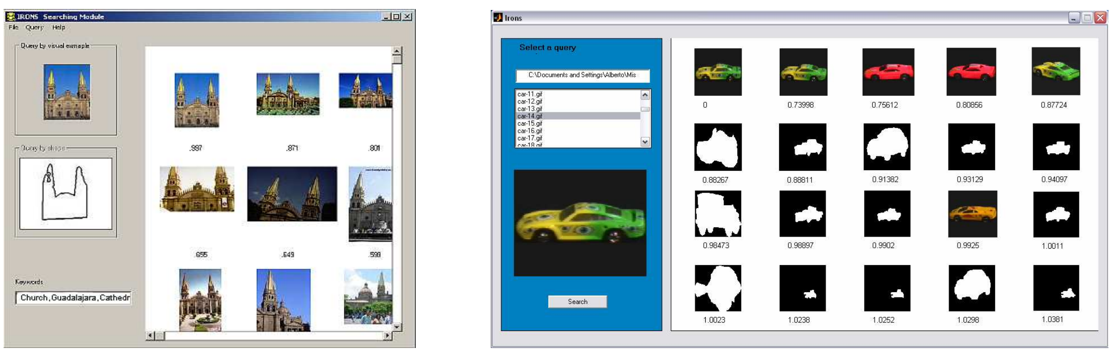

Figure 5. Examples of Image retrieval GUIs of the IRONS system

4 EVALUATION AND RESULTS

In the majority of the experiments of this research we used the on-line image database CE-Shape-1 (MPEG-7 standard) consisting of 1400 images divided into 70 classes with 20 images each. The evaluation of VIR system is non trivial task and it depends on individual perception of the user. Nevertheless, the standard approach is based on computing two measures: recall and precision. The recall means the proportion of relevant images in the entire database that are retrieved for comparison with a query. The precision is proportion of the retrieved images that are relevant to the query from the set of retrieved images. Formally, precision and recall are defined as it follows:

$$recall = P(B \mid A) = \frac{P(B \cup A)}{P(A)} = \frac{a}{a+c}, \quad precision = P(B \mid A) = \frac{P(B \cup A)}{P(B)} = \frac{a}{a+b},$$

where A is the set of relevant items, B is the set of retrieved items, a is the variable associated with retrieved and relevant images (hit), b is defined by retrieved and not relevant ones (partially relevant), and c is the case of not retrieved and relevant (miss). The variables a, b, and c will be used for establishment of relevance to query of each element in data base. We divided the experiments into three groups. In the first experiment we applied our SF technique to a sub-set of images randomly selected from CE-Shape-1 database. The idea of this experiment is to observe how well the SF technique is able to choose the images that belong to the same class of the query. Fig. 6(left) shows the average precision vs the number of random selected images. When our method is applied over a reduced set of images, precision is low. When the number of images increases the precision is also grown up because there are more possibilities of applying the similarity metric over a major number of shapes belonging to the image query's class.

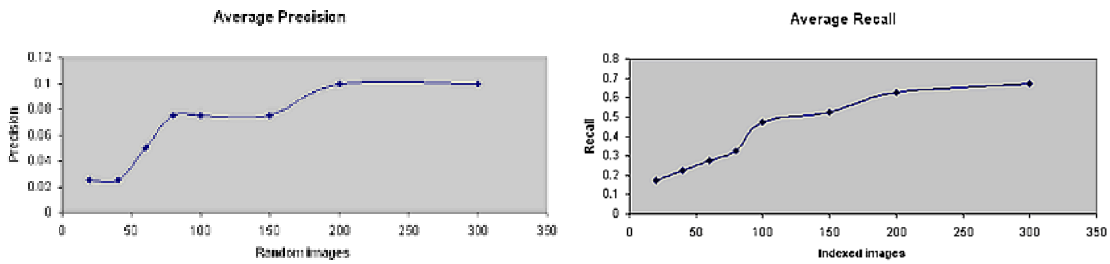

Figure 6. *Average precision and recall random images and for images with structural features*

Second experiment shown in Fig.6 (right) consists in applying our shape retrieval approach over a subset of images that were previously selected using some the eccentricity and solidity structural features.

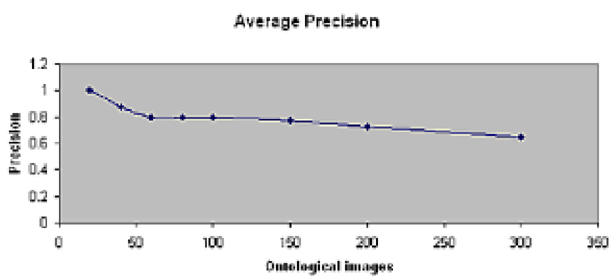

Figure 7. *Average Precision of experiment which uses ontological descriptions*

The third experiment consists in analysis of efficiency of ontological descriptors which allow reducing the semantic gap problem and reduce the number of iteration for VIR. Fig7. shows the precision with ontology which improves the quality of retrieval in case of low number of images maintaining the high precision (about 0.8) for other cases.

5 CONCLUSIONS

The most important contribution of this research is the proposed hybrid method combining the advantages of low-level image characteristics extraction with textual description of image semantics. Satisfactory retrieval of expected images is achieved faster due to the lower number of iterations in the searching process with ontology. The analysis of the indexing approaches shows that proposed Star Field approach is in order faster than 2STF. This occurs because the typical data structures used in indexing tools are hashing tables, which are manipulated with specific keys or signatures representing a shape. The disadvantages of the system are errors in spatial sampling during generation of the image feature vector as well as the required amount of system memory. The proposed image retrieval method is robust to partial occlusion and to small changes in the position of the objects. From the obtained experimental results, we can conclude that the method could be considered as an alternative way for the development of visual information retrieval systems.

References

Chavez, 2006 PhD Thesis. Star Field Approach for Shape-Based Image Retrieval: Development, Analysis and Applications, *Universidad de las Americas- Puebla*, Mexico

Emery J.C. 1987. 'Management Information Systems: The Critical Strategic Resource', Oxford Univ.

Excalibur corp. 2004. 'Excalibur visual retrievalware software developers kit', http://www.excalib.com

Fensel D. 2004. 'Ontologies: a Silver Bullet for Knowledge Management and Electronic Commerce', 2-nd Ed. Springer, USA.

Gupta A. and Malik A. 2005. 'Management Information systems (A computerized Approach to Managerial Aspects), FireWall Media.

Gruber T.R. 1993. 'A translation approach to portable ontology specifications', *Knowledge Acquisition*, 199-220.

Hiroshi H. S. 2005. 'Dynamically visual learning for people identification with sparsely distributed cameras'. *Lecture Notes in Computer Science*, 3540.

Jensen J.y R. and Hersh W. R. 2006. 'Manual query modification and data fusion for medical image retrieval'. *Accessing Multilingual Information Repositories*, 4022.

Kumar N. and Mital R. 2005. 'Management Information System'**,** Anmol Publications PVT. Ltd.

Medina, M. A. et al. 2004. 'A Novel hybrid method for image retrieval by ontological descriptions of sub-regions', *Journal WSEAS Transactions on Systems*, 4 (3): 1301-1306.

MIT Media Lab. 2003. 'Photobook', http://www-white.media.mit.edu/vismod/

Mokhtarian F. 1992. 'A theory of multiscale, curvature-based shape representation for planar curves', *IEEE Trans. On Pattern Anal. Mach Intell., (14) 8*: 789-805.

Smith J. and Chang S. 1996. 'Visual seek: a fully automated content-based image query system', *Proceedings of ACM Multimedia 96*: 87-98.

Starostenko O. et al. 2005. 'Novel algorithms for estimating motion characteristics within a limited sequence of images', *Recent Advances in Interdisciplinary Applied Physics*, Elsevier: 277-281.

Starostenko O. Cruz C. K. A 2007. Novel Shape Indexing Method for Automatic Classification of Lepidoptera, *XVII Int. CONIELECOMP*, Mexico, Publ. IEEE Computer Society , NJ , USA:1-6.

Virage Inc. 2004. 'Vir image & image read/write toolkit', http://www.virage.com/

Westphal M. 2004. 'Westphal. Imatch labs', http://www.mwlabs.de/