Content based Image Retrival

S.Dharani,
Research Scholar,
St.Peter's University,
Chennai-54

B.Justus Rabi
Professor,
Dr.M.G.R. University,
Chennai-95

A.N.Nanda Kumar,
Professor,
City Engineering College,
Bangalore- 560062

Joe Patrick Carmel
III Year Student (Mech)
Dr.M.G.R. University,
Chennai-95

ABSTRACT
In this paper, we propose a content-based image retrieval system which takes into account the local attributes of the image for defining the feature space. This method presents a way to localize the characteristics of the queries by partitioning the image into a rectangular grid and applying a different feature vector to each region in the similar measuring phase. The assignment map specifying the feature space for each image region is implicitly selected by the user, through the system interface, according to perception of the content. This is the most important aspect of the system, which provides flexibility to the user to query at the object level by selecting the "type" of the regions. User only intervention to the system is at this phase, but the way of preparing the query, directs the system's similarity calculation in the later stages. The experiments indicate that the proposed system yields better results for images having distinctive objects compared to global systems using the same features for the entire image.

General Terms
Image Processing, Feature Extraction, Image Segmentation, Object Recognition.

Keywords
DataBase Generation Subsystem, DataBase Retrieval Subsystem, Feature Matching Subsystem.

.

1. INTRODUCTION

In recent years, the rapid increase in the number of images in various databases has led to a parallel increase in the public demand of image retrieval systems having preferably content-based search capabilities. This phenomenon led to the implementation of many content-based image retrieval systems [1, 2, 3]. There are many problems faced in designing a retrieval system. The most basic one is to measure the similarity in terms of content. At this point, one of the two approaches can be adopted which are global image properties and object level semantics into account [1]. The first approach employs major global analysis methods like color histograms or texture descriptors. MPEG-7 [8,9] provides successful descriptors for this approach and presents tools for such global analysis of the image content. Research on the second approach is quite rare due to the inter- dependency of object recognition and feature extraction steps. Without the object recognition, it is not possible to define a consistent set of features, while the specification of the feature space is a necessary preliminary step for object recognition. The current image retrieval systems are not capable for querying image collections at object level [4]. Research on filling the gap between low level feature descriptors and object level semantic information is an important issue in content based retrieval systems. In this paper, we present a retrieval system, which enables the user to query at object level in a simplified manner. We delay the problem of automatic segmentation of objects and let the user prepare a query according to their perception so that the system will focus on some specific region, containing the object of interest, and use the 'appropriate' low level image feature descriptors for the object. This is called as "localization" of the query. In this way the misleading affects of using the same feature vector for the entire image is avoided. The two major challenging problems in this approach are;

i) To partition the image area so that preliminary indexing and feature vector assignment to individual regions are possible.
ii) To use the similarities of these partitions to calculate an overall similarity for a given image pair.

In the proposed system, the image is partitioned into a rectangular grid and the feature vector assignments are implicitly supported by the user via an interface. Thus, the proposed system is quite sensitive to the user's intention and expectation in the very beginning of the retrieval process, in contrast to recently developed adaptive systems [5, 6]. The following sections (2, 3, 4 and 5) explain the major steps of the proposed method, regarding the image area partitioning, preparation of the query, similarity calculation and feature extraction issues. Section 6 presents the experiments on an image collection, selected from Corel Image Gallery. Section 7 describes the over all structures of content based image retrieval model. Finally, in section 8, there is a discussion about the weaknesses of the proposed system and possible improvements are suggested for future work.

2. PARTITIONING THE IMAGE AREA
Segmentation seems to be the most appropriate method for extracting the semantic blobs corresponding to objects in an image. In fact, there are some retrieval systems, which use automatic segmentation in various retrieval architectures [4, 7]. Unfortunately, non- semantic and non-uniform nature of the segments formed by the current state of the art segmentation algorithms complicates the similarity calculation for retrieval. In our implementation, we used a rather simple but "adequate" approach to localize the query by first partitioning the image into an M×N rectangular grid as demonstrated in Figure 1. The aim is to form sufficiently small and manageable rectangular regions, so that they can be grouped together to coarsely correspond to objects or semantically related regions like sky,

bird and log areas in Figure 1. The rectangular grid may not be visible to the user. The user manually selects areas for the objects of interest by for example drawing a closed region using the mouse. The regions covering the object areas form the image partition, as indicated in Figure 2.

Figure 1. The uniform partitioning of an image.

3. QUERY PREPARATION

In this phase user guides all the retrieval process according to their intentions and perception of the query image. Query preparation implicitly determines the feature vectors to be used in the similar measurement phase of each rectangular region. This can be considered as formation of a map showing feature assignments to regions. More than one feature can be relevant for a certain region so more than one feature can be assigned for it, or that region may be out of interest so none of the features may be assigned. As defined in the previous section, each image is partitioned into rectangular regions r_{jk} (j = 1, …, N and k = 1, …, M). At this point, we define a set of attributes A = { ajk }. These attributes are pre-defined based on the image content of the database and can be as simple as the words in a dictionary prepared for the database, such as sky, bird, background, etc., or can be as complicated as denoting the appearance of the object, such as mixed object, smooth object, region with no importance, etc. Each attribute ajk, corresponds to a feature vector fjk.

The attributes can be OR-ed or AND-ed for an object region to take the union and intersection of the feature space. The user assigns a set of attribute from a menu to the selected area. The rectangular regions under this area are assigned to the corresponding feature space. The user is allowed to select and assign as many objects according to the needs, until the whole image is covered. This can be considered as formation of a map showing the attributes of regions. If a region is out of interest, none of the features are assigned to this region and it is considered as irrelevant by the system. Actually, the system indirectly makes the user conduct a coarse semantic segmentation of the image area in terms of uniform rectangular regions. A set of low level color (**C**) and texture (**T**) descriptors from MPEG-7 standard set are used to form feature space for each attribute. Figure 2 shows, a sample query image with mutually exclusive and collectively exhaustive attributes. In this example, sky regions are assigned to a color descriptor (**C**), eagle regions are assigned to both color and texture (**C-T**), log regions are assigned to texture (**T**) and irrelevant regions are not assigned to any feature. Overlaps may also be allowed during the object area selection.

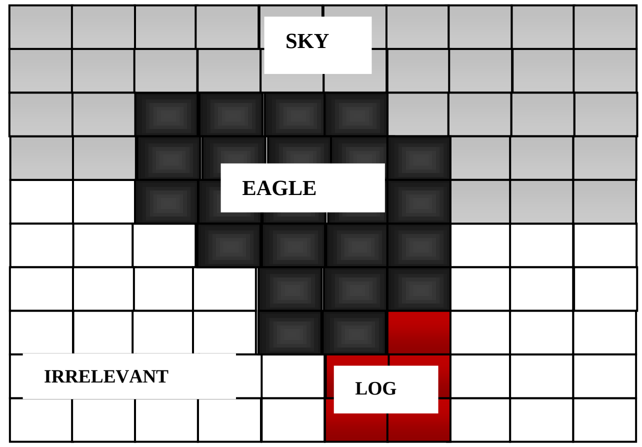

Figure 2. Sample attribute assignments map of the image

4. SIMILARITY MEASUREMENT

After formation of the attribute and the assignment map, the similarity of an image pair is expressed as the similarity of object regions determined by the user in the interface. In our current interface, the user is allowed to choose only one object of interest by assigning attributes for the sake of simplicity. The non-assigned regions are considered to be irrelevant and all the assigned regions are assumed to correspond to the object of interest. The description for the smallest box containing this object is found using the assignment map, in the form {w, h, rqjk, nb}; where w is width, h is height of the box in number of regions, rq_{jk} is the first region of the box in the query image and n_b is the total number of regions (i.e. w×h). This box is slid over the database images and a similarity measure is calculated as the average of similarities of the individual regions contained in it with the corresponding regions of database images.

This procedure is most similar to Local Similarity Patterns method presented in [5] and it can be considered as an improved version allowing object level, location invariant querying. Figure 3 summarizes the pseudo code of our algorithm. Figure 4 summarizes the variables used in the formalization. Given an image pair (I_1, I_2) as query and database images respectively,

extract {w, h, rqjk, nb} from I1's assignment map
initialize distance = MAX_FLOAT
start from the top left corner of I_2
for each possible box in I_2
    start from rqjk
    for each region in the object box
        sum = sum + $D_R(rq_{jk}, rd_{xy}, F_R(rq_{jk}))$
    sum = sum/n_b
    if sum < distance then distance = sum
return distance

Figure 3. Psedocode of the algorithm.

D_R is the region similarity function which takes two regions and a feature vector and return their similarity accordingly. F_R is the feature assignment function, which returns the vector assigned given a region of query image. rq_{jk} denote any region of object box and rd_{xy} denotes any region of sliding box.

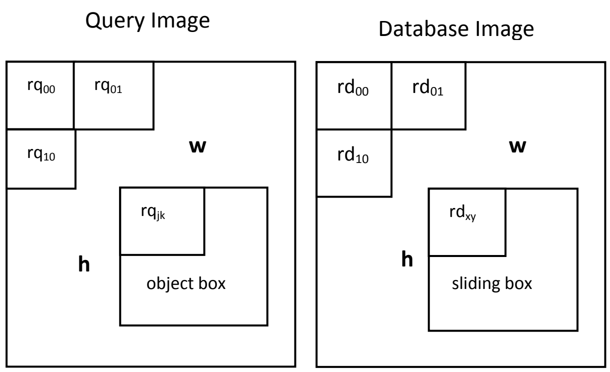

Figure 4. The variable used in the algorithm.

5. FEATURE EXTRACTION

MPEG-7 Experimentation Model [9] provides a descriptor named Multi Image, which enables two of the feature descriptors to be chosen from the MPEG-7 core set and used in combination for retrieval. Multi Image is implemented with Color Structure and Edge Histogram as color and texture descriptors. Then, a higher level descriptor, namely Local Similarity, is implemented using Multi Image descriptor to work in the background of our interface. The Local Similarity descriptor works region-wise, extracting features and calculating similarities as described in the previous section. The features of all the images in the database are extracted and indexed a priori to speed up the query execution.

6. EXPERIMENTS

1000 images having 10 major categories of 100 images each are chosen from Corel Image Gallery, and they are used as the test database to evaluate the performance of the system. The categories contain various outdoor images of landscapes, cityscapes and different kinds of animals. 5 categories (cheetahs, elephants, tigers, eagles and planes) of images with distinct objects are chosen to be the base classes for the experiments, since we are interested in the performance when queries are localized according to objects of interest. For each image in a certain class, appropriate queries are prepared and run with our system and the average of the precision at different match thresholds are taken to be the class precision for that threshold. By match threshold we mean the number of items retrieved as the result of the query, and precision is taken to be the number of relevant items retrieved within the given threshold.

The experiment is repeated for the MPEG-7 Multi Image descriptor when applied to the entire image area in both feature extraction and similarity calculation phases on its own. The results obtained when the same feature vector is applied to the whole image are compared with our system's localized query results. The results are presented in two graphs; the first one shows the average precision for each match threshold and the second one shows the standard deviations. The figures 5 show the graphs of Cheetah class.

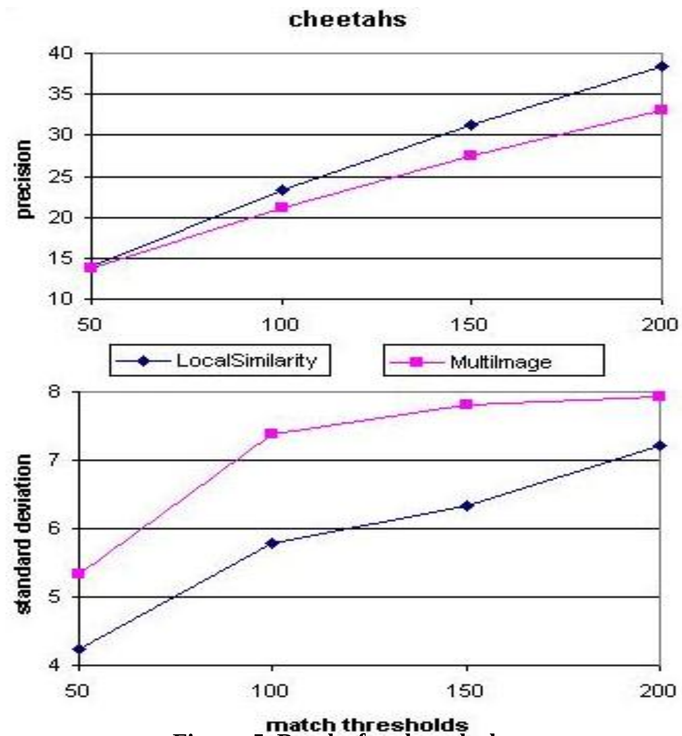

Figure 5. Results for cheetah class

For all the classes except the tiger class, our system indicates better average precision values with smaller standard deviation. The tigers are at various scales and poses in the collection, which is the main drawback of our system due to its region-wise scale variant interpretation of the images. Preparing the query by assigning only color feature to tiger regions gives better results than color-texture assigned queries, this indicates the unsuitable nature of Edge Histogram descriptor for this class and a solution of enlarging feature set suggests itself. If the class contains distinctive objects like elephants usually having a consistent color and pose, our system outperforms the Multi Image descriptor. The power of our system lies behind its ability to retrieve "hard to match" images with higher precision values. By "hard to match", we mean the images having the object on a different background. For example, tigers in the water or cheetahs on the snow. The background changes the appearance of the scene drastically, thus the global method is misled with false matches. Unfortunately, this fact is not observed from the graphs since they give only the number of matches but not the ordering or the quality of the match. Some example query results of this sort are given below with a match threshold of 8 where the first image is the query image, and the left column shows the results of our system. Fig 6 shows the Cheetah query where left column shows the results of Local Similarity descriptor. Fig 7 shows the Tiger query where left column shows the results of Local Similarity descriptor.

Figure 6. A cheetah query

Figure 7. A tiger query

7. CONTENT BASED IMAGE DATABASE MODELLING

Image databases designers use three main inputs: an image collection, user's requirements, and an application domain. Each input induces constraints on the databases to be built. Modelers must find a compromise between these constraints [5]. In order to determine the actual needs of IDB designers, we have studied various projects [2, 3,7]. Our review` of IDB can be summarized as follows:

• Volume of data IDB has to manage huge amounts of data. They use two main strategies (indexing and classification) in order to virtually diminish the amount of data to be researched during the image retrieval.

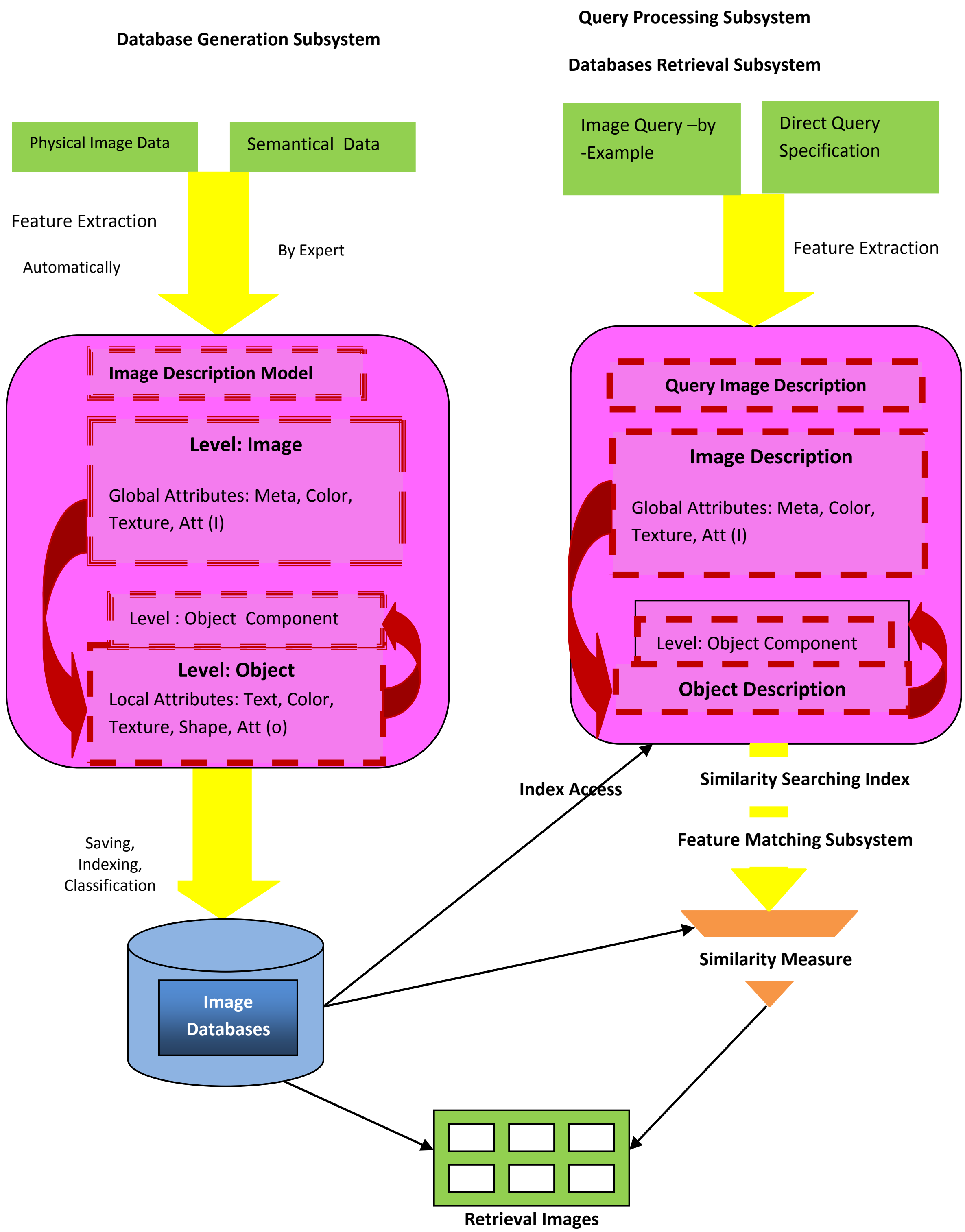

Figure.8 Over All Structures of Content Based Image Retrieval Model

• An image is described by a combination of syntactical (color histograms, textures, shapes) and semantically (also called meta-data) information and a query can be expressed using a combination of syntactical and semantically image features.

• Granularity image descriptions are generally composed of global and local information.

Our conviction is that an efficient framework should:

• be generic enough in order to cover most of modelers requirements, whatever their image corpus and their users' requirements may be.

• provide a convenient support for both syntactical and semantical information

• provide a convenient support for local and global descriptions.

Fig.8 shows the overall Structures of Content Based Image retrieval Model. The creation of our framework of CBIR system steps on the base of the published in [2, 3, 4, 5,10] models but it differs from them by structure, components, and data models. It includes: built-by-modules subsystems, data models, processing and algorithms types. This framework offers a choice of all possible components in order to build up an image database from a corpus of images and users' requirements. As a result of the analysis of the existing IDB, of their structures, organization and abilities for access that they give, a generalized architecture of the possible image data processing in the basic processes of saving and retrieval of IDB images is presented. This generalized our architecture-based framework for building an Image Databases from a Collection of Images shown in Fig.1 and may be used for presentation of the key editions and methods that are proposed in the literature. In our method's architecture, the basic components are the two mutually connected processes: image saving and inserting in IDB and image retrieval by user's query. The two basic subsystems "DB Generation Subsystem" and "DB Retrieval Subsystem" correspond to these main processes. In the first subsystem, the images that are added to IDB are processed by features extraction algorithms or by expert. In this way the images are described by syntactical and semantical attributes, which are presented in our data model. The data model includes connected sub-descriptions on different levels.

Each one of them includes attributes, a set of components and spatial relations between components. The types of the extracted image features, their describing characteristics and the levels of their extraction (image, object, components) are determining for the abilities of access to the saved data. Within our model an image, denoted by I is described in terms of simple and of complex objects. Let us denote OI a set of simple and complex objects. Each object $o \in O_I$ is described by Attributes (semantic and syntactical (color, texture, shape) de-noted by Att(o)), Set of object components Set(o), and Spatial Relation between Objects Rel(o)). We denoted by $Descr_I$ the set of image description:

$$Descr_I = \{< Att(o), Set(o), Rel(o) >\}_{o} \in OI$$

The images inserted in IDB together with their feature description data form a property presentation that is saved in a data structure of IDB. The features in the form of coded vectors can be used as indexes for direct organization of the access to IDB or for data clustering. According to the level of the used presentation of image contents two approaches for indexing techniques may be determined: indexing on the base of global distribution of the image characteristics and indexing on the base of the typical peculiarities of local image areas or regions.

The "DB Retrieval Subsystem" has to search and discover an answer of user's query. This subsystem is divided in two subsystems: "Query Processing Subsystem" and "Feature Matching Subsystem". This process is determining the system rapidity. The first subsystem processes the query primarily. The query specification may be done by example image, drawn by the user's draft, or exact and clear information from the user about the primary features of his interest. The cognitively based presentation has an important role in query processing on different levels. The query presentation is a result of the same processing for the same properties extraction so as inserting image in IDB. The same algorithms for properties extraction are used also for the query image named "Example" and the result is a presentation that is used for the query index forming. The search of a similar index to those in the IDB is implemented by the "Feature Matching Subsystem". The similarity matching of the sample index with the IDB indexes aims parts of the images to be found that are similar to a given sample or to a defined variant of a given sample.The type and depth of the properties extraction from the inserted in IDB images are determining for the functionality and flexibility of every visual information system. As in most cases the extracted characteristics-indexes are multidimensional; the approach of similarity search is perceived. The similarity search uses a similarity measure as a similarity criterion. The measure evaluates the similarity degree between two images and is determining in the process indexing and similarity integration of multi-dimensional index vectors of the image and the query.

8. CONCLUSION AND FUTURE WORK

The results indicate that localizing queries according to their object contents and using different feature vectors according to the object's visual characteristics is a promising technique. Even though, the proposed method has some drawbacks like lacking scale invariance, there are various possible improvements that can increase the power of the system. First of all, we can make the system scale invariant, using a multi-scale partitioning, together with the scale invariant features. Over segmenting the images using an automatic algorithm and letting the user group the smaller regions into objects, may also be an alternative partitioning method. In this way, we may obtain a more accurate partitioning compared to our manual coarse segmentation over a rectangular grid. Similarity measurements for the sliding areas can be formulated as a matching problem, by employing optimal correspondence matching, elastic matching, or relaxation matching. Some special cases like occlusion can be handled, during the attribute assignment phase. The current feature set can be extended to cover a broader class of descriptors, including higher level of texture and shape descriptors. Lastly, the attribute mapping of our sliding areas can be enhanced by allowing the user formulate queries in the form "look for this object on this kind of background" to benefit from the background's distinctiveness.

9. REFERENCES

[1] D. A. Forsyth, J. Ponce, Computer Vision: A modern Approach, Chapter 25, Prentice-Hall, 2001.

[2] Y. Rui, T.S. Huang, "Image Retrieval: Current Techniques, Promising Directions and Open Issues", Journal of Visual Communication and Image Representation, vol. 10, pp. 1–23, 1999.

[3] A. A. Goodrum, "Image Information Retrieval: An Overview of Current Research", Special Issue on Information Science Research, vol. 3, no. 2, 2000.

[4] C. Carson, S. Belongie, H. Greenspan, J. Malik, "Blobworld: Image Segmentation Using Expectation-Maximization and Its Application to Image Querying", IEEE Trans PAMI, vol. 24, no. 8, 2002.

[5] Z. Stejic, Y. Takama, K. Hirota, "Genetic Algorithm-based Relevance Feedback for Image Retrieval Using Local Similarity Patterns", Information Processing & Management, 2002.

[6] X. S. Zhou, T. S. Huang, "Relevance Feedback in Image Retrieval: A Comprehensive Review", CVPR 2001 CBAIVL Workshop.

[7] J. Z. Wang, J. Li, G. Wiederhold, "Simplicity: Semantics-sensitive Integrated Matching for Picture Libraries," IEEE Trans. PAMI, vol 23, no.9, 2001.

[8]MPEG-7 Home page, http://www.darmstadt.gmd.de/mobile/MPEG7/index.html.

[9] MPEG 7 Experimentation Model Software, www.lis.ei.tum.de/research/bv/topics/mmdb/e_mpeg7.html

[10] Marinovic,I.; FurstnerI.; Manufaktura d.o.o., Subotica , "Content-based image retrieval" Intelligent Systems and Informatics, 2008. SISY 2008. 6th International Symposium on Sept. 2008.pp 1-6