[Stammberger 1999] Stammberger T., Eckstein F., Michaelis M., Englmeier K-H., Reiser M. (1999). *Interobserver reproducibility of quantitative cartilage measurements: comparison of B-spline snake and manual segmentation.* In: Magnetic Resonance Imaging, 17(7). p.1033-1042.

[Tsechpenakis 2004] Tsechpenakis G., Rapantzikos K., Tsapatsoulis N., Kollias S. (2004). *A snake model for object tracking in natural sequences.* In: Signal processing: Image communication 19, p.219-238.

[Westermann 2003] Westermann U., Klas W. (2003). *An analysis of XML database solutions for the management of MPEG-7 media descriptions*. ACM Computing Surveys, Volume 35 (4), p.331-373.

[Yuille 1992] Yuille A., Hallinan P., Cohen D. (1992). *Feature extraction from faces using deformable templates.* In: International Journal of Computer Vision, 8(2), p.99-111.

Authors' Information

Jérôme Godard – National Institute of Informatics (NII), Graduate University of Advanced Studies (Sokendai)
Hitotsubashi 2-1-2 Chiyoda-ku, Tokyo 101-8430, Japan; e-mail: jerome@grad.nii.ac.jp

Frédéric Andrès – National Institute of Informatics (NII), Graduate University of Advanced Studies (Sokendai)
Hitotsubashi 2-1-2 Chiyoda-ku, Tokyo 101-8430, Japan; e-mail: andres@nii.ac.jp

Kinji Ono – National Institute of Informatics (NII), Waseda University
Hitotsubashi 2-1-2 Chiyoda-ku, Tokyo 101-8430, Japan; e-mail: ono@nii.ac.jp

MPEG-7 BASED IMAGE RETRIEVAL ON THE WORLD WIDE WEB

Rajeev Agrawal, Farshad Fotouhi, Peter Stanchev, and Ming Dong

Abstract: Due to the rapid growth of the number of digital media elements like image, video, audio, graphics on Internet, there is an increasing demand for effective search and retrieval techniques. Recently, many search engines have made image search as an option like Google, AlltheWeb, AltaVista, Freenet. In addition to this, Ditto, Picsearch, can search only the images on Internet. There are also other domain specific search engines available for graphics and clip art, audio, video, educational images, artwork, stock photos, science and nature [www.faganfinder.com/img]. These entire search engines are directory based. They crawls the entire Internet and index all the images in certain categories. They do not display the images in any particular order with respect to the time and context. With the availability of MPEG-7, a standard for describing multimedia content, it is now possible to store the images with its metadata in a structured format. This helps in searching and retrieving the images. The MPEG-7 standard uses XML to describe the content of multimedia information objects. These objects will have metadata information in the form of MPEG-7 or any other similar format associated with them. It can be used in different ways to search the objects. In this paper we propose a system, which can do content based image retrieval on the World Wide Web. It displays the result in user-defined order.

Keywords: XML, MPEG-7, Metadata, Multimedia, Content Based Image Retrieval (CBIR)

1. Introduction

The CBIR has been a very active research area in the last decade. Conventional content-based image retrieval systems [1, 2, 3] use low-level features such as color, texture, shape, automatically extracted from the images. Another focus of this research is on improving the low level features. The modifying the similarity measures make the retrieval as better as possible. It is argued in [4] that unconstrained object recognition is still beyond of current technology. The content based systems can at best capture only pre-attentive similarity, not semantic similarity. So far there has not been a single system, which can perform this task automatically without human intervention due to the nature of this problem.

The expansion of the World Wide Web (WWW) is making the problem of effective retrieval of images very important for all its users. The complexity of Web documents is rapidly increasing with the wide use of multimedia components, such as images, audio and video, associated to the traditional textual content. This requires extended capabilities of the Web query search engines in order to access images according to their multimedia content. A large number of search engines (e.g. Altavista, Yahoo, HotBot, etc.) support indexing and content-based retrieval of Web documents. Only the textual information is taken into account. Initial experimental systems providing support to the retrieval of Web documents based on their multimedia content (Webseek [5] and Amore [6]) are limited to the use of pure physical features extracted from multimedia data, such as color, shape, texture. These systems do not go beyond the use of pure physical visual properties of the images. They suffer the same severe limitations of today as the general-purpose image retrieval systems [7], such as Virage [8] and QBIC [9]. These systems consider images as independent objects, without any semantic organization in the database or any semantic inter-relationships between database objects. Many image searches also use an approach that filters out less relevant results. They analyze and index the text on the page adjacent to the image, the image link text, text in the HTML alt tag, filename or file path name. Similarly, this approach can also be used with other media files such as audio and video. Even though these search engines do not “look inside” the media files, they can give quite relevant results.

Another approach can be to look into the media file contents itself and trying to mine for textual information in the file for better multimedia indexing. For example, a Portable Network Graphics (PNG) image file can contain textual information such as title, author, description, copyright, creation time, software used, disclaimer, warning, source and comments [10]. Not all file formats contain metadata, and even if they do, an indexing engine should know how to handle all the different file formats and where to find that information in a file. It would be better if we had a data model which could be used with different media formats and utilized a rich set of metadata. There have been many metadata models developed. Some of them are RLG Preservation Metadata Elements, NISO Draft Standard, DIG35 Specification, Data Dictionary for Audio/Video Metadata, Metadata for Long-Term Preservation, Metadata Encoding and Transmission Standard [11]. MPEG-7 is another multimedia metadata standard. The Moving Picture Experts Group (MPEG) was established in 1988 to develop audiovisual compression standards. MPEG-1, MPEG-2, and MPEG-4 all represent the content itself, while MPEG-7 represents information about the content [12]. While the first produces the contents, the latter describes the content. There are number of tools provided in MPEG-7 - descriptors (the elements), description schemes (the structures), a Description Definition Language (DDL) (for extending the predefined set of tools) and a number of system tools. MPEG-7 can support all natural languages. DDL provides the foundation for the standard. It provides the language for defining the structure and content of MPEG-7 documents. The DDL is not a modeling language such as Unified Modeling Language (UML) but a schema language to represent the results of modeling audiovisual data (i.e. descriptors and description schemes) as a set of syntactic, structural and value constraints to which valid MPEG-7 descriptors, description schemes, and descriptions must conform. The purpose of a schema is to define a class of XML documents. The purpose of and MPEG-7 schema is to define a class of MPEG-7 documents. MPEG-7 instances are XML documents that conform to a particular MPEG-7 schema (expressed in the DDL) and that describe audiovisual content. MPEG7 has been developed after many rounds of careful discussion. It is expected that this standard would be used in searching and retrieving for all types of media objects. If we have images stored with MPEG-7 metadata, it would be easier to do semantic retrieval. MPEG-7 files contain a reference to the location of the corresponding image file. It is also possible to exploit other tools and technologies developed for XML like Xquery, XPath, etc. There has been a lot of work on XML schema integration. This plays a central role in numerous applications, such as web-oriented data integration, electronic commerce, schema evolution and migration, application evolution, data warehousing etc. In schema integration, the main objective is to find a suitable technique to match the elements in different schemas. We propose to combine XML schema integration techniques and image retrieval techniques using low-level features with or without semantic annotations.

Rest of the paper is organized as follows: Section 2 describes the motivating examples. Section 3 relates a list of previous work and other literature survey. Section 4 describes our proposed system. Finally, we give concluding remarks in section 5.

2. Motivating Examples

The commercial image search engines available today basically search the images based on keywords. The keywords are extracted from the web page, where image appears. But the keyword based search has its own limitation, which will be clear from the following examples.

1. If we want to search and retrieve the pictures of a person in the different stages of his/her life with respect to the time, available on different websites, that is not possible through keyword search. The keyword search would definitely retrieve the images but not integrate in the order we want. Assumption here is that different websites has the pictures of the person at different stages of his/her life and also incorporate some semantic information, which can be in MPEG-7 or in any other metadata format. The reason is keyword search just looks for the name in the surrounding text, but no in other information. E.g. when we search the pictures of a great person like Mahatma Gandhi images are retrieved, but not in any order. The main reason is that no semantic information is incorporated with the images.

2. Some security agency is interested in getting more information about a person, who has perpetrated some crime and they have a photograph of this person. There is no technique available which can return the information about this person from the Internet, if the agency uses this photograph as input (query by example method). The basic idea of this kind of search is that low level feature of the query image should be compared with all the images available on the Internet and a set of images, which are closer up to certain threshold are returned.

3. We want to search images of two cities, which belong to same country. The keyword search can include some false results. E.g. when we search for the cities Detroit and Flint together, we see some graphs, which are not the images of the cities, but refer their names in the graphics.

4. There is also no method available, which can return the result of following types of query. E.g. search the pictures about American history between the year 1900 and 1950.

There is no method of defining the queries between certain time range and/or any other metric. One of the problems of not getting the desired results is that there in no or little metadata available with the images available on the Internet. Second reason is that the algorithms employed by the search engines, does not have the capability to do search based on a specific criteria like these. As we can see in the above examples, that there is still a long way to be able to apply complex queries to search the images from the World Wide Web. In addition to above examples, we may encounter large number of other kinds of queries, which are not possible through existing search engines.

3. Literature Survey

The CBIR on World Wide Web involves two research areas: images classification and, images search and retrieval techniques.

3.1. Image Classification

In the literature, a wide variety of content-based retrieval methods and systems may be found. In [13] authors have reviewed about 200 references in CBIR up to the year 2000. There are three broad classes of applications user aims when using the system: search by association, search at a specific image, and category search. [14] identifies other patterns of use: searches for one specific image, general browsing to make an interactive choice, searches for a picture to go with a broad story, searches to illustrate a document. An attempt to formulate a general categorization of user requests for still and moving images are found in [15]. This and similar studies reveal that the range of queries is wider than just retrieving images based on the presence or absence of objects of simple visual characteristics. To describe the image, we have to extract certain low level features from it. There are a number of image processing operations that translate the image data into some other spatial data array. These operations may use local color, local texture, or local geometry. The main purpose of image processing in image retrieval must be to enhance aspects in the image data relevant to the query and to reduce the remaining aspects. There are several color representations like RGB, HSV, YUV and their variations.

Local shape characteristics derived from directional color derivatives have been used in [16] to derive perceptually conspicuous details in highly textured patches. In [17] a series of Gabor filters of different directions and scale have been used to enhance image properties [18]. Combining shape and color both in invariant fashion is a powerful combination as described by [19]. The texture is defined as all what is left after color and local shape have been considered or it is defined by such terms as structure and randomness. Basic texture properties include the Markovian analysis and other generalized versions [20, 21]. Other texture analysis methods are MRSAR-models [22], Wavelets [23], fractals [24] etc. A comparative study on texture classification from mostly transform-based properties can be found in [25].

In CBIR, the image is often divided in parts before features are computed from each part. There are four types of partitioning identified in [13]: string segmentation, weak segmentation, sign detection, data independent image partitioning. In [26] knowledge-based type abstraction hierarchies are used to access image data based on context and a user profile, generated automatically from cluster analysis of the database. Also in [27] the aim is to create a very large concept-space inspired by the thesaurus-based search from the information retrieval community. In [28] a variety of techniques is discussed treating retrieval as a classification problem. One approach is principal component analysis over a stack of images taken from the same class of objects. This can be done in feature space [29] or at the level of the entire image [30]. In [31] binary Bayesian classifiers are used to capture high-level concepts from low-level image features under the constraint that the test image belongs to one of the classes. Specifically, the hierarchical classification of vacation images is considered. At the highest level, images are classified as indoor or outdoor; outdoor images are further classified as city or landscape. Finally, a subset of landscape images is classified into sunset, forest, and mountain classes. A large number of systems have ignored two distinct characteristics of CBIR systems: the gap between high level concepts and low level features, subjectivity of human perception of visual content. A relevance feedback based approach has been suggested in [32]. Other interactive approaches have been suggested in [33, 34, 35]. Example include interactive region segmentation [36]; interactive database annotation [34, 37]; usage of supervised learning before the retrieval [38, 39]; and interactive integration of keywords and high level concepts to enhance image retrieval performance [40, 41]. In [42] an image retrieval system called SIMPLIcity (Semantics-sensitive Integrated Matching for Picture Libraries), which uses semantics classification methods, a wavelet-based approach for feature extraction. An integrated region matching based upon image segmentation, has been proposed. There are several domain-dependent ontology based systems [43, 44]. In [45] system uses a neural network to identify objects present in the images.

3.2. Image Search and Retrieval Techniques

There are a large number of papers published in the area of image search and retrieval. We are restricting our discussion here related to image search on World Wide Web. A system is implemented in [46] by which visual information on the web is (1) collected by agents, (2) processed in both text and visual feature domains, (3) catalogued and (4) indexed for fast search and retrieval. A typical web image search engine will first traverse the Web by following the hyperlinks between documents using several autonomous Web agents or spiders. These agents detect images and download and process them and add the new information about the image to the catalog.

A *perception-based search component,* which can learn users' subjective query concepts quickly through an intelligent sampling process, is proposed in [47]. A multi-resolution feature extractor extracts perceptual features from images and a high-dimensional indexer performs non-supervised clustering using Tree-structured Vector Quantization (TSVQ) [48] to group similar objects together. iFind is a web-based image retrieval system developed at Microsoft Research, China [49]. It provides the functionalities of text based image search, query by example, and their combination. Images in the database are indexed by their low-level (visual) features, high level (semantic) features (collected from image's environment), and optionally, annotations if they are available. In [50] MISE (The MediaSys Image Search Engine) is described. This system enables the users to search, to browse, to process, and to store images according to the combination of visual and textual features with meta-data related to the images. The MediaSys servers store the meta-data, visual and textual features, and the images themselves over a large scale distributed and heterogeneous system. The article [51] investigates what MPEG-7 means to Multimedia Database systems (MMDBSs) and vice-versa. It is argued that MPEG-7 has to be considered complementary to, rather than competing with, data models employed in it. [52] describes the use of stylesheets in the search and retrieval process of multimedia information, especially for audiovisual information. MPEG-7 has been used to describe the contents of the information. The use of stylesheets over the MPEG-7 data gives flexibility during both query formulation and the presentation of search results, and it allows a personalized way of querying and presenting.

4. The Proposed System

We discuses some of the example queries in section 2, which can not be answered by any of the existing systems to the best of our knowledge. We propose a system, which will exploit the XML technology and new MPEG-7 media metadata standard. In this section, we briefly describe the Image Integration Architecture.

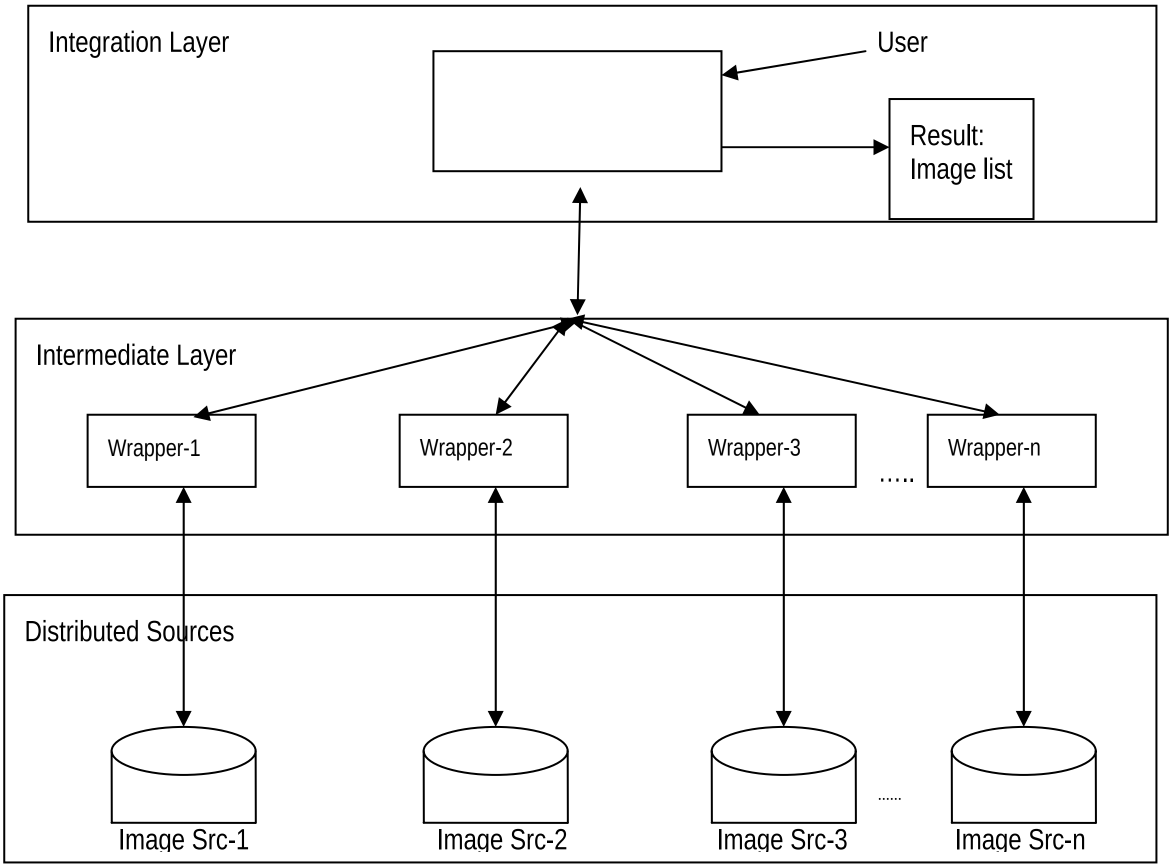

Figure 1. Image Integration Architecture

Figure 1 shows three-layered image integration architecture. At the lowest level, we have different Image sources. These sources have images and have not been designed on certain agreed schema. In other words, images in these sources may be in raw JPG, BMP, GIF or any other format without any semantic information. They may contain images clustered in certain groups. They may contain metadata in the form of MPEG-7 with partial annotation or they may contain MPEG-7 metadata with structured annotation. There may be other possibilities also.

Intermediate layer focuses on extracting image information by extracting low level features, metadata or any other semantic information available. If there is no semantic information available, we have to rely on low level features. We are considering images which are embedded in a webpage or stored in the image database. Each image source has to be treated in a different way.

Image source with raw image formats. At intermediate level, we extract low level features and store as MPEG - 7 metadata. Since this procedure has to be automatic, we can not do annotations at this level. There is no automatic annotation technique available so far.

Image source with raw image formats but clustered in groups. We extract low level features from the image and also store cluster information in MPEG - 7 metadata. Some intelligent technique has to be used to make cluster information useful in retrieval. We can also use traditional image search retrieval methods and look for important keywords stored in and around the image.

Image source with raw image format and with some metadata but not in MPEG – 7 standard. We extract low level features of the image and use the metadata while creating MPEG-7 metadata.

Image source with MPEG-7 partial or full annotation. We do not need to extract the low level features, since they are already available in MPEG-7 metadata.

Our emphasis here is to get information about all the images in MPEG-7 format, which is essentially XML data. Then we can use XML tools to query the images in Integration layer. There has been a lot of work on XML schema Integration [53, 54, 55, 56, 57, 58]. In this paper we are not discussing about XML schema integration.

The user will make a query at the top level using any of the methods using keyword, query by example, range queries discussed in section 2. This architecture may use agent based method or the popular directory based indexing method to search the image data sources. Integration process consists of querying the results returned by the intermediate layer, refine them according to user demand and return the results back to him/her. The relevance feedback and/or other long term learning technique can be used at the highest level to improve the results. The queries similar to the examples mentioned in section 2 can be successful if we combine low level features and semantic information together to produce the results. This architecture does not merely return the search results based on the keywords associated with the image, but also takes into account the low level features of the image.

5. Conclusion and Future Research

In this paper we suggest three - layered image integration architecture at a conceptual level. This approach takes care of images stored on the websites/image databases with or without semantic information. There are many challenges we have to face in this approach like selecting appropriate schema integration technique. MPEG-7, though already declared standard, will still take some time before images have their metadata stored in this format. Therefore it would be a grade mistake to rely on the assumption that metadata would be easily available in MPEG-7. Similarly, there are a large number of low level features suggested by different researchers, but MPEG-7 has included only some of them. There are possibilities that better features may be released in future and we have to consider these new features in any content-based image retrieval system. We are trying to set up an experimental environment based on the approach suggested in this paper, taking into account the methodology suggested in [59]. We are in the process of collecting the images with the properties described in section 4. We believe that the proposed system would enhance the quality of content based image retrieval.

Bibliography

[1] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker. Query by Image and video content: The QBIC system. IEEE Computer, vol. 28, no. 9, pp. 23-32, 1995.

[2] W. Y. Ma and B. S. Manjumath. Netra: A toolbox for navigating large image databases. ACM Multimedia System, vol. 7, pp. 184-198, 1999.

[3] Y. Rui, T. S. Huang. S. Mehrotra, and M. Ortega. A relevance feedback architecture for content-based multimedia information retrieval systems. IEEE Workshop on Content-based Access of Image and Video Libraries, pp. 82-89, 1997.

[4] S. Santini and R.Jain. Visual navigation in perceptual databases. International Conference on visual Information systems, San Diago, CA, Dec. 1997.

[5] J. R. Smith and S. Chang, Visually searching the Web for content, IEEE Multimedia, July-September 1997.

[6] S. Mukherjea, K. Hirata and Y. Hara. Towards a multimedia World Wide Web information retrieval engine. 6th WWW International Conference, S. Clara, CA, 6–11 May 1997.

[7] C. Meghini, f. Sebastiani and U. Straccia. Modelling the retrieval of structured documents contaning texts and images. 1st ECDL, Pis, Italy, Sep. 1997

[8] J.R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jain and C.F. Shu. The Virage image search engine: An open framework for image management. SPIE 96, 1996.

[9] M. Flickner et al., Query by image and video content: the QBIC system, IEEE Computer, 28(9), September 1995.

[10] PNG (portable Network Graphics) Specification, Version 1.2., http://www.libpng.org/pub/png/spec/

[11] Metadata Standards. http://www.chin.gc.ca/English/Standards/metadata_multimedia.html

[12] B S Manjunath et. El. Introduction to MPEG-7. John Wiley, 2002.

[13] Arnold W. M. Smeulders et. el. Content-Based Image Retrieval at the End of Early Years. IEEE transactions on Pattern Analysis and Machine Intelligence, vol. 22, No. 12, Dec 2000.

[14] S. Ornager. Image Retrieval: Theoretical and Empirical User Studies on Accessing Information in Images. 60th Am. Soc. Information Science Ann. Meeting, vol. 34, pp. 202-211, 1997.

[15] L. Armitage and P. Enser. Analysis of User Need in Image Archives. J. Information Science, vol. 23, no. 4, pp. 287-299, 1997.

[16] A. Mojsilovic, J. Kovacevic, J. Hu, R.J. Safranek, and S.K. Ganapathy. Matching and Retrieval Based on the Vocabulary and Grammar of Color Patterns. IEEE Trans. Image Processing, vol. 9, no. 1, pp. 38-54, 2000.

[17] B.S. Manjunath and W.Y. Ma. Texture Features for Browsing and Retrieval of Image Data. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, no. 8, pp. 837-842, Aug. 1996.

[18] R. Rodriguez-Sanchez, J.A. Garcia, J. Fdez-Valdivia, and X.R. Fdez-Vidal. The RGFF Representational Model: A System for the Automatically Learned Partitioning of `Visual Pattern' in Digital Images. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21, no. 10, pp. 1,044-1,073, Oct. 1999.

[19] T. Gevers and A.W.M. Smeulders. Content-Based Image Retrieval by Viewpoint-Invariant Image Indexing. Image and Vision Computing, vol. 17, no. 7, pp. 475-488, 1999.

[20] S. Krishnamachari and R. Chellappa. Multiresolution Gauss-Markov Random Field Models for Texture Segmentation. IEEE Trans. Image Processing, vol. 6, no. 2, 1997.

[21] G.L. Gimel'farb and A.K. Jain. On Retrieving Textured Images from an Image Database. Pattern Recognition, vol. 29, no. 9, pp. 1,461-1,483, 1996.

[22] J. Tatemura. Browsing Images Based on Social and Content Similarity. Proc. Int'l Conf. Multimedia and Expo, 2000.

[23] I. Daubechies. Ten Lectures on Wavelets. Philadelphia: SIAM, 1992.

[24] L.M. Kaplan et al. Fast Texture Database Retrieval Using Extended Fractal Features. Storage and Retrieval for Image and Video Databases, VI, vol. 3,312, pp. 162-173, SPIE Press, 1998.

[25] T. Randen and J.H. Husoy. Filtering for Texture Classification: A Comparative Study. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21, no. 4, pp. 291-310, Apr. 1999.

[26] C.C. Hsu, W.W. Chu, and R.K. Taira. A Knowledge-Based Approach for Retrieving Images by Content. IEEE Trans. Knowledge and Data Eng., vol. 8, no. 4, pp. 522-532, 1996.

[27] H. Chen, B. Schatz, T. Ng, J. Martinez, A. Kirchhoff, and C. Lim. A Parallel Computing Approach to Creating Engineering Concept Spaces for Semantic Retrieval: The Illinois Digital Library Initiative Project. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, no. 8, pp. 771-782, Aug. 1996.

[28] N. Vasconcelos and A. Lippman. A Probabilistic Architecture for Content-Based Image Retrieval. Proc. Computer Vision and Pattern Recognition, pp. 216-221, 2000.

[29] H. Murase and S.K. Nayar. Visual Learning and Recognition of 3D Objects from Appearance. Int'l J. Computer Vision, vol. 14, no. 1, pp. 5-24, 1995.

[30] R.W. Picard and T.P. Minka. Vision Texture for Annotation. Multimedia Systems, vol. 3, pp. 3-14, 1995.

[31] Aditya Vailaya et. el. Image Classification for content-Based Indexing. IEEE reanscations on Image Processing, Vol. 10, No. 1, Jan. 2001.

[32] Yong rui et. el. Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval. IEEE transactions on circuits and Video Technology. Volume 8, Number 5, 1998, pp. 644-655.

[33] A. D. Narasimhalu, Multimedia Syst. (Special Section on Content-Based Retrieval), 1995.

[34] W. Niblack, R. Barber et al. The QBIC project: Querying images by content using color, texture and shape. In Proc. SPIE Storage and Retrieval for Image and Video Databases, Feb. 1994.

[35] P. P. Ohanian and R. C. Dubes. Performance evaluation for four classes of texture features. Pattern Recognition, vol. 25, no. 8, pp. 819–833, 1992.

[36] M. Ortega, Y. Rui, and K. Chakrabarti, S. Mehrotra, and T. S. Huang. Supporting similarity queries in MARS. In Proc. ACM Conf. Multi-media, 1997.

[37] A. Pentland and R. Picard. IEEE Trans. Pattern Anal. Machine Intell. (Special Issue on Digital Libraries), 1996.

[38] A. Pentland, R. W. Picard, and S. Sclaroff. Photobook: Content-based manipulation of image databases. Int. J. Comput. Vision, 1996.

[39] R. W. Picard and T. P. Minka. Vision texture for annotation. Multi-media Syst. (Special Issue on Content-Based Retrieval).

[40] J. Dowe. Content-based retrieval in multimedia imaging. In Proc. SPIE Storage and Retrieval for Image and Video Databases, 1993.

[41] Y. Rui, T. S. Huang, and S. Mehrotra. Content-based image retrieval with relevance feedback in MARS. In Proc. IEEE Int. Conf. ImageProcessing, 1997.

[42] J. Z. Wang, G. Li, and G. Wiederhold. SIMPLIcity: Semantics-sensitive Integrated Matching for Picture LIbraries. In IEEE Trans. on pattern Analysis and Machine Intelligence, volume 23, pages 947--963, 2001.

[43] C. Breen, L. Khan, and A. Kumar. Image Classification Using Neural Networks and Ontologies. IEEE DEXA, International Workshop on Web Semantics, France, Sept 2002.

[44] L. Khan and D. McLeod. Audio Structuring and Personalized Retrieval Using Ontologies. IEEE Advances in Digital Libraries, Library of Congress, Washington, DC, May 2000.

[45] Casey Breen, Latifur Khan, Arun Kumar, and Lei Wang. Ontology-based image classification using neural networks", SPIE 2002

[46] J. R. Smith, S.-F. Chang, "Searching for Images and Videos on the World-Wide Web," Columbia University, No. CU/CTR/TR 459-96-25, Aug. 1996.

[47] Wei-Cheng Lai, Edward Chang, and Kwang-Ting (Tim) Cheng. An Anatomy of a Large-scale Image Search Engine. WWW 2002, 7-11 May 2002, Honolulu, Hawaii.

[48] A. Gersho and R. Gray. Vector Quantization and Signal Compression. Kluwer Academic, 1991

[49] Hong-Jiang Zhang, Zheng Chen, Wen-Yin Liu and Mingjing Li. Relevance Feedback in Content-Based Image Search. Invited Keynote, 12th Int. Conf. on New Information Technology (NIT), May 29-31, Beijing.

[50] Panrit Tosukhowong, Frederic Andres, Kinji Ono, Jose Martinez, Noureddine Mouaddib, Nicolas Dessaigne and Douglas C. Schmidt A Flexible Image Search Engine. ACM, MM99, Oct 30 - Nov 5, 1999, Orlando, Florida

[51] Harold Kosch. MPEG-7 and Multimedia Database Systems. SIGMOD Record, Vol. 31, No.2, June 2002.

[52] Mark van Setten, Erik Oltmas, Mettina Veenstra. Personalized Video Search and Retrieval using MPEG-7 and Stylesheets. https://doc.telin.nl/dscgi/ds.py/Get/File-8842/

[53] V. S. Subrahmanian, Sibel Adali, Anne Brink, Ross Emery, James J. Lu, Adil Rajput, Timothy J. Rogers, Robert Ross, and Charles Ward: HERMES: A heterogeneous Reasoning and Mediator System. Technical Report, University of Maryland, Maryland, 1995.

[54] S. Chawathe, H. Garcia-Molina, J. Hammer, K. Ireland, Y. Papakonstantinou, J. Ullman, and J. Widom: The TSIMMIS Project: Integration of Heterogeneous Information Sources. IPSJ Conference, pp. 7-18, Tokyo, Japan, October 1994.

[55] Rajeev Agrawal, Mukesh Mohania, Yahiko Kambayashi, S S Bhowmick, S Madria: An Architecture for XML Schema Integration. ICDL: Research and Practices, Kyoto, Japan, 2000.

[56] Ralf Behrens: A Grammar Based Model for XML Schema Integration. BNCD, 2000

[57] Ronaldo dos Santos Mello, Carlos Alberto Heuer: A Bottom-Up Approach for Integration of XML Sources. WIIW 2001

[58] Dongwon Lee, Murali Mani, Wesley W. Chu: Effective Schema Conversions between XML and Relational Models. In European Conf. on Artificial Intelligence (ECAI), Knowledge Transformation Workshop (ECAI-OT), Lyon, France, July 2002.

[59] P. Stanchev, G. Amato, F. Falchi, C. Gennaro, F. Rabitti, P. Savino, Selection of MPEG-7 Image Features for Improving Image Similarity Search on Specific Data Sets, 7-th IASTED International Conference on Computer Graphics and Imaging, CGIM 2004, Kauai, Hawaii, 395-400, 2004.

Authors' Information

Rajeev Agrawal – Kettering University, Flint, MI 48504, USA; Email: ragrawal@kettering.edu

Farshad Fotouhi – Wayne State University, Detroit, MI 48202, USA; Email: fotouhi@cs.wayne.edu

Peter Stanchev – Kettering University, Flint, MI 48504, USA; Email: pstanche@kettering.edu

Ming Dong – Wayne State University, Detroit, MI 48202, USA; Email: mdong@cs.wayne.edu