International Journal of Image and Graphics, Vol. 1, No. 1 (2001) 83–91


# CONTENT-BASED RETRIEVAL OF MULTIMEDIA INFORMATION

(DAVID) DAGAN FENG*

*Biomedical and Multimedia Information Technology (BMIT) Group,*
*Basser Department of Computer Science,*
*The University of Sydney, NSW 2006, Australia*
*and*
*Center for Multimedia Signal Processing (CMSP),*
*Department of Electronic & Information Engineering,*
*The Hong Kong Polytechnic University, Hong Kong, P. R. China*

The recent information explosion has led to massively increased demand for multimedia data storage and retrieval techniques. Content-based retrieval is an important alternative and complement to the traditional keyword-based searching for multimedia data and can greatly enhance information management. For the last ten years, the Biomedical and Multimedia Information Technology (BMIT) Group and recently the Center for Multimedia Signal Processing (CMSP) have conducted systematic studies and research activities on this topic. Some of the works relating to content-based image/video retrieval and their applications are briefly presented in this paper.

*Keywords*: Content-Based Retrieval; Image/Video Data Management; Multimedia Application.

## 1. Introduction

With the remarkable growth in the use of computers to process multimedia data, there is an ever-increasing demand for the storage and retrieval of such data. As a result, multimedia database applications have become an area of intense research interest. Since digital image and video are the most important components of multimedia data, retrieval techniques for image and video are becoming the major and crucial issues for multimedia database applications. Due to the rapid increase of digital image/video collection, there is a pressing need for the development of efficient methods of image/video browsing, searching, and retrieval which are quite different from the traditional keyword-based information query. For the last ten years, the Biomedical and Multimedia Information Technology (BMIT) Group and recently the Center for Multimedia Signal Processing (CMSP) have conducted systematic studies and research activities including those on content-based retrieval of multimedia information particularly for image and video data. In this paper,

*E-mail: feng@cs.usyd.edu.au

we briefly present some of the related research works and their applications in biomedicine.

## 2. Content-Based Image Retrieval

In recent years, various content-based image retrieval systems have been developed in research prototypes and commercial systems which include the QBIC (Query By Image Content) system at IBM Almaden,[1] the Photobook system at the MIT Media Lab,[2] the Virage system developed by Virage,[3] the CANDID system at the Los Alamos National Laboratory,[4] and the MARS system at the University of Illinois at Urbana–Champaign.[5] These content-based image retrieval systems usually use some major visual features such as color,[6–8] texture[9–11] and shape.[12–14] In addition, some other features such as icons,[15] Wold[16] and eigen vector features[17] may also be included. Color is one of the most widely used visual features. It is relatively robust to background complication and independent of image size and orientation. As one of the major visual features, texture plays an important role in human perception. It refers to visual patterns with homogeneity properties that cannot be provided by color and intensity alone. Shape includes boundary-based and region-based shape representations such as the Fourier Descriptor and Moment Invariants. These features are extracted automatically in the indexing process when images are entered into a multimedia database. Queries and retrievals can be based directly on the visual properties of the images and return results ranked by the degree of content matching.

However, the above visual features cannot perform consistently over a variety of images. Most visual features in indexing and retrieval either take statistical information but loose spatial information or take spatial information but loose statistical information. Therefore, we have proposed a new visual feature which contains both spatial and statistical information for the similarity measurement in content-based image retrieval.[18]

The retrieval method we developed is based on the Radon transform.[19] It actually calculates the projection of images in the parameter space. The projection is normalized in the parameter space to preserve the translation and rotation invariants. The new visual feature is also insensitive to lighting, reflection and contrast variations, and noise. A look-up table is used to speed up the transformation. The computational complexities of signature calculation and similarity comparison are $O(n)$ and $O(\sqrt{n})$ respectively. We tested the new retrieval method by using a set of 158 images from PhotoDisc image library (volume OS8). Figure 1 shows the retrieval results using our new method. The query image is the first image on the top-left corner. The similarity of each image to the query image is displayed at the right corner of the image. The results demonstrate that our new retrieval method works well and the query can emphasize three kinds of characteristics of such images: vertical elongated shape (shape); one region on top of another region (composition); and the top region darker than the bottom region (spatial

Fig. 1. Query results using the proposed retrieval method.

distribution). It calculates the statistical information while preserving geometrical distribution information in the image. The new retrieval method is robust against noise and efficient in calculation when certain mapping and indexing technologies are used.

## 3. Content-Based Video Retrieval

There are also many ongoing developments in video retrieval systems such as VideoQ system at the Columbia University,[20] VIDSEEK system at the University of Kansas,[21] and Visual-Retrieve-Ware at Excalibur Inc.[22] In these retrieval systems, color content, color structure, texture, shape, and motion were used as visual features. We have recently set up a Semantic and Visual Content-based Retrieval system (SVCR) for news video.[23] The SVCR provides multiple ways of video query. Users can query specific video sessions not only by semantic meaning like peoples, places, topics, etc. but also by particular visual contents like colors and textures. Figure 2 shows the main GUI of the SVCR.

For color feature extraction, we developed a new color histogram method which incorporates nonuniform quantization and time relevant weighting in contrast to the traditional color histogram that is based on the uniform quantization technique. A color-based video retrieval software has been developed as a component for feature extraction subsystem of SVCR.

For texture feature analysis, we proposed a new method[24] of region mosaic based Wold decomposition which exploits the spatial distribution of harmonic peak points. Compared with other Wold decomposition methods, our experimental results indicate this new approach can achieve better retrieval performances. Based on this technique, a texture based image retrieval for Brodatz database and a video retrieval subsystem have been developed.

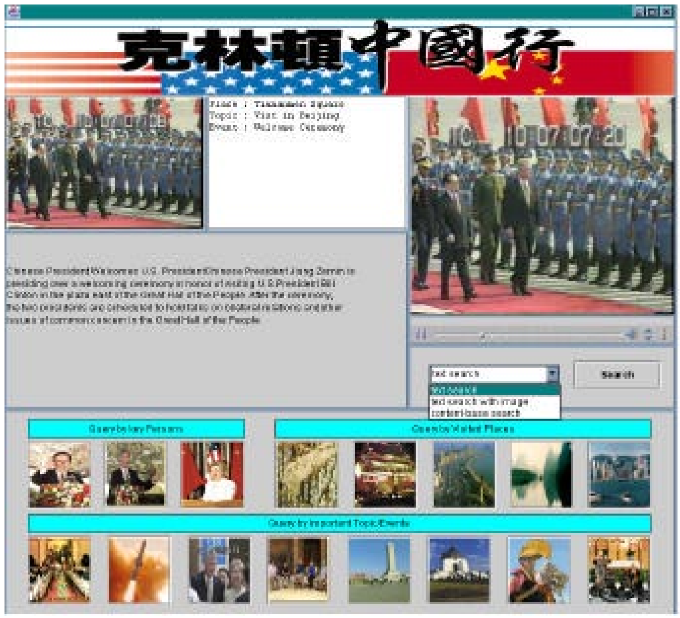


Fig. 2. The main GUI of SVCR system.

We have also developed a dynamically adjustable data structure for content-based video retrieval system management. In this structure, video attributes and indexing methods are distributed and dynamically stored on different attribute agents and indexing agents, and video attribute and indexing methods are separately encapsulated and dynamically linked with a reference mechanism.

For semantic feature extraction, our work mainly focuses on automatic video segmentation and semi-automatic annotation based on visual features and human interaction. Other works include mosaic-based video compression for video database, object motion tracking for video object extraction, etc.

## 4. Content-Based Retrieval of Biomedical Multimedia Data

In the past three decades, medical imaging techniques have advanced rapidly providing powerful tools for patient diagnosis, treatment, and surgery. These techniques include ultrasound, X-ray computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) and single photon emission computed tomography (SPECT). Especially, functional imaging systems such as PET and SPECT provide image-wide quantification of physiological and biochemical processes within the body.[25] For example, dynamic $[^{18}F]$ *2-fluoro-deoxy-glucose* (FDG) PET images which are widely used to determine the local cerebral metabolic rate of glucose (LCMRGlc) depict the glucose consumption and energy requirements of various structural and functional components in human brain.[26] Such quantitative physiological information inside the functional image content is unlikely to be retrieved by common image retrieval techniques using color, texture or shape index.

We have recently presented a prototype design for a content-based functional image retrieval database system (FICBDS) for dynamic PET images[27] based on dynamic feature classification and a three-step functional image data compression technique which can achieve very high compression ratios without degradation of image quality.[28] Pixel kinetics are encoded during image data compression to achieve image indexing and compression simultaneously. The proposed functional image retrieval system not only supports efficient content-based retrieval based on physiological kinetic features of the stored functional images but also greatly reduces image storage requirements.

The physiological tissue time-activity curve (TTAC) feature extraction from functional image contents can be implemented using a knowledge-based cluster analysis (KCA). In dynamic PET images, for each pixel, a physiological TTAC can be extracted from a sequence of image frames. However, pixels in physiologically similar regions should have similar kinetics. We therefore, applied the KCA method to automatically classify image-wide TTACs into a certain number of typical TTAC groups corresponding to different physiological kinetic patterns. The clustering algorithm classifies the image-wide TTACs $Z_i(t)$ (where $i = 1, 2, \ldots, R, R$ is the total number of image pixels and $t = 1, 2, \ldots, 5$) into $S$ cluster groups $C_j$ (where $S << R$ and $j = 1, 2, \ldots, S$) based on the magnitude of natural association (a similarity measure). The mean TTAC values for each identified cluster group are stored in an index table, indexed by cluster group. The averaging yields a set of highly smoothed TTAC features. The average TTACs represent $j$ distinct physiological behaviors.

In the FICBDS database system, we anticipate that content-based functional image retrieval on TTAC characteristics will be the most useful query method. In this query method, the user defines a TTAC feature vector or selects from a set of pre-defined TTAC samples as a querying example. Images containing similar TTAC feature vectors are then retrieved from the database. Let $\mathbf{TTAC}_{EX}$ be the example TTAC feature vector and $\mathbf{TTAC}_{DB}$ be the all extracted TTAC feature vectors stored in the database. The measure of similarity can be computed by $D_E(\mathbf{TTAC}_{EX}, \mathbf{TTAC}_{DB})$, the Euclidean distance between $\mathbf{TTAC}_{EX}$ and $\mathbf{TTAC}_{DB}$. All $\mathbf{TTAC}_{DB}$ satisfying $\{D_E(\mathbf{TTAC}_{EX}, \mathbf{TTAC}_{DB}) \leq M\}$ are the final retrieval results where $M$ is a user-defined TTAC matched degree.

To assess the feasibility of our content-based functional image retrieval database design, we used a set of dynamic PET images to create a prototype database. The image data sets with 15 clinical FDG PET studies were acquired by a SIEMENS ECAT 951R PET scanner at the PET and Nuclear Medicine Department, Royal Prince Alfred Hospital, Sydney. The number of cross-sectional image planes was 31. A typical conventional sampling schedule (CSS) consisting of 22 temporal frames

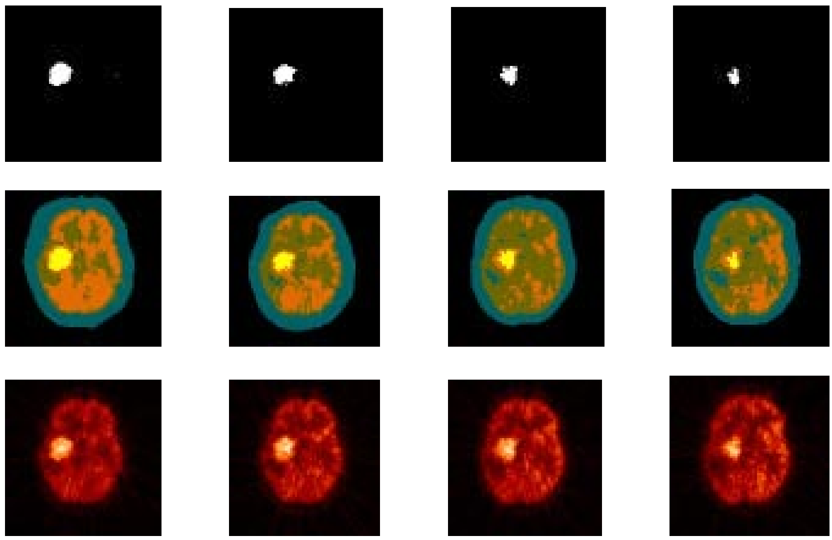

Fig. 3. The results of TTAC query.

was used to acquire the PET projection data. The prototype database contained 465 plane images (15 studies, 31 planes for each study). Figure 3 shows some results of TTAC query with a high metabolic rate. Images in the first row are tissue regions with similar TTAC features to the sample TTAC. Their cluster index images and original temporal images are shown in the second and third rows respectively. The results show the query successfully identified brain regions with relatively high metabolic rate. It is possible to identify and retrieve images containing tissue regions which exhibit physiological behavior similar to a pre-defined pattern. Our content-based functional image database system could facilitate identification of patients with particular disease states and potentially increase our understanding of underlying disease processes and improve specificity in diagnosing disease.

Our content-based functional image database retrieves index images corresponding to specific physiological kinetic features. A dynamic image sequence representing the original kinetic information and image data are then reconstructed from the index image. Kinetic feature indexing allows our database to avoid the need to store the raw image data but still enables rapid access to the dynamic functional image data. In contrast, conventional image retrieval techniques based on features such as color, texture and shape require the entire image to be stored in the database. Also, since dynamic images are compressed and indexed simultaneously, the complexity of database design and implementation is significantly reduced. The reduction in database size speeds web-based access to dynamic data for interpretation by the physician.

In order to facilitate examination of functional image data via the Internet, we have recently developed a Java-based functional image browser (FIB) system.

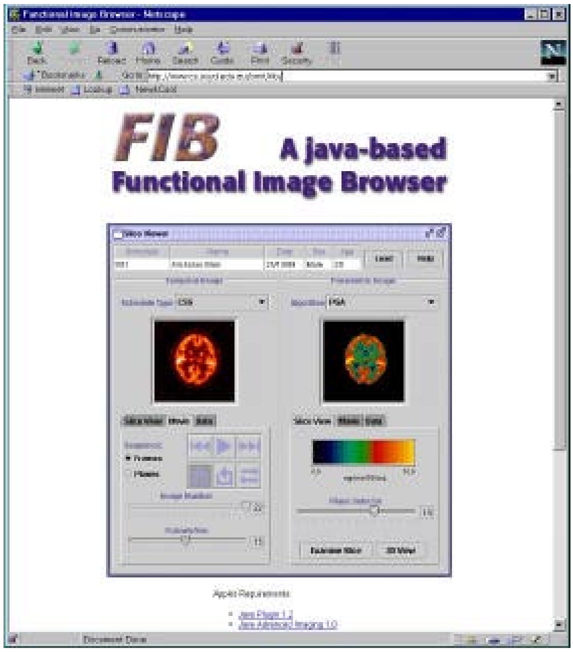


Fig. 4. The web-page of the Functional Image Browser (FIB).

The FIB supports on-line temporal image and physiological parametric image slice viewer, dynamic temporal image and parametric image movie playback, slice examiner, and 3D viewer. Figure 4 shows the web-page of the FIB.

## 5. Summary

Content-based retrieval is an important alternative and complement to the traditional keyword-based searching for multimedia data and can greatly enhance information management. It is expected to play an important role in web-based multimedia information systems.

## Acknowledgments

This research is supported by the ARC and UGC Grants.

## References

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, *IEEE Comput.* **28**(9), 23 (1995).
2. A. Pentland, R. W. Picard, and S. Sclaroff, *Proc. SPIE* **2185**, 34 (1994).
3. J. R. Bach, C. Fuller, A. Gupta, A. Hampaur, B. Horowitz, R. Humphrey, R. Jain, and C. Shu, *Proc. SPIE* **2670**, 76 (1996).

4. P. M. Kelly and T. M. Cannon, *Proc. IS&T/SPIE* **2670**, 42 (1996).
5. Y. Rui, T. S. Huang, and S. Mehrotra, *Proc. IEEE Int. Conf. Image Processing '97*, p. II815, CA, 1997.
6. E. Binaghi, I. Gagliardi, and R. Schettini, *Proc. IFIP Working Conf. Visual Database Systems II* (Elsevier Science Publishers, 1992), p. 79.
7. M. J. Swain, *Proc. SPIE* **1908**, 193 (1993).
8. J. R. Smith and S. F. Chang, *Proc. SPIE* **2670**, 98 (1995).
9. J. Weszka, C. Dyer, and A. Rosenfeld, *IEEE Trans Sys. Man and Cyb.* **SMC-6**(4), 211 (1976).
10. P. P. Ohanian and R. C. Dubes, *Pattern Recognition* **25**(8), 819 (1992).
11. H. Tamura, S. Mori, and T. Yamawaki, *IEEE Trans. Sys. Man and Cyb.* **SMC-8**(6), 322 (1978).
12. W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin, *Proc. SPIE* **1908**, 173 (1993).
13. K. Tanabe and J. Ohya, *Progress in Image Analysis and Processing* (Wiley, NY, 1989), p. 138.
14. B. M. Mehtre, M. Kankanhalli, and W. F. Lee, *Information Processing and Management* **33**(3), 160 (1997).
15. F. Rabitti and P. Stanchev, *Proc. IFIP Working Conf. Visual Database Systems* (1989), p. 415.
16. R. W. Picard, T. Kabir, and F. Liu, *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, p. 638, NY, 1993.
17. A. Pentland, R. W. Picard, and S. Sclaroff, *Proc. SPIE*, **2185**, 34 (1994).
18. H. Wang, F. Guo, D. Feng, and J. Jin, *Proc. 6th ACM Int. Multimedia Conf. MM'98*, p. 229, UK, 1998.
19. A. K. Jain, *Fundamentals of Digital Image Processing* (Prentice Hall, NY, 1989).
20. S. F. Chang *et. al.*, *IEEE Trans. Circuits and Systems for Video Technology* **8**, 938 (1998).
21. http://www.tisl.ikans.edu/∼sgauch/DVLS.html.
22. http://vrw.exalib.com:8015/cst.
23. F. H. Long, D. Feng, F. N. Yik, J. C. Ren, and W. C. Siu, in *Multimedia Videos and Applications — Proc. Int. Symp. Signal Processing and Intelligent System*, p. 639, China, 1999.
24. F. H. Long, W. Y. Chan, D. Feng, and W. C. Siu, in *Multimedia Videos and Applications — Proc. Int. Symp. Signal Processing and Intelligent System*, p. 628, China, 1999.
25. D. Feng, D. Ho, H. Iida, and K. Chen, in *Medical Imaging Systems Techniques and Applications: General Anatomy*, ed. C. T. Leondes (Gordon and Breach Science Publishers, Amsterdam, 1997), p. 85.
26. D. Feng, D. Ho, K. Chen, L. C. Wu, J. K. Wang, R. S. Liu, and S. H. Yeh, *IEEE Trans. Med. Imag.* **14**, 697 (1995).
27. W. Cai, D. Feng, and R. Fulton, *IEEE Trans. Info. Tech. Biomed.* **4**(2), 152 (2000).
28. D. Ho, D. Feng, and K. Chen, *IEEE Trans. Info. Tech. Biomed.* **1**(4), 219 (1997).

**(David) Dagan Feng** (S'88-M'88-SM'94) received his M.E. degree in Electrical Engineering & Computing Science (EECS) from Shanghai JiaoTong University in 1982, M.Sc. degree in Biocybernetics and Ph.D. degree in Computer Science from the University of California, Los Angeles (UCLA) in 1985 and 1988 respectively. After briefly working as Assistant Professor at the University of California, Riverside, he joined Department of Computer Science, the University of Sydney at the end of 1988 as Lecturer, Senior Lecturer, Reader, and then Professor. Since 1997, he has also been appointed as Professor at the Department of Electronic & Information Engineering, Hong Kong Polytechnic University. He is currently Head of Department of Computer Science, the University of Sydney.

Prof. Feng's research interests include Biomedical & Multimedia Information Technology, Functional Imaging, Modeling & Simulation, Fast Algorithms & Data Compression. He has published over 200 scholarly research papers, made several landmark contributions in his field, and received a number of awards including the Crump Prize for Excellence in Medical Engineering from USA. He is Founder and Director of the Biomedical & Multimedia Information Technology (BMIT) Group at the University of Sydney. He is also Deputy Director of the Center for Multimedia Signal Processing (CMSP) at the Hong Kong Polytechnic University.

He is Special Area Editor for the IEEE Transactions on Information Technology in Biomedicine (IEEE-T-ITB), Guest Editor for the Special Issue on Multimedia Information Technology in Biomedicine for IEEE-T-ITB, and Vice-Chair of International Federation of Automatic Control (IFAC), Technical Committee on BIOMED. Prof. Feng has also been appointed as Honorary Research Consultant at the Royal Prince Alfred Hospital, Australia, Guest Professor at Northwestern Polytechnic University, Shanghai JiaoTong University and Tsinghua University, China.