# Shape Indexing and Semantic Image Retrieval Based on Ontological Descriptions

Oleg Starostenko, Leticia Flores-Pulido, Roberto Rosas, Vicente Alarcon-Aquino
Universidad de las Americas-Puebla
Cholula, Mexico
`oleg.starostenko,leticia.florespo,roberto.rosas,vicente.alarcon@udlap.mx`

Oleg Sergiyenko
Autonomous University of Baja California
Mexicali-Ensenada, Mexico
`srgnk@iing.mxl.uabc.mx`

Vera Tyrsa
Polytechnic University of Baja California
Mexicali, Mexico
`vera-tyrsa@yandex.ru`

**Abstract**

This paper presents some hybrid approaches for visual information retrieval that combine image low-level feature analysis with semantic descriptors of image content. The aim of this proposal is to improve retrieval process by reducing nonsense results to user query. In the proposed approach user may submit textual queries, which are converted to image characteristics providing in this way searching, indexing, interpretation, and retrieval. In the case of visual query, both an image and sketch may be used. Approaches for image interpretation and retrieval are applied to color filtering, shape indexing and semantic. In order to assess the proposed approaches, some systems for image retrieval have been designed. The simplest system uses color region arrangement and neural network or wavelet based classifiers. Then this system has been improved using shape analysis with its indexing by ontological descriptions. For shape matching two proposed approaches are used such as star field or two-segment turning functions, which are invariant to spatial deformation of objects in image. The ontological annotations of objects in image provide machine-understandable semantics. The evolution of the proposed approaches and improvement of retrieval process are described in this paper. Four designed systems are assessed: RetNew, IRWC, Butterfly, and IRONS tested on standard COIL-100 and CE-Shape-1 image collections. The obtained results will allow to develop novel methods for solving efficient image retrieval processes.

## 1 Introduction

Recently, a quick increasing in amount and complexity of digital collections on web requires the development of new treatment and search engines as well as methods and algorithms for multimedia information retrieval. There are two different approaches for retrieval of relevant multimedia documents. The first one is based on the traditional techniques, which use a text as the main descriptor of images. This approach requires that images in digital collections must be previously indexed by textual labels. As a result, a human must annotate each image in collection manually and then use only this restricted set of images for searching. This is a very time consuming activity, and the fact that all images have a description does not guarantee that retrieved documents are relevant. The second approach provides the indexing and analysis of images based on their visual content. Usually, the low-level features of images such as color, texture, and shape are used. High-level methods try to recognize the particular patterns and interpret the content of whole image in order to make retrieval more efficient and reduce the number of iterations with nonsense results.

Among the visual information retrieval (VIR) systems that have been reported we have QBIC, VisualSeek, AMORE, SQUID, etc. The Query by Image Content system (QBIC) by IBM Almaden Research Center retrieves images, graphics and video data from online collections using color, texture and shape [1]. VisualSeek (Columbia University) is a web-based system, where the user requests images by description of spatial arrangements of color regions [2]. The AMORE (Advanced Multimedia Oriented Retrieval Engine by NEC USA Inc.) provides image retrieval in Web by keywords or by specification of a similar image [3]. The SQUID (Shape Queries Using Image Databases, University of Surrey) provides the image retrieval only by analysis of shapes in pre-processed database [4].

There are also other modern systems for VIR, namely, ImageFinder (Attrasoft Inc.) that uses probabilistic artificial neural network for similarity analysis of visual queries, MARS (University of Illinois) based on computing distance between two color histograms, texture, shape and indexed hybrid trees, Photobook by MIT Media Laboratory in Cambridge based on 2D shapes and texture image analysis computing similarity by amount of strain energy needed to deform one shape to match other, VP Image Retrieval System by University of Tokyo using silhouettes extracted by its decomposition into convex parts matching shifting query shape over the concatenated primitive signatures [2], [5].

Although the contributions of these systems to field of VIR were important, they do not provide mechanisms to represent meaning of images operating only at the low feature level. The analysis of factors like tolerance to deformation, robustness against noise, feasibility of indexing, presence of errors during spatial sampling, restrictions for input visual queries that must have small number of well-defined and separated objects without occlusion, sensitivity to scaling or rotation of analyzed regions, low precision of recognition if objects in image have week borders or complex background are the other factors that do not permit to develop high efficient VIR systems. In order to overcome some of these problems, we propose to apply novel approaches of color and shape indexing providing machine-understandable semantics for search, access, and retrieval of multimedia data using ontology. The results of our research may be used for development of image retrieval facilities such as, systems for supporting digital image processing services, design of software for high performance exchange of multimedia material within distributed collaborative and learning environments, distance education, digital libraries; design of searching facilities for personalized information among federated collections; image based navigators, etc.

The remainder of this paper is organized as follows. In Section 2 the analysis of well-known methods for image segmentation and retrieval is presented. In Section 3 the four designed systems are presented. Section 4 reports a description of experiments, evaluation and comparative analysis of the proposed and well-known content based image retrieval (CBIR) systems.

## 2   Low and high level feature extraction and generation of feature vector.

The VIR includes some processes applied to queried image such as image pre-processing, low-level image feature extraction and generation of feature vector, computing similarity of that feature vector to other ones from processed image collection and finally, retrieving images with high grade of similarity to user query. There are some paradigms of image processing used in VIR that may be subdivided into two groups. The first group is based on analysis of low-level characteristics such as, pixel position, color, texture, shape, region distribution, etc. [6], [7]. However, these approaches generate results that are not useful for image comparison because spatial image variations make impossible efficient image matching. The second group of methods retrieves images after recognition of image content using models based on symbolic representation of images [8], probabilistic techniques for analysis of image content, perceptual grouping of lower-level image features into a meaningful higher-level interpretation [9] based on fuzzy reasoning [10] and semantic Web approaches [5]. The advantage of these approaches consists

in speeding up retrieving process due to the reduced number of iterations required during image seeking and matching.

The most widely used image pre-processing technique is based on color feature vector that does not depend on image size or orientation and provides visual document segmentation into regions with similar colored characteristics. The most useful techniques used currently in CBIR are segmentation in RGB or HIS color space, histogram processing, color slicing, tone and color correction [11]. The advantage of color feature is the reduction of data amount being processed during matching the queried and reference images in digital collection.

We also focused on shape of objects in images because it has a meaning by itself. However, shape matching is considered as one of the most difficult aspects of CBIR because a shape needs many parameters to be described explicitly. It is possible to distinguish three main categories of techniques that represent shape: the feature vector approach, the relational and the transformation approaches [12]. The feature vector approach represents a shape as a numerical vector, and the similarity between two shapes is calculated by using metrics like Euclidian or Hausdorff distance. In the relational approach a shape is broken down into a set of salient components; the total descriptor is made up of descriptors of each salient components and the relation among them. Finally, the transformation approach is used to describe a shape based on the effort needed to transform one shape into another. The well-known methods for global shape description such as Elasticity correspondence, curvature scale space approach, B-splines and chain case codes, Fourier and heuristic descriptors sometimes are too complex for fast processing and sensitive to spatial variations [13]. In this paper, we propose some novel approaches for shape indexing and matching, particularly, methods based on Two segment turning function (2STF) and star field (SF), which have been implemented in designed CBIR systems improving retrieval process.

One way to solve problems of low-level feature based systems is to use artificial intelligence methods facilitating knowledge sharing and reusing. It permits to promote the development of machine-understandable semantics for representation and exchange of visual information. This alternative is known as ontology approach. There are several definitions of ontology but the more acceptable is the Grubber's one [14], who establishes that ontology is a formal, explicit specification of a shared conceptualization. In VIR applications, ontology permits to describe 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. The ontology is described by a directed acyclic graph in which each node has a feature vector that represents the concept associated with that node. Concept inclusion is represented by IS-A inter-relationship made by manual indexing of preprocessed images. For example, particular elements of buildings, churches, etc. correspond to specific concepts of shapes defining these buildings. If a query describes an object using this ontology, the system would recover shapes that contain windows, columns, facades, etc. even though, those images have not been labeled as geometric figures for the retrieved object. In order to support the ontology management, the Resource Description Framework (RDF) language has been used. It defines a syntactic convention and a simple data model to implement machine-readable semantics [15]. Concluding, we believe that the proposed hybrid approach combining color, shape and ontological descriptors is a promising technique that permits to develop high performance CBIR systems.

## 3   Evolution of image retrieval approaches in the designed systems

After analyzing well-known systems and approaches used for pre-processing, indexing, classification, matching and retrieval of images in digital collections, we conclude that the basic low-level image features are color, texture, and shape. Some CBIR systems have been designed to improve performance and retrieval parameters of well-known CBIR systems such as, processing speed, grade of similarity to

input visual query, low number of iteration in retrieved process and simplicity in description of image semantics.

## 3.1    Image Retrieval by Neural Network (RedNew) System

The first VIR system designed by us is the prototype called Image Retrieval by Neural Network (RetNeu) system shown in Fig. 1. The system consists of two channels: the first generates low-level feature vector of input image while the second retrieves the centroid from pre-processed image collection with similar features. The centroid is the generalized image representing class or group of images in collection with similar content providing in this way the distribution of images in collection on base of their meaning. This distribution process of images and organization of collection is made by using the Resonance Adaptive Theory ART2 neural network with three layers and 12 input and 20 output nodes. In Table of Fig. 1 the input features extracted by system for ART2 are presented. The variables ra and rb are the largest and the smallest radiuses of segmented region respectively and the variable r is the radius of circular region.



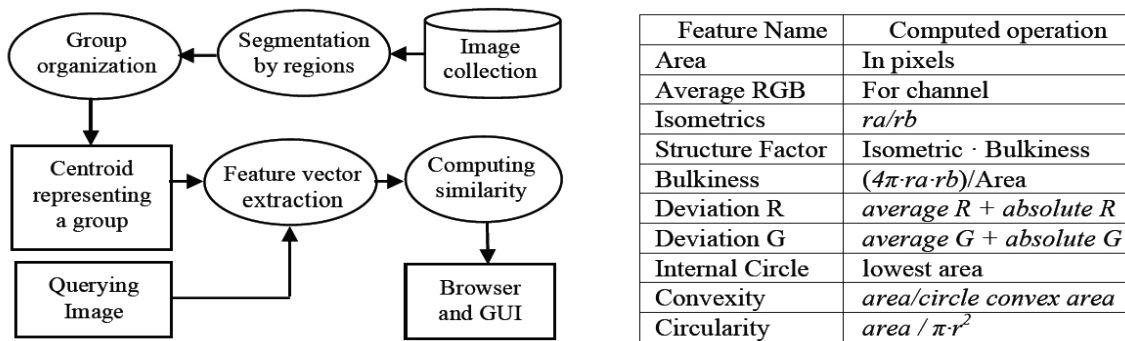| Feature Name | Computed operation |
|---|---|
| Area | In pixels |
| Average RGB | For channel |
| Isometrics | $ra/rb$ |
| Structure Factor | Isometric · Bulkiness |
| Bulkiness | $(4\pi \cdot ra \cdot rb)$/Area |
| Deviation R | $average\ R + absolute\ R$ |
| Deviation G | $average\ G + absolute\ G$ |
| Internal Circle | lowest area |
| Convexity | $area/circle\ convex\ area$ |
| Circularity | $area\ /\ \pi \cdot r^2$ |

Figure 1: Block diagram of RetNew VIR system and list of image features that compose a feature vector.

The image preprocessing in RetNeu implies image segmentation, low-level-feature extraction, computing a corresponding centroid, and calculation of similarity measure as it is shown in Fig. 2. The image segmentation and feature extraction process is based on YIQ decomposition and region growing [16].

An image is divided into small windows, for example, 8x8 pixels and the average of window color is assigned to a single pixel in new smaller image. The set of the pixels with similar color is interpreted as a particular region. The main descriptor of each region is generated according to a set of color parameters
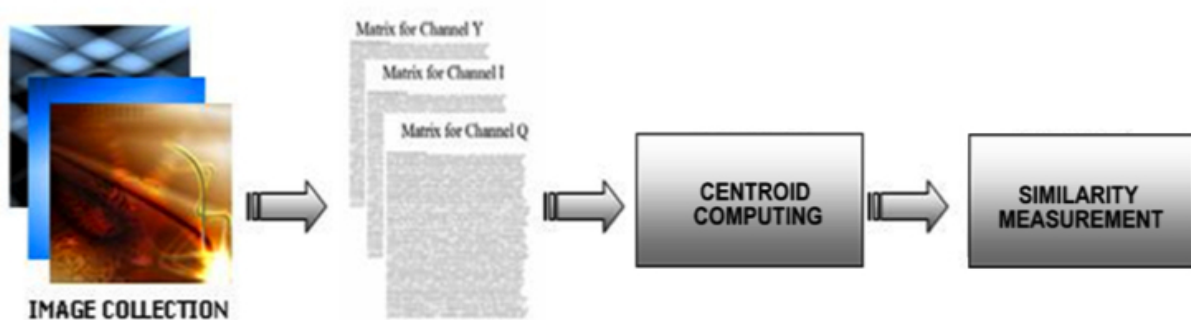


Figure 2: Image retrieval process in RetNew system.

61

and low-level features, which are used as input of ART2. According to the results obtained by neural network, the processed image is indexed by corresponding parameters that define to which group belongs this image, for example, to group of buildings, maps, animals, landscapes, etc. The input visual query that may be thumbnail image, sketch, or shape is converted to its feature vector and then, the computing similarity is provided by matching this vector with all centroids in collection. Using Jacobs' metrics shown in eq. 1, the highest grade of similarity between feature vectors defines the group of images that may be candidates for retrieval [17].

$$\| Q,T \| = w_{0,0} \mid Q[0,0] - T[0,0] + \Sigma_{i,j} w_{i,j} \mid Q[i,j] - T[i,j] \mid \tag{1}$$

where Q represents the feature vector of a querying image, T is the centroid corresponding to particular class of image collection, and w(i,j) is semantic weight associated with a class of images to which centroid belongs in T. When the best correspondence with one or more centroids is found a comparison of the feature vectors of images within selected class is applied to find images candidate for retrieval. Thus, RetNeu system has two steps of matching: the first one consists in selection of a class or centroid with similar to queried image content, and the second one includes a selection of the most similar images within the class using low-level features of image. Therefore, in the indirect manner RetNeu system retrieves images using not only low-level image features but also takes into account a meaning of visual query. The disadvantages of the proposed approach implemented in RetNeu system are the limited number of low-level features used for image indexing, significant time that it takes for organization of collection with pre-processed reference images, which additionally requires the user feedback for textual description of image semantics, necessary training of neural network and double image matching process.

## 3.2 Image Retrieval by Wavelet Coefficients (IRWC) System

The improvement of RetNeu System has been done in another proposed Image Retrieval by Wavelet Coefficients IRWC System [16]. It is based on only YIQ color decomposition of image reducing significantly a set of images to be processed, and the neural network of RetNeu is substituted by wavelet transform (Haar, Daubechines, Biorthogonal wavelets, and Symlet transforms have been evaluated) reducing two-steps image matching process to one step of feature vector comparison. The best results have been obtained with the Symlet transform. It is similar to Daubechines wavelet but it provides the best symmetry and may be obtained by the following equation

$$Symlet(x) = e^{\phi(i)\omega} \tag{2}$$

where x is input image, $\phi(i)$ is Daubechines transform for each pixel i, and $\omega$ is the moment of an image. The block diagram of the designed system is shown in Fig. 3. The input query is converted to color regions extracted by YIQ decomposition generating a feature vector that consists of wavelet coefficients representing those regions. Computing similarity of queried image to pre-processed images in collection is provided by methods adopted in VIR such as, Jacobs', Euclidian and Q.Tian metrics. Finally, the retrieved images are presented on user interface with the corresponding degree of similarity. The best similarity metrics was Euclidian distance due to more precise comparison of each characteristic in feature vectors.

The Euclidian distance between feature vectors, particularly, between wavelet coefficients is com-
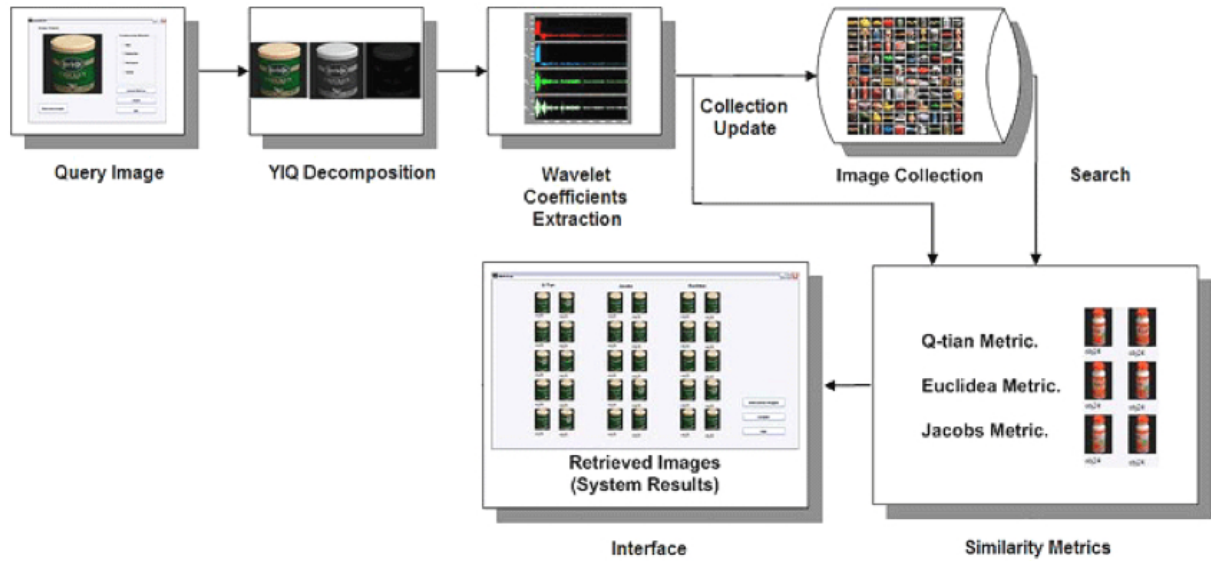
Figure 3: IRWC system for image retrieval using wavelet characteristics.

puted by applying the following:

$$\| P, Q \| = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2} \tag{3}$$

where P and Q represent the wavelet characteristics of pre-processed image from collection and input querying image respectively. The Symlet transform and Euclidian similarity metrics have been finally adopted in proposed the IRWC system. In Fig. 4 the user interfaces of IRWC system is presented. It is possible to note that there are some images, which do not correspond to querying image reducing in this way the performance of used approach.

The advantages of wavelet-based approaches consist in better convergence of results due to the big length of feature vector providing more precise comparison, symmetry and regularity useful for image processing with presence of noise.

### 3.3   Shape Indexing System for Automatic Classification of Lepidoptera

Shape Indexing System for Automatic Classification of Lepidoptera (Butterfly) is another example that shows how to improve CBIR process. This system has been designed for classification of rare specimens of Lepidoptera (butterflies) in Mexico. The color filter and shape descriptors based on tangent space are used for butterfly indexing. That establishes the relationship between object and its explicit definition in textual form using domain of butterfly?s specimens. The block diagram of the system is shown in Fig. 5.

The input of the system may be an image, its shape, manual sketch or a keyword, which describes an object. The retrieved images with classification of butterfly?s specimen will be those, which have more similarity to the low-level features of a query. Once user generates input query, the system applies color filter for reduction of set of images candidates for retrieval. Empirically, the threshold for color variations has been defined on the level about 200 for images with 4000 colors. The shape indexing unit generates feature vector described silhouette of butterfly. The shape is presented by polygon, where each linear segment is labeled by corresponding length and angle. Therefore, a polygon composed by segments may be presented as a step-case function in tangent space, where axis x represents a length of each segment
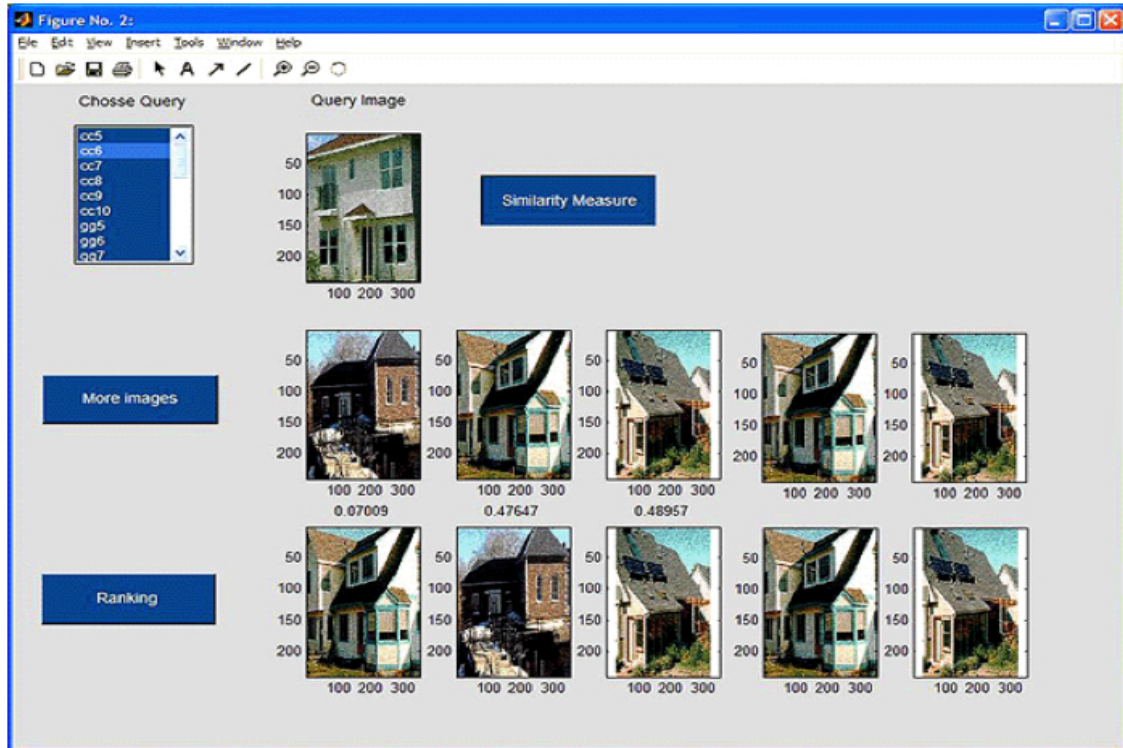
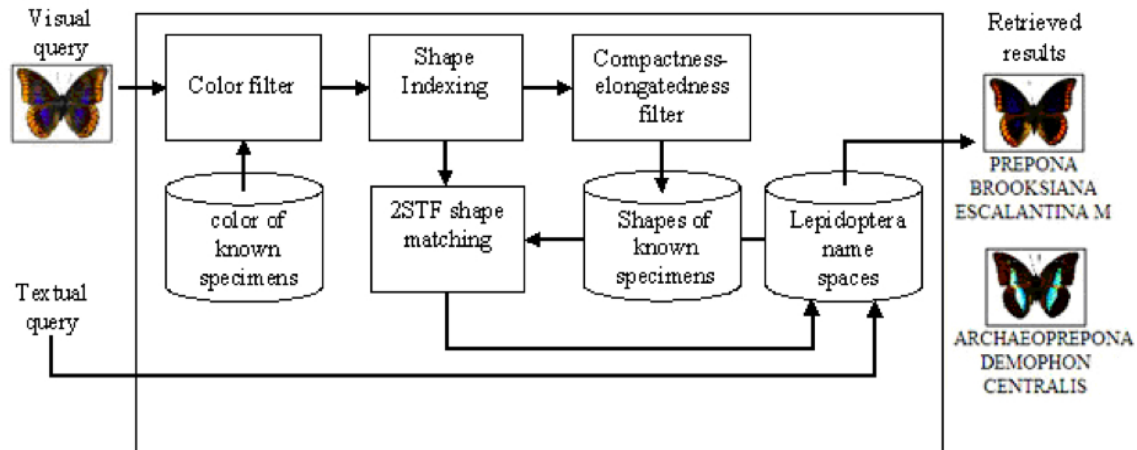Figure 4: GUI of IRWC system for image retrieval



Figure 5: Block diagram of the proposed system for classification of Lepidoptera specimens.

and axis y represent an angle of a segment [5]. The disadvantage of this approach is the sensitivity to spatial distortions and rotation of a shape. To solve this problem, the two segment turning function (2STF) has been proposed. It is described in the step 4 during discussing the proposed approach [19]. Matching of shapes is done by comparison of 2STFs of a query and pre-processed images in collection after applying the additional filter based on compactness and elongatedness. In case, when the user's query is a textual description of particular specimen only a keywords matching is applied to Lepidoptera name space. This collection is connected to another one with Shape and images of known specimens.

When a textual query is found in Lepidoptera name space the image corresponding to this description is retrieved.

The proposed approach for indexing and retrieval consists of the following procedures:

1. The spatial sampling of the original image is provided by computing the average values of color regions via the I1I2I3 color model generating the color descriptor of each region.

2. In order to define butterfly shape the SUSAN corner detection method is used [18]. The extracted principal corners present the points that define butterflies by polygonal shape.

3. Using discrete curve evolution, a set of vertices of complex polygon is reduces to a small subset of vertices containing relevant information about the original outline [19]. The relevance measure $K$ is computed as follows

$$K(s1,s2) = \frac{\beta(s1,s2)l(s1)l(s2)}{l(s1)+l(s2)} \qquad (4)$$

   where $\beta(s1,s2)$ is the turning angle at the common vertex of the segments $s1, s2$ and $l$ is the length function normalized with respect to the total length of polygon. The lower value of $K(s1,s2)$ corresponds to less contribution to a shape of the curve of arc $s1 \cup s2$.

4. The next step is indexing of the simplified shape by its transformation to 2STF, because the polygonal representation is not a convenient form for calculating similarity between two shapes. We propose to compute the similarity between two shapes comparing their corresponding polygons represented by the cumulative angle function or turning function. The step turning function of a polygon A represents the angle between the counterclockwise tangent and the x-axis as a function of the arc length $s$ $\theta_A(s)$. This function is invariant to translation and scale of a polygon, but it is sensitive to rotation. To overcome this problem, we propose 2STF, where the y-value of step function must be now directional angle of the linear segment with respect to its previous segment. Fig. 6 a) shows the angle between two consecutive linear segments, and how to adjust the turning angle between those segments.
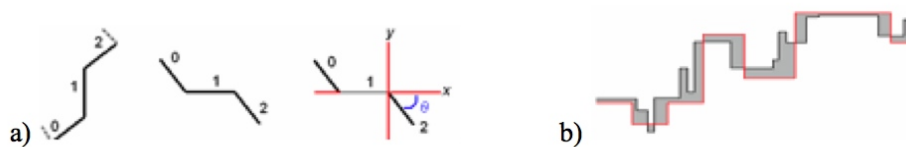


Figure 6: a) adjustment of the turn angle between two consecutive linear segments; b) matching strategy for computing the similarity between two polygons using 2STF.

5. For comparison of two polygons the scaling 2STFs to the same length is done by factor $sf = \frac{l(a_1)}{l(a_2)}$, where $l(a)$ is the length of curve $a$ and $l(a1) > l(a2)$. Then the curve $a_2$ is shifted as it shown in Figure 6 b), and the area between two 2STF curves is computed. The shady area represents how similar two arcs are. Additionally, compactness and elongatedness filter reduce number of images to be compared and provides fast and satisfactory image retrieval by accelerating the convergence to expected result.

6.  Finally, it is possible to establish the relationship between butterfly and its formal explicit definition by textual descriptions found in domain of particular butterfly specimen.

The user may submit a query via GUI shown in Fig. 7 where searching results are presented with textual description. For reasonable number of presents retrieved images, the matching of query must be more than 80%.
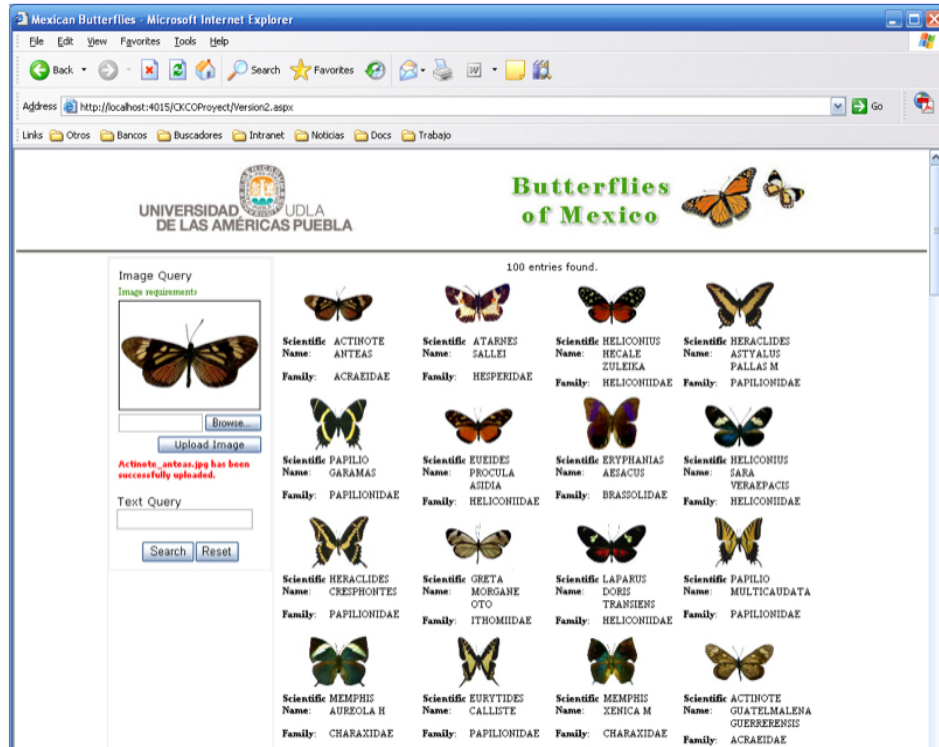


Figure 7: GUI of the system for classification of butterfly's specimens.

The disadvantages of system are the errors in spatial sampling during generation of image feature vector, application of color thresholds, the additional time for computing of shape elongatedness and compactness, as well as the required amount of system memory. However, the proposed image retrieval method is invariant to rotation, translation, scale, and partial occlusion of objects in image.

### 3.4 Image Retrieval by Ontological Description of Shapes (IRONS) System.

Image Retrieval by Ontological Description of Shapes (IRONS) System has been designed in order to improve indexing and shape matching process extending also the semantic aspect of image retrieval. The disadvantage of 2STF representation is the significant time that it takes to find the best correspondence between two curves. Experimentally, we determined that comparison of shapes with more than ten arcs takes more than 10 seconds. However, the advantages of 2STF are simple implementation and independence from scale, reflection, translation, and rotation of objects in image. The problem of slow matching process may be solved by the proposed technique called Star Field (SF) representation of shapes that allows comparing polygons with any number of arcs without significant increment of time. Our Star Field method combines 2STF with the maxima of curvature zero-crossing contours of Curvature Scale Space (CSS) as a feature vector to represent shapes [4], [20].
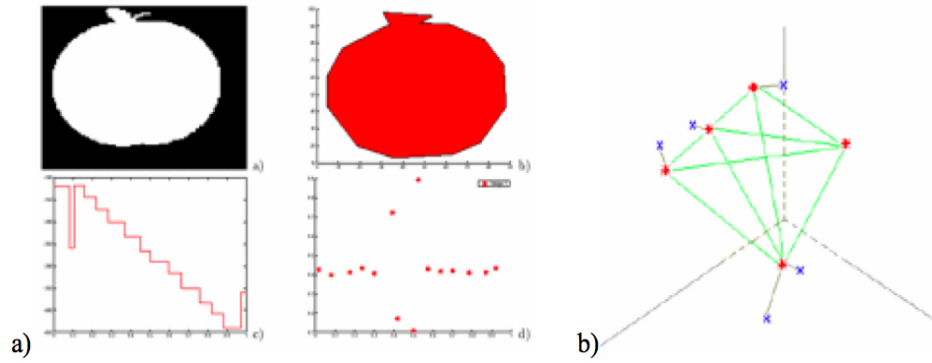
Figure 8: a) original image, its 15-segments polygon, and corresponding 2STF and SF, b) SF matching graph.

Formally, a SF representation is a set of marks or stars $M1; M2 ::: Mnm$, where $nm$ is the number of vertices of polygonal curve that it represents. $Mn$ is defined by means of two coordinates $(x; y)$. The $x$ - coordinate indicates the normalized distance from the starting point to the corresponding vertex, making sure that in the middle of the SF plane is the star that corresponds to the most important vertex of the polygon. The $y$ - coordinate is the normalized angle between two consecutive segments that share the corresponding point as it is shown in Fig. 8 a).

For computing similarity between two SF the graph and adjacency matrix are used. Given two polygonal curves $P1$ and $P2$ and their star field representations $SF1$ and $SF2$, 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 $V$ the vertex set and $E$ the edge set of $G$. Our particular graph $G$ has a set $V$ that consists of two subsets of vertices $v_1$ and $v_2$. Set $V = v_1 \cup v_2$, where $v_1$ is the set of point of $SF1$ and $v_2$ is the set of points of $SF2$. On the other hand, $E$ is the set of pairs $(r; s)$, where $r \in v_1$ and $s \in v_2$ Then 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, which includes every vertex, 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 (distances between corresponding stars in matching graph, see Figure 8 b).

The block diagram of IRONS system is shown in Fig. 9. The input for the system may be an image, its shape, or a keyword, which describes the object in image to be retrieved. The retrieved images will be those, which have a high degree of matching in color/shape and ontological annotations defining the content of a querying image [5].

The system consists of low-level feature extraction module used for visual queries and high-level feature module for textual ones. The previously pre-processed images and their manually defined descriptions are organized in specific ontological structure of trees that implies. They are stored in collection of classified images and in ontological namespace respectively. The feature vectors of each node in the ontology name space consist of keywords linking the previously classified images to the characteristics of new shape represented by SF. In this 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. Additionally in the IRONS sys-
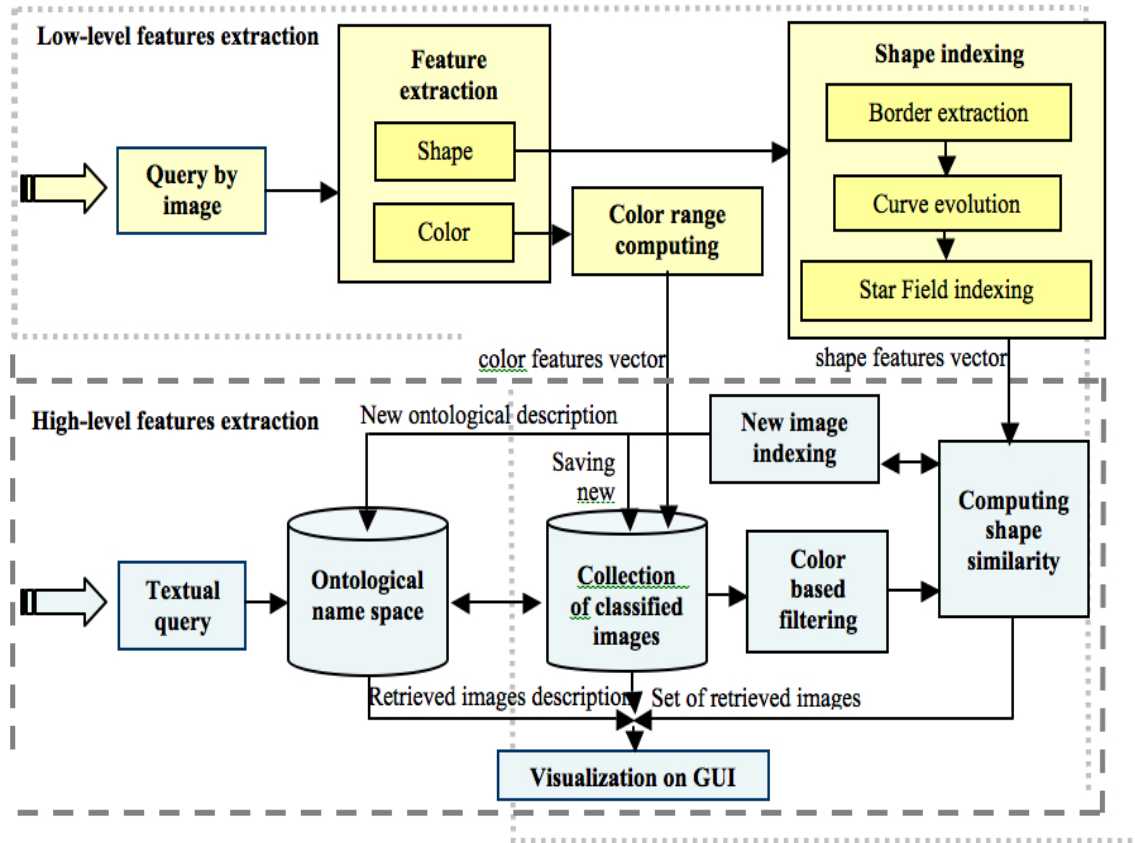
Figure 9: IRONS system for shape indexing using ontological descriptions.

tem a user may define his proper user-oriented annotations for new input images or sketches adjusting the ontological name space according to his perception of objects and scenes in image. The indexing and ontology annotation processes may be quantified computing the final grade of similarity $d_{FG}$ using color/shape and ontological description feature vectors in the equation of Euclidian distance as

$$d_{FG}(d_T d_O = \sqrt{\sum [(d_{T1} + d_{O1}) - (d_{T2} + d_{O2})]^2} \qquad (5)$$

where $d_{Ti}$ and $d_{Oi}$ denote total color/shape and ontological description vectors of i-th image respectively. The ontological annotation tool is used for searching matches in the ontology name space. The images with higher matching are retrieved and visualized on GUI with a certain degree of similarity as it is shown in Fig. 10.

The proposed approach provides speeding up the matching process as well as reduces the number of iterations with nonsense results using ontology. The analysis of the indexing approach shows that SF is at least in one order as fast as 2STF. This occurs because the typical data structures used in indexing tools are hashing tables, which are manipulated fast with specific keys or signatures representing a shape.

## 4 Experiments and discussions.

Some experiments with the designed systems have been done with the proposal to evaluate novel approaches and define the next steps in development of CBIR. For analysis of system performance and
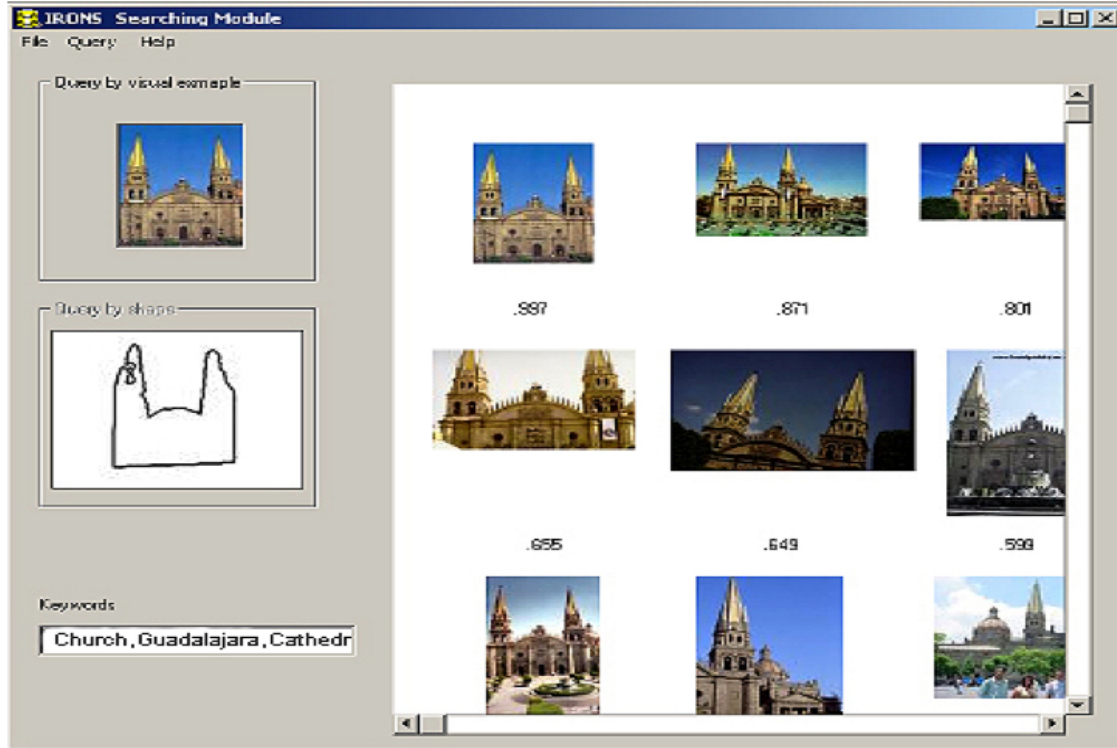
Figure 10: GUI of IRONS system.

efficiency of matching process the experiments have been divided in two groups. The first one is Candidate Images Selected by Different Similarity Metrics experiment, which consists in applying Jacobs', Euclidian, and Q.Tian metrics over some sets of previously classified standard collections of images (Columbia Object Image Library collection COIL-100, and CE-Shapes-1). Each experiment consisted in random selection of 100 querying images from 72 different classes in COIL collection of 7200 images and 50 images from 60 classes in CE-Shape-1 collection of 1400 images. The purpose of this experiment is to observe how well used matching techniques are able to choose relevant images from collection. In Table 1 the results of experiments for designed systems comparing them with some prototypes are presented. The reported retrieved accuracy describes ability of system to find relevant images in the worst and the best cases, which depend on the used approaches for feature vector generation, image indexing and matching strategy. The second Candidate Relevant Images experiment consists in evaluation of the retrieval process in the proposed systems. The evaluation of VIR system is a non-trivial task. This is because there is subjectivity involved into query interpretation by the user. Nevertheless, there is a standard way of judging the obtained results. This technique consists in calculating two metrics known as recall and precision. The recall measures the ability of a system to retrieve relevant information from the whole collection. The precision is the ratio between a number of relevant retrieved images and total number of relevant images in the collection.

$$Precision = \frac{A}{B}; Recall = \frac{A}{C} \tag{6}$$

where A is a set of relevant images retrieved by system, B is a set of relevant and irrelevant images retrieved by system for particular query, and C is a set of all relevant images in collection for particular

Table 1: Characteristics of CBIR systems and reported retrieval accuracy.

| System source | Feature vector | Metrics type | Shape indexing | Matching technique | Semantic analysis | Corpus size | Retrieval accuracy |
|---|---|---|---|---|---|---|---|
| SQUID [17] | shape | Euclidian | circularity, curvature Scale Space | CSS matching with Gaussian function | no | 1100 | 68-85 % |
| CIRES [9] SQUID | color, texture, structure | Euclidian | no, low level, structure element | feature wighting and Gaussian normalization | perceptual grouping of structures | 10221 | 58.5-87.4 % |
| RetNew [5], [16] | color, regions, | Jacobs | no, color region decomposition | centroid selection neural networks for low level features | centroid organization | 7200+ 1400 | 55-88.3 % |
| IRWC [5] | color, region wavelet symlet | Jacobs Euclidian Q. Tian | no, color region decomposition | wavelet coefficiens | semantic grouping | 7200+ 1400 | 78-92.5 % |
| Butterfly [18] | color shape | Euclidian | Tangent space and 2STF | shape and textual descriptor matching | textual descriptions | 140 in 15 classes | 77.5-93 % |
| IRONS [13], [19] | color, shape ontology | Euclidian | Star field ontological descriptions | shape, ontological descriptions | ontology | 7200+ 1400 | 82-95 % |

query

Fig. 11(a) and (b) show the average recall and precision in the experiments on RetNeu and IRWC systems respectively. The x-axis represents the number of tests with 100 images in 20 classes of collection; y-axis shows the average recall and precision computed according to eq. 6. Taking the results presented in Fig. 11, the improved IRWC systems provides better retrieval. In RetNeu system for different sets of images the recall/precision values lie in the range about of (0.1 over 0.9)/(0.06 over 0.4) respectively. In IRWC system the average recall/precision have higher values, which lie in the range of (0.3 over 1.0)/(0.1 over 0.55) respectively. For the IRONS system, the Candidate Images Selected Using an Image Ontology experiment has been done to observe how well the proposed approach is able to choose the images without and with ontology. In Fig. 12(a) the results without ontology are presented on GUI. In the retrieved set there are images from different classes such as fruits, trees, geometrical patterns, etc., which have similar color and shape features with query (image in left-upper corner). The application of ontology allows reducing semantic gap problem retrieving relevant images only from the same class to which the particular query belongs as it is shown in Fig. 12(b) (query image is in the left-upper corner).

Experiments for the IRONS system without ontology gives recall/precision in the intervals about (0.5 over 0.9)/(0.28 over 0.4) respectively. In the experiments applying the ontological descriptions the recall/precision metrics lie in the range of (0.7 over 1.0)/(0.35 over 0.65) respectively. The evaluation of
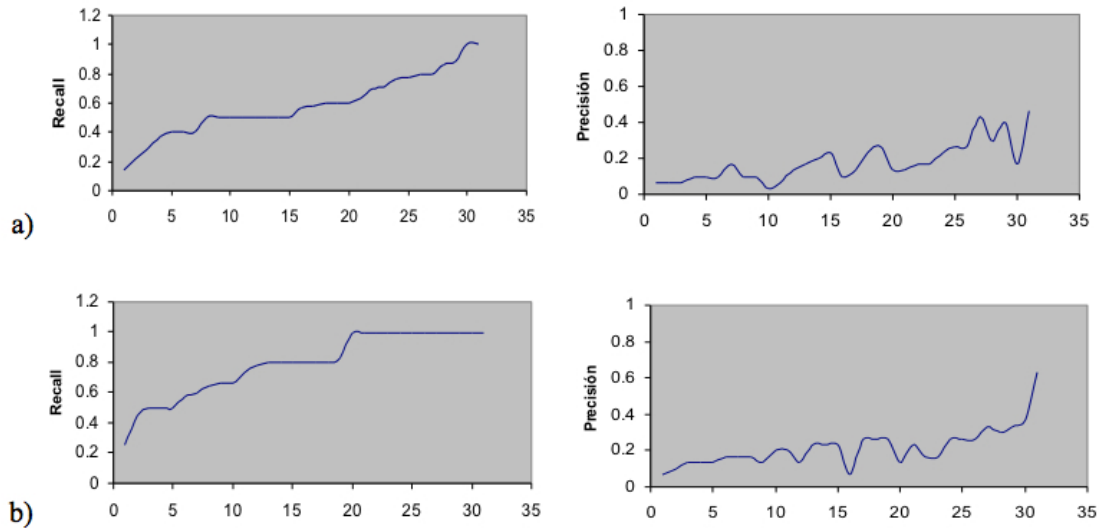
a)

b)

Figure 11: a) Recall and precision metrics for a) RetNew and b) IRWC systems.
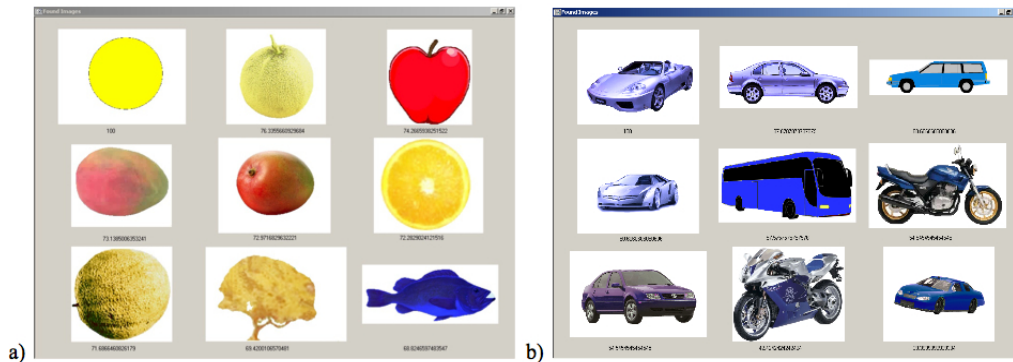


a)                                          b)

Figure 12: Retrieved images a) using color/shape feature vector, b) additionally, using ontology.

the proposed approaches and testing the designed systems show the ability of them to retrieve relevant images from different collections is satisfactory and achieve up to 88-95 % of retrieval accuracy as it is shown in Table 1. Using semantic aspects of CBIR, the IRONS system with ontology provides correct retrieval of expected images without nonsense results and makes it faster due to the lower number of iterations in a searching process. The found images belong to the same class as a query that may be seen in Fig. 12(b). The disadvantages of the system are the presence of errors during image segmentation applying spatial transforms, generation of feature vectors and the restrictions for input visual queries, which must have small number of well-defined and separated objects. Additionally, significant occlusions between objects, week borders or complex background in image, noised or incomplete images are not recommended in this application. From the obtained experimental results we conclude that the proposed approaches could be considered as alternative way for the development of visual information retrieval facilities.

# 5   Conclusion

The evaluation of the proposed approach and testing of the designed systems show that the hybrid feature vector composed by color, shape and ontological descriptors provides quite acceptable image retrieval from standard collections. The conceptual contribution of this paper is to introduce and to evaluate the proposed approaches for image indexing and retrieval using low-level image features and semantic aspect of images. The tangent space for shape indexing has been improved by the proposed 2STF and SF speeding up the matching process invariant to shape partial occlusion, rotation, and scaling. Satisfactory automatic retrieval of relevant images is achieved faster due to the lower number of iterations in a searching process with ontology providing textual description of image semantics. The ontological annotations allow simple and fast estimation of meaning of a shape and sometimes of a whole image. Application of diverse metrics for the evaluation of system performance confirms that the proposed approaches are robust enough to be used for efficient image retrieval. The analysis of factors like tolerance to deformation, to complex images with significant number of objects, to presence of noise, and feasibility of indexing are considered as possible extension of the proposed approaches. The practical contribution of this paper consists in the description of four novel designed systems, evaluation of their performance and analysis of improvement of well-known CBIR systems solving still open problem of VIR.

# References

[1] QBIC (TM), Motion estimation algorithms of image processing services for wide community, http://wwwqbic.almaden.ibm.com/, 2003

[2] R. C. Veltkamp, M. Tanase. Content-Based Image Retrieval Systems: A Survey: http://give.lab.cs.uu.nl/cbirsurvey/, last viewed October 2010.

[3] The Amore, Advance multimedia oriented retrieval engine, http://www.ccrl.com/amore/, 2003.

[4] S. Abbasi, SQUID, http://www.ee.surrey.ac.uk/CVSSP/, last viewed 2010

[5] L. Flores-Pulido, O. Starostenko, I. Kirschning, J. Chávez-Aragón. Wavelets vs Shape-Based Approaches for Image Indexing and Retrieval, in the book Novel Algorithms and Techniques in Telecommunications, Automation and Industrial Electronics, Sobh, T. (Eds.), Springer, 2008.

[6] A. Chávez, O. Starostenko, Image Retrieval by Ontological Description of Shapes (IRONS), Proc. 1st Canadian Conf. on Computer and Robot Vision (CRV), 2004, pp 341-346.

[7] T. Gevers, Classification on Color Edges in Video into Shadow-Geometry, Highlight or Material Transitions, IEEE Trans. On Multimedia. 2002.

[8] Virage Autonomy Systems, http://www.virage.com/content/, 2006.

[9] Q. Iqbal, CIRES, http://amazon.ece.utexas.edu/ qasim/research.htm, 2007

[10] B. Kovalerchuk, J. Schwing, Visual and Spatial Analysis. Advances in Data Mining, Reasoning, and Problem Solving, Springer, USA, 2004.

[11] W. Burger, M Burge, Principles of Digital Image Processing: Core Algorithms, Springer; 1st Edition, 2009.

[12] R. C. Gonzalez, R.E. Woods, Digital Image Processing, Prentice Hall, USA, 2008.

[13] O. Starostenko, J. A. Chávez-Aragón, G. Burlak, R. Contreras. A Novel Star Field Approach for Shape Indexing in CBIR System, J. of Engineering Letters, vol. 5, Is.2, 2007, pp. 287-295.

[14] T. R. Gruber, A translation approach to portable ontology specifications, J. Knowledge Acquisition, 1993, pp. 199-220.

[15] D. Beckett, The design and implementation of the Redland RDF application framework, Proc. 10th Int. WWW Conf., 2001, pp. 120-125.

[16] L. Flores-Pulido, W. E. Estrada-Cruz, J. A. Chavez-Aragon. An Image Retrieval System based on Feature Extraction for Machine Vision Using Three Similarity Metrics, Proc. 23-rd ISPE Int. Conference on CAD/CAM Robotics, Colombia, 2007, pp. 832?837.

[17] A. Jacobs, Fast Multiresolution Image Querying. SIGGRAPH, Conf. N.Y.1995

[18] O. Starostenko, C. K. Cruz, A. Chávez-Aragon, R. Contreras. A Novel Shape Indexing Method for Automatic Classification of Lepidoptera, Proc. of XVII Conielecomp, Mexico, 2007, pp.1-6.

[19] A. Chávez-A., O.Starostenko, L. Flores P., Star Fields: Improvements in Shape-Based Image Retrieval, J. Research on Computing Science, Mexico, Vol. 27, 2007, pp.79-90.

[20] F. Mokhtarian, A theory of multiscale, curvature-based shape representation for planar curves, IEEE Trans. On Pattern Anal. Mach Intell. 14, 8, 1992, pp. 789-805.