Local Concept-Based Medical Image Retrieval with Correlation-Enhanced Similarity Matching Based on Global Analysis

Md Mahmudur Rahman, Sameer K. Antani, George R. Thoma
U.S. National Library of Medicine,
National Institutes of Health,
Bethesda, MD, USA
{rahmanmm,santani,gthoma}@mail.nih.gov

Abstract

A correlation-enhanced similarity matching framework for medical image retrieval is presented in a local concept-based feature space. In this framework, images are presented by vectors of concepts that comprise of local color and texture patches of image regions in a multi-dimensional feature space. To generate the concept vocabularies and represent the images, statistical models are built using a probabilistic multi-class support vector machine (SVM). For the similarity search, the concept correlations in the collection as a whole are analyzed as a global thesaurus-like structure and incorporated in a similarity matching function. The proposed scheme overcomes some limitations of the "bag of concepts" model, such as the assumption of feature independence. A systematic evaluation of image retrieval on a biomedical image collection of different modalities demonstrates the advantages of the proposed retrieval framework in terms of precision-recall.

1. Introduction

Medical and health-care service is a big industry, which is directly related to every citizen's quality of life. Among all the issues in this sector, improving the image based medical diagnosis efficiency and accuracy is one of the main focuses in both research and practitioner circles. In recent years, rapid advances in software and hardware technology facilitate the generation and storage of large collections of biomedical images by hospitals and clinics every day. These images of various modalities constitute an important source of anatomical and functional information for the diagnosis of diseases, medical research and education [1, 2]. The exponential growth of the biomedical image data has created a compelling need for innovative tools for managing, retrieving, and visualizing images from large collections to support the clinical decision making.

Medical images are commonly stored, retrieved and transmitted in the DICOM (Digital Imaging and Communication in Medicine) format in a Picture Archiving and Communications System (PACS) [3] and image search is carried out on the textual attributes, such as person information, other health meta data, often found in image headers. The available attributes (annotations) are often very brief and there is no header information attached for many biomedical images other then the DICOM format. The search results might be improved by combining the text attribute-based search capability with low-level visual content, commonly known as the content-based image retrieval (CBIR) [4]. In CBIR, the access to information is performed at the perceptual level based on automatically extracted visual features (e.g., color, texture, shape, edge, etc.) with appropriate similarity model.

During the last decade, several medical CBIR prototypes have been proposed [2, 5, 6, 7, 8]. Majority of these are developed around a specific imaging modality, e.g., the ASSERT system [5] is designed for high resolution computed tomography (HRCT) images of the lung and the SPIRS system [6] for digitized X-rays of the spine. These systems are task-specific and cannot be transferred to other domains or modalities. To date, few research projects have a goal create CBIR systems for heterogeneous image collections [7, 8]. For example, the IRMA (Image Retrieval in Medical Applications) system [7] is an ongoing project that retrieves from a large set of radiological images of different anatomical regions, acquisition views, and biological systems based on various low-level texture features. The medGIFT project [8] is based on the open source image retrieval engine GNU Image Finding Tool (GIFT). It aims to retrieve diverse medical images where a very high-dimensional feature space of various low-level features is used as visual terms analogous to the use of keywords in a text-based retrieval approach.

In general, CBIR systems for the medical image collection can be classified on the basis of their task specification as well as image feature extraction and processing capabilities. Some systems utilize low-level visual features computed over the entire image while others compute the features on localized regions. Although there exists a strong correlation between the segmented regions and the regions of interest (ROI) in medical images, accurate and semantically valid automatic segmentation is a very difficult task and requires some kind of user intervention. These low-level global or local features have limited discriminative power to close the *semantic gap*, which is the extent of mismatch between user's requirements as high-level concepts and the low-level representation of images [4].

In an effort to minimize the *"semantic gap"*, some recent approaches in general domain have used machine learning techniques on locally computed image features to represent images based on the *"bag of concepts"* model [9, 10, 11]. The model is applied to images by using a visual analogue of a word (e.g., *"bag of words"*) in text documents by automatically extracting different predominant local color and texture patches. In general domain, it has shown that images represented as *"bags of concepts"* are suitable for scene classification and showing impressive levels of performance [9]. For example, a semantic modeling approach is investigated in [10] for a small collection of images based on the binary classification of semantic patches of local image regions. Similarly, binary SVM is employed for the model generation of 23 selected concepts in [11] for natural photographic images. In the testing stage, unlabeled regions are fed into all the models, the concept from the model giving the highest positive result is associated with the region. Recently, a similar approach is applied [12] in medical domain with improved performances when compared to using only low level features. However, in the *"bag of concepts"* model, each concept is considered independent of all the other concepts besides the loss of all ordering structure. This independent assumption might not hold in many cases as in general there exists correlated or co-occurring concepts in individual images as well as in a collection as a whole. For example, there is a higher probability of co-occurrence between the white teeth and red color tissue of the mouth in a dental photographic image whereas most of the time the teeth are surrounded by jaw bones and black background in dental X-ray images. In these examples, individual objects, such as teeth, mouth tissue, jaw bones and black background can be considered as the local concepts with their distinct color and texture properties. So, there is indeed a need to exploit the correlation or co-occurrence patterns among the concepts to improve the effectiveness of the retrieval process.

To overcome the limitations of both the low-level and concept-based feature representations, this paper presents a correlation-enhanced similarity matching technique for medical images retrieval. In this approach, the concept similarities and correlations are analyzed in the collection as a whole to construct global similarity thesauruses, which are finally utilized in a quadratic form of the distance measure to compare query and target images in a database. The rest of the paper is organized as follows. In Section 2, the local concept-based image representation approach is described. Section 3 presents the similarity matching approach based on the global correlation analysis. Exhaustive experiments and analysis of the results are presented in Sections 4 and 5. Finally, Section 6 provides our conclusions.

2. Image Representation on Local Concept Space

By the term *"local concept"*, we refer to the perceptually and semantically distinguishable local patches or regions in individual images. For example, in a heterogeneous collection of medical images, it is possible to identify such local patches as homogeneous texture patterns in grey level radiological images, and differential color and texture structures in microscopic pathology and dermoscopic images. The variation in these local patches can be effectively modeled by using supervised learning based classification techniques such as the SVM [13]. In this context, the SVM will create statistical models for concepts from the training data, where an instance (e.g., local patches) in the training set is represented by a feature vector and contains category specific labels.

In its basic formulation, the SVM is a binary classification method that constructs a decision surface and maximizing the inter-class boundary between the samples. Stated mathematically,

$$f(\mathbf{x}) = \text{sign}\left(\sum_{i=1}^{N} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b\right) \quad (1)$$

where $\mathbf{x} \in \Re^d$ is an input vector, $\mathbf{x}_i$ is a training sample vector along with its label $y_i \in (+1, -1)^N$, b is a bias and K is a kernel function which maps the vectors into a higher dimensional space by the non-linear mapping function $\phi : \Re^d \rightarrow \Re^l$, where $l > d$ or l could even be infinity.

A number of methods have been proposed for the extension of the SVM to multi-class classification problems [14]. We utilize one such a method by combining all pairwise comparisons of binary SVM classifiers, known as *one-against-one* or pairwise coupling (PWC) [16]. The first step of this approach is to construct a training set of local semantic patches from individual image regions. We consider a semi-automatic approach for the semantic patch generation. In this approach, the training images are at first fixed partitioned equally into an even grid of non-overlapping regions. Due to the fixed partitioning scheme, some regions or patches would contain multiple local concepts. Hence,

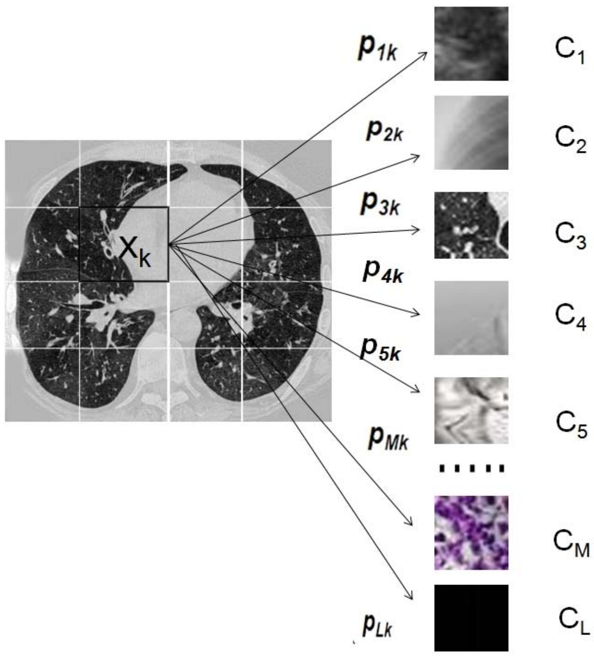

Figure 1. Probabilistic membership scores of a region

instead of considering all the patches from fixed partitioning, a subset of these are manually selected based on the criteria that the area of each region should correspond to a particular concept by at least 70-80%.

In order to perform the multi-class SVMs training based on the local concept categories, a set of L labels are assigned as $C = \{c_1, \cdots, c_i, \cdots, c_L\}$, where each $c_i \in C$ characterizes a local concept category. The training set of the local patches is annotated manually with the concept labels in a mutually exclusive way. Hence, each patch is labeled with only one local concept category and is represented by a combination of color and texture moment-based features. The color feature is composed of the mean and the standard deviation of each channel in the RGB color space and the texture feature is composed of second order texture moments, such as energy, maximum probability, entropy, contrast, and inverse difference moment from a grey level co-occurrence matrix (GLCM) [15]. For SVM training, the initial input to the system is the feature vector set of the patches along with their manually assigned corresponding concept labels.

2.1. Image Encoding and Feature Representation

Images in the data set are annotated with local concept labels by partitioning each image I_j into an equivalent $r \times r$ grid of l vectors as $\{\mathbf{x}_{1_j}, \cdots, \mathbf{x}_{k_j}, \cdots, \mathbf{x}_{l_j}\}$, where each $\mathbf{x}_{k_j} \in \Re^d$ is a d-dimensional combined color and texture feature vector. For each $\mathbf{x}_{k_j}$, the local concept category probabilities are determined by the prediction of the multi-class SVMs as [16]

$$p_{ik_j} = P(y = i \mid \mathbf{x}_{k_j}),\ 1 \le i \le L. \qquad (2)$$

For example, Figure 1 shows a particular region in a segmented image and its probabilistic membership scores to different local concept categories. The probability or confidence scores of all categories now forms an L-dimensional vector for a region x_{k_j} of I_j as

$$\mathbf{p}_{k_j} = [p_{1k_j} \cdots p_{ik_j} \cdots p_{Lk_j}]^{\mathrm{T}} \qquad (3)$$

Here, each $p_{ik_j}, 1 \le i \le L$, denotes the probability that a region x_{k_j} belongs to the category $c_i \in C$. Based on the probability scores, the category label of x_{k_j} is determined by $m = \arg\max_{1 \le m \le L}[p_{m_k}]$ that is the label of the category c_m with the maximum probability score. Hence, the region x_{k_j} is annotated with the label m and the entire image is thus represented as a two-dimensional index linked to the concept or localized semantic labels assigned for each region. Based on this encoding scheme, an image I_j is represented as a vector in a local semantic concept space as

$$\mathbf{f}_j = [f_{1_j}, \cdots, f_{i_j}, \cdots f_{L_j}]^{\mathrm{T}} \qquad (4)$$

where each f_{i_j} corresponds to the normalized frequency of a concept $c_i, 1 \le i \le L$ in image I_j.

However, this representation scheme captures only a coarse distribution of the concepts and is analogous to the distribution of quantized color in a global color histogram. The image representation based on the hard encoding scheme, i.e., to find only the best concept prototype for each region, is very sensitive to quantization/classification errors and ignores correlations information among concepts. Two regions within an image will be considered different if their corresponding concept labels are predicted as different even though they might be very similar or correlated with each other. In the following section, we propose an effective correlation-enhanced similarity matching scheme to address the above limitation.

3. Correlation-Enhanced Similarity Matching

This section presents the similarity matching approach in the local concept space by considering the correlations between the concepts in the collection based on their similarity and co-occurrence patterns. For the correlation analysis, we construct two different global thesauruses in the form of a similarity and a correlation matrix where each element of the matrices defines concept similarities or co-occurrence relationships respectively. Finally, this global matrices are utilized in a quadratic form of distance measure to compare a query and database images as

$$Dis_{\mathrm{S}}(I_q, I_j) = \sqrt{(\mathbf{f}_q - \mathbf{f}_j)^T \mathbf{A} (\mathbf{f}_q - \mathbf{f}_j)} \qquad (5)$$

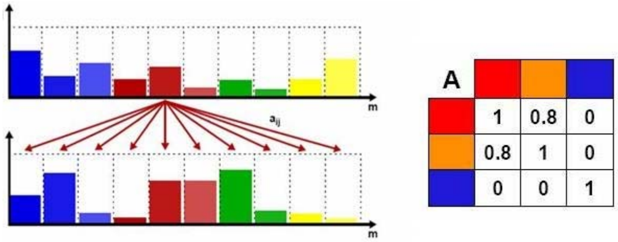

Figure 2. Example of quadratic similarity matching in color histograms

Here, $\mathbf{f}_q$ and $\mathbf{f}_j$ are the feature vector for the query image I_q and a target image I_j respectively and $\mathbf{A}$ is a matrix, where each element of it represents some kind of similarity or co-relationships between the attributes of the feature vectors.

The quadratic distance measure is first implemented in the QBIC [17] system for the color histogram-based matching. It overcomes the shortcomings of the L-norm distance functions by comparing not only the same bins but multiple bins between color histograms. For example, Figure 2 shows the concept of the quadratic distance measure function, where instead of a *one-to-one* matching of the bins between two histograms, each bin (e.g., red here) in one histogram is matched with all other bins in the other histogram. The matching is performed with a weighted similarity score based on the matrix $\mathbf{A}$ on the right side of Figure 2, where each element a_{ij} of the matrix depicts the similarity between two colors i and j in some color space. Due to this property, distance matching scheme performs better compared to the Euclidean and histogram intersection-based distance measures for the color-based image retrieval [17].

Although, a similarity based on only the color feature does not always indicate semantic similarities between the images due to the *semantic gap* problem as mentioned above and does not imply any hidden correlation between feature attributes in a collection. The local concept-based feature representation as described in Section 2 is at an intermediate semantic level due to the incorporation of multi features (e.g., color and texture) in a local region level when compared to the low level pixel-based color feature representation. To measure the similarities between the concepts, we rely on how the concepts in the collection are indexed by images, i.e., for each concept there is image vector space. This idea of measuring similarity was originally proposed in [19] for the textual query expansion in information retrieval with moderate success.

In this approach, each concept $c_i, i \in \{1, \cdots, L\}$ is associated with a vector $\mathbf{c}_i =< w_{i1}, \cdots, w_{ij}, \cdots, w_{iM} >$ where M is the number of images in the collection. The element w_{ij} is the weight for the concept c_i in image I_j, which is computed in a rather distinct form as [19]:

$$w_{ij} = \frac{(\frac{f_{ij}}{max_j\ (f_{ij})})\ ikf_j}{\sqrt{\sum_{l=1}^{M} (\frac{f_{il}}{max_l\ (f_{il})})^2\ ikf_l^2}} \tag{6}$$

where f_{ij} be the frequency of occurrence of the concept c_i in the image I_j and $max_j\ (f_{ij})$ computes the maximum frequency of c_i under all images in the collection. Further, the inverse concept frequency ikf_j for I_j, (e.g., analogous to the inverse image (document) frequency), is computed as $ikf_j = log\frac{L}{k_j}$, where k_j be the number of distinct concepts in the I_j. After generating the concept vectors, a similarity matrix $\mathbf{S}_{L \times L} = [s_{u,v}]$ is built through the computation of each element $s_{u,v}$ as the normalized cosine relationship or dot product between two concept vectors $\mathbf{c}_u$ and $\mathbf{c}_v$ as

$$s_{u,v} = \mathbf{c}_u \cdot \mathbf{c}_v = \sum_{j=1}^{M} w_{uj} * w_{vj} \tag{7}$$

Although, the construction of the matrix $\mathbf{S}$ is prohibitively difficult for large collections. Many collections are available now-a-days, with several hundred thousand images. However, the matrix needs to be computed only once and can be computed off-line. The only component done on a per query basis is the utilizing the matrix elements in the distance matching function.

The similarity matrix $\mathbf{S}$ is created based on how the concepts in the collection are indexed by images and does not take into account the frequency of co-occurrence of pairs of concepts in images. Hence, we propose to use another global matrix that is built by considering the co-occurrence of concepts inside images and in a collection. Let $\mathbf{A}^* = [a_{uv}]$ be a $L \times L$ matrix in which the rows and columns are associated with the concepts in a collection. Each entry a_{uv} expresses a normalized correlation factor between concepts c_u and c_v as

$$a_{uv} = n_{uv}/(n_u + n_v - n_{uv}) \tag{8}$$

where n_u be the number of images in S_l that contain the concept c_u, n_v be the number of images that contain the concept c_v, and n_{uv} be the number of the images in the collection that contain both the concepts.

The entry a_{uv} measures the ratio between the number of images where both c_u and c_v appear and the total number of images in the collection where either c_u or c_v appear and its value ranges to $0 \leq a_{uv} \leq 1$. If c_u and c_v have many co-occurrences in images, the value of a_{uv} increases and the images are considered to be more correlated. This matrix is termed as a *connection matrix* in [18], which is successfully utilized in a fuzzy information retrieval approach. Finally we can easily replaces the above matrices with the distance matching function in (5) and perform the similarity search effectively.

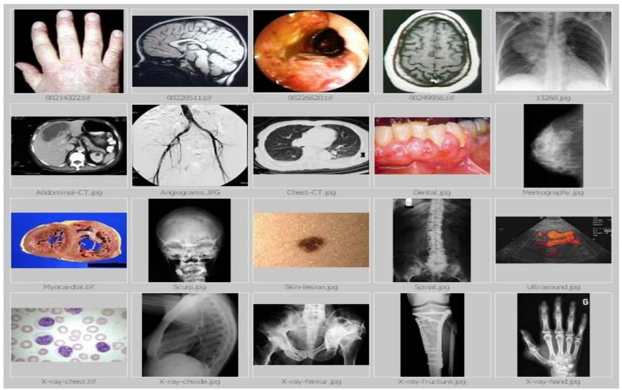

Figure 3. Example images of the Medical collection.

Table 1. Statistics of the Training Set for Local Concepts

Concept	#	Concept	#
CT-Tissue-Brain	400	CT-Tissue-Abdomen	380
CT-Tissue-Lung	350	EC-Tissue-Gastro	300
CT-Bone-Whole	350	CT-Bone-Corner	300
Skin-Normal	420	Skin-Melanoma	370
Teeth-White	330	Red-Tissue-Mouth	310
MRI-Tissue-Leg	320	MRI-Tissue-Brain	350
Xray-Tissue-Lung	400	Xray-Hepatic-Veins	450
Xray-Chest-bone	260	Micro-Infection	200
US-Grey-Texture	445	Doppler-Color	400
Micro-Blood-Blue	490	Histo-Blue	480
Micro-Pink-Bacteria	520	Micro-Tissue-Muscle	500
Photo-Tissue-Brain	520	Photo-Tissue-Cardiac	500
Backgrnd-White	500	Backgrnd-Grey	510
Backgrnd-Blue	350	Backgrnd-Brown	550
Backgrnd-Black	350	Line-Drawing	550

4. Experiments

The image collection for experiment comprises of 5000 bio-medical images of 30 manually assigned disjoint global categories, which is a subset of a larger collection of six different data sets used for medical image retrieval task in ImageCLEFmed 2007 [21]. In our collection, the images are classified into three levels as shown in Figure 4. In the first level, images are categorized according to the imaging modalities (e.g., X-ray, CT, MRI, etc.). At the next level, each of the modalities is further classified according to the examined body parts (e.g., head, chest, etc.) and finally it is further classified by orientation (e.g., frontal, sagittal, etc.) or distinct visual observation (e.g. CT liver images with large blood vessels). The disjoint categories are selected only from the leaf nodes (grey in color) to create the ground-truth data set. The categories are selected based on analyzing the visual and some mixed-mode query topics during the first three years (2005, 2006, and 2007) of ImageCLEFmed retrieval campaign [21]. Few example images of different modalities in the medical collection are shown in Figure 3.

Table 2. CV Accuracy of Local Concept Classification

Kernel	C	γ	Degree	Accuracy (%)
RBF	200	0.5		81.40%
Polynomial	100		1	78.76%
Polynomial	100		2	79.41%

4.1. Training

For the SVM training, we defined 30 local concept categories from the patches, such as X-ray lung tissue and bone, normal and abnormal skin, microscopic images of different color and texture patterns, and so on. Table 1 shows the statistics of each local concept with the number of regions to present them in the training set. The training set used for this purpose consist of only 5% images of all global categories of the entire data set. The local concepts are selected as the ones that exhibit some meanings to the physicians with distinct visual appearances. To generate the local patches, each image in the training set is at first partitioned into an 8×8 grid generating 64 non-overlapping regions. Only the regions that conform to at least 80% of a particular concept category are selected and labeled with the corresponding category label.

For the SVM training, we utilized both the radial basis function (RBF) and the polynomial kernels. There are two tunable parameters while using RBF kernels: C and γ. It is not known beforehand which values of C and γ are the best for the classification problem at hand. Hence, a 10-fold cross-validation (CV) is conducted. Basically pairs of (C, γ) are used and the one with the best CV accuracy is picked. We also experimented with the polynomial kernel of degree 1 and 2 with $C = 100$. However, the best accuracies are achieved by the RBF kernel as shown in Table 2. Hence, after finding the best values of parameters C and γ of the RBF kernels, they are utilized for the final training to generate the SVM model files for the local concepts. We utilized the *LIBSVM* software package [22] for the implementation of the multi-class SVM classifiers.

5. Results

For a quantitative evaluation of the retrieval results, we selected all the images in the collection as query images and used *"query-by-example"* as the search method, where a query is specified by providing an example image to the system. A retrieved image is considered to be a correct match if it is in the same category (based on the ground truth) as the query image. Precision (percentage of retrieved images that are also relevant) and recall (percentage of relevant images that are retrieved) are used as the basic evaluation measure of retrieval performances [20]. The average precision and recall are calculated over all the queries to generate the precision-recall (PR) curves in different settings.

We at first evaluated the effectiveness of the concept-

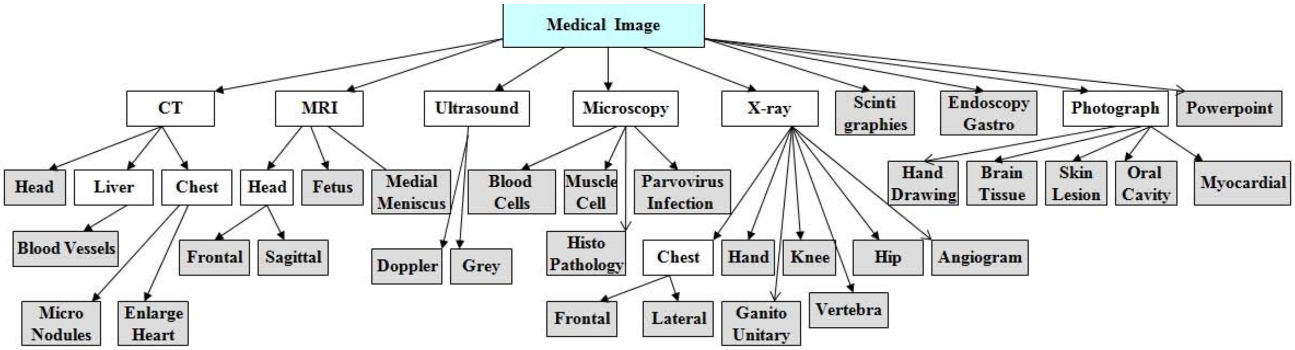

Figure 4. Classification structure of the medical image data set.

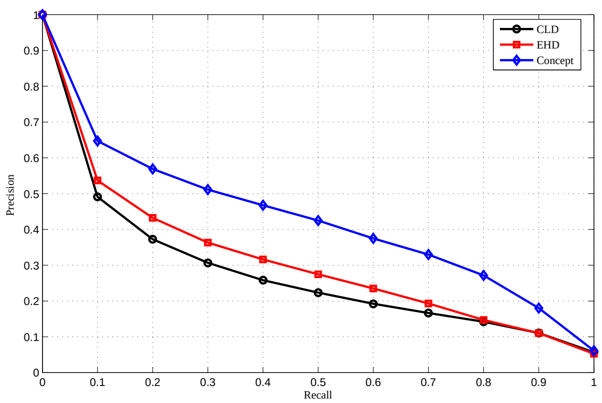

Figure 5. PR graph of different feature representations.

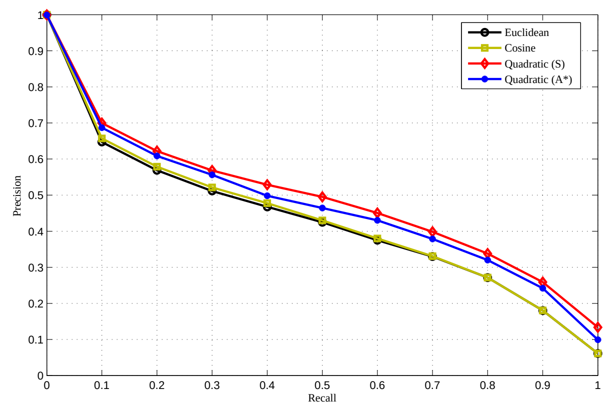

Figure 6. PR graph of different similarity matching schemes.

based feature representation scheme compared to the two well known low-level feature representation schemes, the MPEG-7 based Color Layout Descriptor and Edge Histogram Descriptor (EHD) [23]. The CLD represents the spatial layout of the color images in a very compact form and the EHD represents spatial distribution of edges as a global shape feature. Figure 5 shows the PR curves based on the Euclidean similarity matching in the different feature spaces. By analyzing the Figure 5, we can observe that proposed concept-based feature representation schemes performed much better for both the collections compared to the CLD and EHD based features in terms of precision at each recall level. The better performances are expected as the concept features are more semantically oriented that exploits the domain knowledge of the collections at a local level.

Figure 6 shows the effectiveness of the quadratic similarity matching as described in Section 3. It is compared with the Euclidean and Cosine similarity matching in the same concept-based feature space. From Figure 6, we can observe that, the quadratic similarity matching based on utilizing matrices $\mathbf{A}^*$ and $\mathbf{S}$ (e.g., Quadratic ($\mathbf{A}^*$) and Quadratic ($\mathbf{S}$)) performed better when compared to the Euclidean and Cosine similarity matching. The performances of the Euclidean and Cosine similarity matching are almost identical, whereas the best precisions in terms of all the recall levels are obtained when search is performed with a the Quadratic similarity matching by utilizing the global matrix $\mathbf{S}$. Overall, the improved results of the Quadratic similarity matching indicate that the correlations among the concepts are not negligible and can be exploited effectively in the similarity matching function.

6. Conclusions

We have investigated correlation-enhanced similarity matching in a local concept-based feature space in CBIR domain. The proposed technique exploit the similarities and correlations between the local concepts based on the global analysis at the collection level. Due to the nature of the image representation schemes in the concept-based feature space, there always exists enough correlations between the concepts. Hence, exploiting this property might improve the retrieval effectiveness. Experimental results validated the assumption and showed that the proposed similarity matching schemes improved retrieval accuracies when compared to the Euclidean and Cosine similarity matching in the same concept-based feature space.

Acknowledgment

This research is supported by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine (NLM), and Lister Hill National Center for Biomedical Communications (LHNCBC). We would like to thank the CLEF [21] organizers for making the database available for the experiments.

References

[1] T. C. Wong, Medical Image Databases. New York, LLC: Springer Verlag; 1998.

[2] H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler, A Review of Content-Based Image Retrieval Systems in Medical Applications Clinical Benefits and Future Directions. *International Journal of Medical Informatics*, 73:1–23, 2004.

[3] R. H. Choplin, J. M. Boehme, and C. D. Maynard, Picture archiving and communication systems: an overview. *RadioGraphics*, 12:127–129, 1992.

[4] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, Content-Based Image Retrieval at the End of the Early Years. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 22(12):1349–1380, 2000.

[5] C. R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. M Aisen, and L. S. Broderick, ASSERT: a physician-in-the-loop content-based image retrieval system for HRCT image databases. *Comput Vis Image Understanding*, 75:111–132, 1999.

[6] W. Hsu, S. Antani, L. R. Long, L. Neve, and G.R. Thoma, SPIRS: a Web-based Image Retrieval System For Large Biomedical Databases. *International Journal of Medical Informatics*, 78:13–24, 2008.

[7] T. M. Lehmann, B.B. Wein, J. Dahmen, J. Bredno, F. Vogelsang, and M. Kohnen, Content–based image retrieval in medical applications-A novel multi-step approach. *Proc SPIE*, 3972:312–320, 2000.

[8] H. Müller, A. Rosset, J. Vallee, and A. Geissbuhler, Integrating content-based visual access methods into a medical case database. *Proc Med Inform Europe (MIE 2003)*, St Malo, France, 480–485, 2003.

[9] Y. Liua, D. Zhang, G. Lu, and W. Y. Ma, A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 40:62–282, 2007.

[10] J. Vogel, and B. Schiele, Semantic Modeling of Natural Scenes for Content-Based Image Retrieval. *International Journal of Computer Vision*, 72(2):133–157, 2007.

[11] R. Shi, H. Feng, T. S. Chua, and C. H. Lee, An adaptive image content representation and segmentation approach to automatic image annotation, *International Conference on Image and Video Retrieval (CIVR)*, 545-554, 2004.

[12] M. M. Rahman, S. K. Antani, and G. R. Thoma, A Medical Image Retrieval Framework in Correlation Enhanced Visual Concept Feature Space. *22nd IEEE International Symposium on Computer-Based Medical Systems (CBMS)*, Albuquerque, New Mexico, USA, August 3-4, 2009.

[13] J. C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, vol. 2:121-167, 1998.

[14] K. Duan and S. S. Keerthi, Which is the best multiclass SVM method? An empirical study. *Sixth International Workshop on Multiple Classifier Systems*, In Proceedings of the LNCS. 3541, 2005.

[15] R. M. Haralick, Shanmugam, and I. Dinstein, Textural features for image classification. *IEEE Trans. Syst. Man Cybernetics*, 3:610-21, 1973.

[16] T. F. Wu, C. J. Lin, and R. C. Weng, Probability Estimates for Multi-class Classification by Pairwise Coupling. *Journal of Machine Learn Ressearch*, 5:975–1005, 2004.

[17] J. Hafner, H. S. Sawhney, W. Equitz, M. Flickner, and W. Niblack, Efficient color histogram indexing for quadratic form distance functions. *IEEE Trans. Pattern Anal. Machine Intell.*, 17(7):729-736, 1995.

[18] O. Yasushi, M. Tetsuya, and K. Kiyohiko, A fuzzy document retrieval system using the keyword connection matrix and a learning method. *Fuzzy Sets and Systems*, 39(2): 163–179, 1991.

[19] Y. Qiu, and H. P. Frei, Concept Based Query Expansion. *Proc. 16th Int. ACM SIGIR Conf. on R&D in Info. Retrieval*, SIGIR Forum, ACM Press, June, 1993.

[20] R. B. Yates, and B. R. Neto, Modern Information Retrieval. 1st ed., Addison Wesley, 1999.

[21] H. Müller, T. Deselaers, E. Kim, C. Kalpathy, D. Jayashree, M. Thomas, P. Clough, and W. Hersh, Overview of the ImageCLEFmed 2007 Medical Retrieval and Annotation Tasks. *8th Workshop of the*

Cross-Language Evaluation Forum (CLEF 2007), Proceedings of LNCS, 5152, 2008.

[22] C. C. Chang, and C. J. Lin, LIBSVM : a library for support vector machines. 2001, Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm

[23] S. F. Chang, T. Sikora, and A. Puri, Overview of the MPEG-7 standard. *IEEE Trans Circ Syst Video Technology*, 11:688–695, 2001.