

Interactive Image Retrieval Using Text and Image Content

*B. Dinakaran*¹, *J. Annapurna*², *Ch. Aswani Kumar*¹

¹ *School of Information Technology and Engineering (Presently with Cognizant Technology Solutions, India)*

² *School of Computing Sciences and Engineering, VIT University, Vellore, India.*
E-mail: aswanis@gmail.com

Abstract: *The current image retrieval systems are successful in retrieving images, using keyword based approaches. However, they are incapable to retrieve the images which are context sensitive and annotated inappropriately. Content-Based Image Retrieval (CBIR) aims at developing techniques that support effective searching and browsing of large image repositories, based on automatically derived image features. The current CBIR systems suffer from the semantic gap. Though a user feedback is suggested as a remedy to this problem, it often leads to distraction in the search. To overcome these disadvantages, we propose a novel interactive image retrieval system, integrating text and image content to enhance the retrieval accuracy. Also we propose a novel refining search algorithm to narrow down the search further from the retrieved images. The experimental results demonstrate the performance of the proposed system.*

Keywords: *Color histogram, color quantization, image descriptor, refining search, region-based segmentation and term feedback.*

I. Introduction

The interest towards image retrieval is increased due to the rapid growth of the World Wide Web. The need to find a desired image from a collection is shared by many groups, including journalists, engineers, historians, designers, teachers, artists, and advertising agencies. The image needs and usages vary considerably among the users in these groups [1]. The users may require access to the images, based on primitive features, such as color, texture or shape, or associated text. The technology to access these images has also accelerated phenomenally. The current

approaches are broad and inter-disciplinary, mainly focused on three aspects of image research which are text-based retrieval, content-based retrieval and interactive based image retrieval. Early techniques are based on the textual annotation of images.

Many techniques have been developed for text-based information retrieval [2] and they proved to be highly successful for indexing and querying web sites. Their success may also shed some light on the area of image retrieval, because the relatively mature theories and techniques of text-based information retrieval may be applicable to the image domain. Text-based image retrieval uses traditional database techniques to manage images. Through text descriptions, images can be organized by topical or semantic hierarchies to facilitate easy navigation and browsing based on standard Boolean queries. Although text-based methods are fast and reliable when images are well annotated, they are incapable of searching in unannotated image collections. The generalization of the information retrieval from the text domain to the image database is, however, non-trivial. The greatest obstacle arises from the intrinsic difference between the text and image in representing and expressing information [3].

Recently Content-Based Image Retrieval (CBIR) has become an active research area [4]. CBIR refers to techniques used to index and retrieve images from databases based on their visual content. Visual content is typically defined by a set of low level features extracted from an image that describe the color, texture and/or shape of the entire image [5]. CBIR methods are capable of searching in large database collections. However, the present methods used in CBIR are neither fully reliable, nor fast enough to handle image databases beyond a closed domain and the techniques used are still subject to development. Various image retrieval systems [6], including Query By Image Content (QBIC), VisualSeek and Virage have been built, based on the low-level features for general or specific image retrieval tasks. However, effective and precise image retrieval still remains an open problem because of the extreme difficulty in fully characterizing images. Successful techniques have been developed for some specific applications, such as face and finger-print recognition [7, 8]. An effective approach for querying and browsing images still remains elusive.

Both the text and content based techniques have their own characteristics, advantages and disadvantages. By combining them, parts of their disadvantages can be overcome. The existing image retrieval systems are with either text-based or image-based queries, but not both. Hence, a system with integrated methods is highly needed [9, 10]. Fig. 1 presents a constructed example of the additional semantics gained by combining text and sample image query. Fig. 1a) shows the first three retrieved images for user query "Bike". For the given query the third image is irrelevant because of poor annotation. Fig. 1b) shows the retrieval result for the query image. Fig. 1c) presents the constructed example of the additional semantics gained by combining text and image query. Thus significant improvements in the retrieval of image can be gained by embedding the image with text [11, 12]. In this paper our objective is to design an effective and efficient

hybrid image retrieval system having the capabilities of searching image database with both text and/or image based query.

Combining text and image based query for retrieving the images from the database may lead to huge selection for the user. Thus to pick a desired image more appropriately from the retrieved set, an interactive approach is needed. In interactive image retrieval, the user has to interact with the system to locate an interested image, which is a time consuming task and it also needs more user involvement. Query refinement is an interactive process to access the desired image quickly. Query refinement comprises query expansion and query re-weighting. Query expansion allows the user to expand the query for desired image search. To expand the query, the user has to find the other relevant terms which are again time consuming and cause distraction in the search. We present a novel refining search algorithm to optimize the user searching time on a retrieved image and improve the system quality. The proposal is that, based on the feedback from the user about the retrieved objects, the system will filter the retrieval results further to narrow the user's information needs.












	Query	Result		
a	“Bike”	 Bike field woman	 Bike mountain outside	 Bike motor red
b		 Bike field woman	 Wheel grass	 Ferry wheel forest
c	“Bike” 	 Bike field woman	 Bike black	 Bike garage outside

Fig. 1. a) Text user query b) Sample image c) Both text and image

Our paper is organized as follows. Section II reviews current image retrieval approaches. In Section III we present a novel method for CBIR and propose a new refining search algorithm. Section IV discusses the implementation issues and result analysis, followed by conclusions, acknowledgment and references.

II. Current image retrieval approaches

Current image retrieval techniques can be classified into four categories: text-based, content-based, composite and interactive approaches. Abby described an overview about the current image retrieval techniques and issues [4]. The text-based approach is a traditional simple keyword based search. The images are indexed according to

the content, like the caption of the image; filename, title of the web page, and alternate tag, etc. and stored in the database. Processing a user query could involve a stop word removal, stemming and tokenizing. Some of the keyword based image retrieval approaches are bag of words, natural language processing and Boolean model [11]. Image retrieval is then shifted to standard database management capability combined with information retrieval techniques. Some commercial image search engines, such as Google Image Search and Lycos Multimedia Search, are keyword-based image retrieval systems.

In content-based approach, the processing of a query image involves extraction of visual features and perform search in the database for similar images [15]. A typical CBIR system views the query image and images in the database (target images) as a collection of features, and ranks the relevance between the query image and any target image in proportion to a similarity measure, calculated from the features. The low level image features can be used to compute similarity between images [16]. Valova and Rachev [3] have described the method of image retrieval using color features. Retrieving images, using a region, was done by Mezaris and Doulaverakis [5]. A review on technology and issues related to CBIR was made by Kerfi et al. [14]. Despite the recent progress, content-based image retrieval has its own limitations because of the semantic gap between the low level image features and high level semantic content of images (like sunset, flowers, etc.). Many approaches have been proposed to reduce the semantic gap. They generally fall into two classes, depending on the degree of user involvement in the retrieval: relevance feedback and image database preprocessing using statistical classification.

Relevance feedback is a powerful technique, originally used in the traditional text-based information retrieval systems. In CBIR a relevance-feedback-based approach allows the user to interact with the retrieval algorithm by providing information regarding the images which the user believes to be relevant to the query. Most of the current term feedback techniques are described by Chai et al. [13]. For interactive CBIR the studies suggested various query expansion or modification mechanisms. Kriengkari et al. [18] have discussed query refinement for content-based image retrieval. Based on the user feedback, the model of similarity measure is dynamically updated to give a better approximation of the perception subjectivity. Empirical results demonstrate the effectiveness of a relevance feedback for certain applications. Nonetheless, such a system may add burden to the user especially when more information is required than just a Boolean feedback (relevant or non-relevant)

III. The architecture proposed

We propose a novel image retrieval system, which uses the text and/or visual contents of the images in the database. Fig. 2 shows the architecture of our proposed system. The system proposed has user interface, where the user enters the text and/ or sample image as a query. The textual and visual content descriptors are

generated from the text query and image query. The descriptors are converted into a vector format.

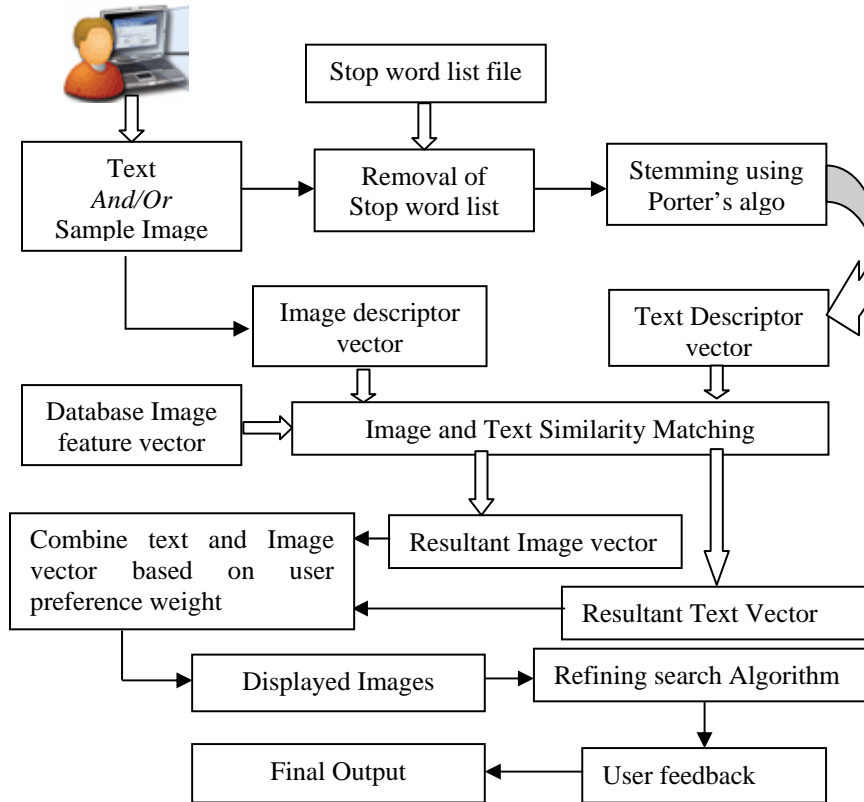


Fig. 2. Architecture proposed

Similarly textual and visual descriptors are calculated and converted into vector representation for the images stored in the database. The vector, generated by the user query is then matched with the vectors stored in the database. The text and content-based methods return two independent lists of images with different weights. These two lists must be combined in a meaningful way to give the user a combined image list [17]. The user preference weight for text and color is incorporated. The default weight is 50% each for both methods. The relevant images are retrieved and displayed according to their relevance. The relevance feedback helps to search the desired image quickly.

a) Text description vector

Text-based image retrieval can be based on the traditional information retrieval technique. However, to improve retrieval performance, we should make use of the structure of HTML documents. This is because words or terms appearing at different locations of an HTML document have different levels of importance or

relevance to related images [19]. We classify the terms into groups, based on their locations, reflecting their importance for image retrieval. Table 1 shows the image description source and its corresponding weight, assigned for our system. In text-based, the words are subjective and it is difficult to choose the ultimate word, as describing visual impressions remains a tedious task. Embedding the sample image with text optimizes the image retrieval.

Table 1. Image description source and its weights

Description source	Weights
Html page title	10
Filename	10
Same Paragraph	5
Image path	10
Alternate tag	10
Caption	5

b) Color and region based image retrieval

Among all the visual features, color is one of the most dominant and important features for image representation [20]. In the system proposed we implemented the color and region based retrieval technique, as shown in Fig. 3. The color based technique has been reported to produce good retrieval performance for WWW images [19]. Each color channel is quantized into 8 intervals and produce color histograms for the user query image. Then a key identifier is generated by sorting the color histograms vector to optimize the search in the database [24, 26]. The image key identifier and color histograms are the feature vectors which are to be stored as the index of the image database [21]. We used histogram Euclidean distance to measure the distance between the color histograms of the query image and the images in the database [22].

The color based techniques are not suitable for the same image on different resolution because of a difference in the number of pixels. To overcome this we can combine the region with a color feature. From the user queried sample image we calculate the number of regions and the pixel count in each region (the region vector) using a region growing algorithm, described by Li [23] and the similarity measure between the image region vectors, using Jaccard's coefficient [22],

$$(1) \quad \text{Similarity}(I, Q) = \frac{|I \cap Q|}{|I \cup Q|},$$

I = Colour descriptors, representing the 2D index vector of the images in the database;

Q = Colour descriptors representing the 2D query vector.

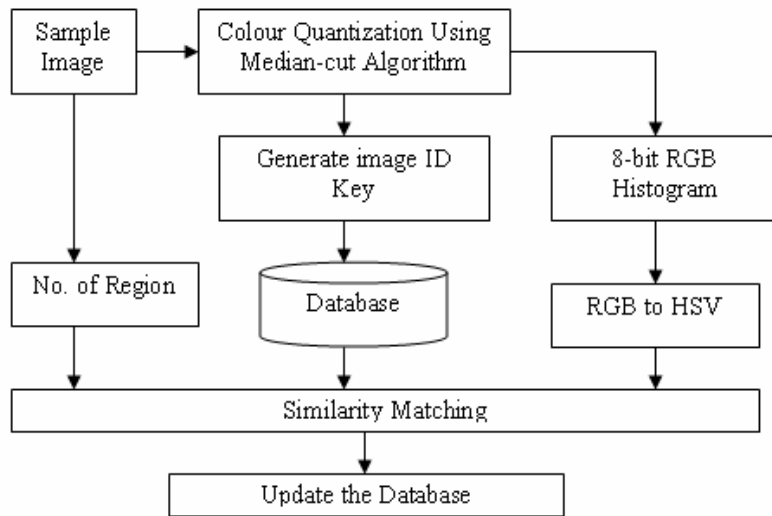


Fig. 3. Image descriptor vector

Chai et al. [13] proposed an interactive image retrieval techniques using user term feedback for a text-based approach. Their approach is to collect terms from all the fields which leads to more confusion and causes more possibilities to produce unrelated terms. Instead of collecting the terms from the image database, we collected only from retrieved image sets for the given user query. The filename, the alternate tag and caption fields other than the user input query terms, have more probability for getting relevant terms to narrow the search. Fig. 4 shows the refining search algorithm suggested.

Let A be a retrieved image set for the user query.

Step 1. Collect terms from fields, such as filename, alternate and caption tag from A .

Step 2. Filter the repeated words and user query terms and calculate each term occurrence.

Step 3. Sort the terms in a descending order based on term occurrence.

Step 4. Allow the user to select one or more terms, which are relevant for his/her interest image.

Step 5. Perform the simple keyword search technique (natural language processing) for the user selected terms on A and display the resultant images.

Fig. 4. Propose refine search algorithm for refining the search

IV. Implementation and result analysis

The system proposed accepts the input from the user query text and/or sample image. Later the system performs the pre-processing steps, like tokenizing, stop word removal and stemming on the input query. Thus we obtain a processed query, which is used to search the database and further the results are ranked and displayed. Then the system calculates the user queried sample image feature vectors by performing the preprocessing steps, like color quantization, finding the number

of regions, using a stack based region growing algorithm and generating an 8-bin HSV histogram. Fig. 5 shows the stack based region growing algorithm.

Step 1. Initialize a two dimension array of the image size.
Step 2. Find a pixel which is not labeled. Label it and store its coordinates on a stack.
 Do:
 2.1) get a pixel from the stack;
 2.2) check its neighbours to see, if they are unlabeled and close to the considered pixel; if so, label them and store them on the stack.
Step 3. Repeat the above step until there are no more pixels on the image.

Fig. 5. Stack based region growing algorithm

The similarity between the sample image feature vectors with database image vectors is calculated, using a histogram Euclidean distance, and a final image vector is generated [2, 11, 14, 25] . The default 50% weight is provided for both text and image vector to generate a final vector. The major difficulty in combining the two weight lists is that the two lists are obtained, using two totally different weighting schemes (one based on term weights and the other on color and number of regions). So we cannot simply add the weights of each image in the two lists to obtain a combined weight. The solution we used is to normalize the similarities calculated based on the text, the color histograms and the region, so that the normalized similarities are within the common range of 0 and 1. Fig. 6 shows the retrieval results obtained for the user query with both a keyword and a sample image. It retrieves more relevant images for the combined user query than separately.

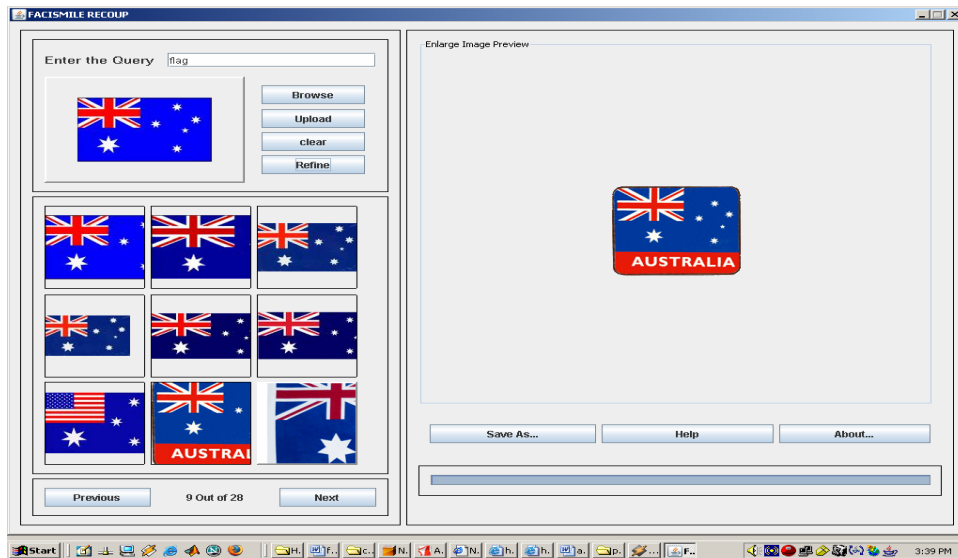


Fig. 6. Both combined text and sample image query

Table 2 shows the performance of our system. We measured the performance of the system, using two retrieval statistics precision and recall.

Table 2. Performance analysis of the proposed system

User query	Total No of Relevant images	No of Retrieved images	Relevant image retrieved	Precision	Recall
Text	30	24	20	0.83	0.66
Image	30	20	18	0.9	0.60
Combined	30	28	25	0.89	0.83

The definition of the recall and precision are given in (2) and (3). Our experimental results show that the system proposed is able to retrieve more relevant images than the text-based and the content-based techniques. The advantage of using both text and color return results is that they are more relevant when compared to searches based only on a text or a sample image:

$$(2) \quad \text{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Total relevant images in collection}},$$

$$(3) \quad \text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}}.$$

The second benefit is that a larger set of results is produced in the system. The images that match either one or both of the specified search criteria for terms and color matching are retrieved. The third one is the ease of carrying out the image search. The interactive search is an iterative process, whereby the user may refine a search on the retrieved images, as shown in Fig. 7.

These benefits allow the user to narrow down the range of the results found, so that a more precise range of images is found. Fig. 7 presents the term feedback for the user query “flag”, such as stickers, robe and ring, etc. In refined search, the lists of other related terms for the given query are shown to the user. The users can choose terms from this list to further refine their queries. The list of terms is automatically generated by the system from the retrieved images. Table 3 shows the result obtained for the given query, before and after the refined search. Thus our system reduces the user's search time for his/her image of interest.

Table 3. Performance analysis of the resultant image obtained before and after refined search

Search time	Number of retrieved images
Before refined search	28
After refined search	8

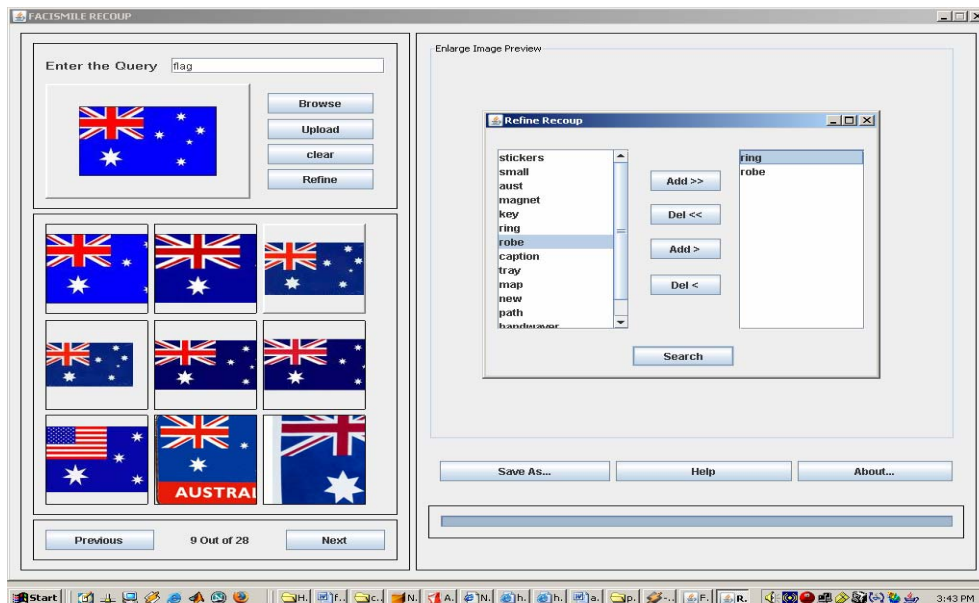


Fig. 7. Refining the search results

V. Conclusion

In this paper we propose a novel image retrieval approach which combines text, content and interactive based retrieval. The accuracy is higher in comparison to using the techniques separately. We designed a hybrid image retrieval system with the method proposed, which successfully achieves the demands with respect to the system requirements (i.e., allows the users to retrieve their desired images based on the text and/or sample image query). A new refining search algorithm has been provided, which optimizes the search results. The experiments on the sample data sets prove the effectiveness of the system. Further it is proposed to integrate and classify the techniques and add low level features, like texture, shape to the system.

Acknowledgements: The authors would like to thank the reviewers for their useful comments. One of the authors, Ch. Aswani Kumar, gratefully acknowledges the financial support from Dept. of Science and Technology, Govt. of India under Grant No SR/S3/EECE/25/2005.

References

1. Miller, H., A. Geissbuhler, S. Marchand, P. Clough. Benchmarking Image Retrieval Applications. – In: Proc. of the 10th International Conference on Distributed Multimedia Systems, San Francisco, CA, USA, 2004.
2. Vikhar, P. A. Content Based Image Retrieval (CBIR): State-of-the-Art and Future Scope for Research. – The IUP Journal of Information Technology, Vol. 6, 2010, No 2, 64-84.
3. Valova, I., B. Rachev. Retrieval by Color Features in Image Databases. – In: Proc. of Adbis'04, University of Rouse, Budapest, Hungary, 2004.
4. Abby, A., Goodrum. Image Information Retrieval: An Overview of Current Research. – Informing Science, Vol. 3, 2000, No 2, 63-66.

5. Mezaris, V., H. Doulaverakis. A Test-Bed for Region Based Image Retrieval Using Multiple Segmentation Algorithms and the MPEG-7experiment Model: The Schema Reference System. Network of Excellence in Content-Based Semantic Scene Analysis and Information Retrieval, IST-2001-32795, 2001.
6. Borlund, P., P. Ingwersen. The Development of a Method for the Evaluation of Interactive Information Retrieval Systems. – Journal of Documentation, Vol. **53**, 1997, No 3, 225-250.
7. Kuilenburg, H., M. Wiering, M. Uyl. Model Based Methods for Automatic Analysis of Face Images. – In: Proc. of 16th European Conference on Machine Learning, 2005, 194-205.
8. Sandeep, K., A. N. Rajagopalan. Human Face Detection in Cluttered Color Images Using Skin Color and Edge Information. – In: Proc. of Indian Conference on Computer Vision, Graphics and Image Processing, 2002.
9. Alexandre, L. A., M. Pereira, S. C. Madeira, J. Cordeiro, G. Dias. Web Image Indexing: Combining Image Analysis with Text Processing. – In: Proc. of 5th International Workshop on Image Analysis for Multimedia Interactive Services – WIAMIS 2004. Lisboa, Portugal, 2004.
10. Westerveld, T., J. C. Gemert, R. Cornacchia, D. Hiemstra, A. P. Vries. An Integrated Approach to Text and Image Retrieval. – In: Proc. of the TRECVID Workshop, November 2005.
11. Gemert, J. C. Retrieving Images as Text. Master's Thesis, University van Amsterdam, 2003.
12. Deselaers, T., T. Weyand, D. Keyers. FIRE in ImageCLEF 2005: Combining Content-Based Image Retrieval with Textual Information Retrieval. Working Notes of CLEF Workshop, Austria, 2005.
13. Chai, J. Y., C. Zhang, R. Jin. An Empirical Investigation of User Term Feedback in Textbased Targeted Image Search. – ACM Trans. Inform. Systems, Vol. **25**, 2007, No 1.
14. Kherfi, M. L., D. Ziou, A. Bernadi. Image Retrieval from the World Wide Web: Issues, Techniques and Systems. – ACM Computing Surveys, Vol. **36**, 2004, No 1, 35-67.
15. Burns, M., K. Leung, A. Rowland, J. Vickers, J. V. Hajnal, D. Rueckert, D. L. G. Hill. Information Extraction from Images (IXI): Grid Services for Medical Imaging. – In: Proc. of Distributed Databases and Processing in Medical Image Computing (DiDaMIC), Rennes, France, 2004, 65-73.
16. Anne, H., H. Ngu, Quan Z. Sheng, Du Q. Huynh, Ron Lei. Combining Multi-Visual Features for Efficient Indexing in a Large Image Database. – VLDB Journal, 2001.
17. Lin, C. H., R. T. Chen, Y. K. Chan. A Smart Content Based Image Retrieval System Based on Color and Texture Feature. – Image and Vision Computing, Vol. **27**, 2009, No 6, 658-665.
18. Kriengkrai, P., M. Sharad, M. Ortega. Query Refinement for Content Based Multimedia Retrieval in MARS. – In: Proc. of 7th ACM Int. Conference on Multimedia, 1999, 235-238.
19. Lu, G. B. Williams. Integrated www Image Retrieval System. – In: Proc. of the 5th Australian WWW Conference, Southern Cross University, 2000.
20. Ko, B. C., H. Lee, H. Byun. Image Retrieval Using Flexible Image Sub Blocks. – In: Proc. of the ACM Symposium on Applied Computing, 2000, 574-578.
21. Ting-sheng, L. CHROMA: A Photographic Image Retrieval System. PhD Thesis, University of Sunderland, United Kingdom, January 2000.
22. Sangoh, J., Histogram-Based Color Image Retrieval. Psych221/EE362 Project Report, Stanford University, 2002.
23. Li, Y., D. Lu, X. Lu, J. Liu. Interactive Color Image Segmentation by Region Growing Combined with Image Enhancement Based on Bezier Model. – In: Proc. of the 3rd International Conference on Image and Graphics, 2004, 96-99.
24. El Alami, M. E. A Novel Image Retrieval Model Based on Most Relevant Features. – Knowledge Based Systems, Vol. **24**, 2011, No 1, 23-32.
25. Muller, H., W. Muller, D. Mc G. Squire, S. Marchand-Maillet, T. Pun. Performance Evaluation in Content-Based Image Retrieval: Overview and Proposals. – Pattern Recognition Letters, Vol. **22**, 2001, No 5, 593-601.
26. Lin, C. H., R. T. Chen, Y. K. Chan. A Smart Content Based Image Retrieval System Based on Color and Texture Feature. – Image and Vision Computing, Vol. **27**, 2009, No 6, 658-665.