Medical Image Retrieval with Query-Dependent Feature Fusion based on One-Class SVM

Yonggang Huang[1,2],Jun Zhang[3],Yongwang Zhao[1,2],Dianfu Ma[1,2]
[1]Institute of Advanced Computing Technology,BeiHang University
[2]National Lab of Software Development Environment,BeiHang University
[3]School of Computer Science and Software Engineering,University of Wollongong
yonggang.h@gmail.com,zhaoyw@act.buaa.edu.cn

***Abstract*—Due to the huge growth of the World Wide Web, medical images are now available in large numbers in online repositories, and there exists the need to retrieval the images through automatically extracting visual information of the medical images, which is commonly known as content-based image retrieval (CBIR). Since each feature extracted from images just characterizes certain aspect of image content, multiple features are necessarily employed to improve the retrieval performance. Meanwhile, experiments demonstrate that a special feature is not equally important for different image queries. Most of existed feature fusion methods for image retrieval only utilize query independent feature fusion or rely on explicit user weighting. In this paper, we present a novel query dependent feature fusion method for medical image retrieval based on one class support vector machine. Having considered that a special feature is not equally important for different image queries, the proposed query dependent feature fusion method can learn different feature fusion models for different image queries only based on multiply image samples provided by the user, and the learned feature fusion models can reflect the different importances of a special feature for different image queries. The experimental results on the IRMA medical image collection demonstrate that the proposed method can improve the retrieval performance effectively and can outperform existed feature fusion methods for image retrieval.**

I. Introduction

In the medical field, with the rapid advances in imaging technology, medical image are produced in ever-increasing quantities by hospitals, pharmaceutical companies, and academic medical research centre. These images of various modalities are playing an important role in detecting the anatomical and functional information about different body parts for the diagnosis, medical research, and education. Currently, many hospitals and radiology departments are equipped with Picture Archiving and Communications Systems (PACS) [1]. The images are commonly stored, retrieved and transmitted in the DICOM (Digital Imaging and Communication in Medicine) [2] format. The search for images is carried out according to the textual attributes of image headers (such as study, patient) and usually has many limitations.

Due to the huge growth of the World Wide Web, medical images are now available in large numbers in online repositories, atlases, and other heath related resources [3]. In such a web-based environment, medical images are generally stored and accessed in common formats such as JPEG (Joint Photographic Experts Group), GIF (Graphics Interchange Format), etc. These formats are used because they are easy to store and transmit compared to the large size of images in DICOM format [4], but also for anonymization purposes [3]. However, there is no header information attached to the images with these image formats other than DICOM format . In this case, the text-based approach is both expensive and ambiguous due to the fact that manually annotating these images are extremely time-consuming, highly subjective and requires domain-related knowledge. The content-based image retrieval (CBIR) [5] systems overcome these limitations since they are capable of carrying out a search for images based on the modality, anatomic region and different acquisition views [3] through automatically extracting visual information of the medical images. Currently, there exist some CBIR systems on medical image such as MedGIFT [3], COBRA [6], IRMA [1] and KmED [7].

The CBIR extract the low level visual features such as color, texture, or spatial location automatically and the images are retrieved based on the these low level visual features. Experiments [8] demonstrate that the image retrieval performance can be enhanced when employing multiple features, since each feature extracted from images just characterizes certain aspect of image content and multiple features can provide an adequate description of image content. Further experiments [9] [10] also show that a special feature is not equally important for different image queries since a special feature has different importances in reflecting the content of different images. Although some research efforts have been reported to enhance the image retrieval performance taking the feature fusion approaches, most of existed feature fusion methods for image retrieval only utilize query independent feature fusion which usually apply a single feature fusion model for all the image queries and do not consider that a special feature is not equally important for different image queries, the others usually require the users to tune appropriate parameters for the feature fusion models for different image queries.

Motivated by this observation, in this paper, with multiply image examples provided by the user, we propose a new query dependent feature fusion method for medical image retrieval based one-class support vector machines. The query dependent feature fusion problem was formulated as a one class classification problem in our work and we solved it

978-0-7695-4323-9/10 $26.00 © 2010 IEEE
DOI 10.1109/CSE.2010.30

with one-class support vector machines because of its good generalization ability. Having considered that a special feature is not equally important for different image queries, the proposed query dependent feature fusion method for medical image retrieval can learn different feature fusion models for different image queries only based on multiply image samples provided by the user, and the learned feature fusion models can reflect the different importances of a special feature for different queries .

The remaining of the paper is organized as follows.Next section discusses some related work. In Section III , we give the formal definition of the query dependent feature fusion problem as one class classification problem. In Section IV, the one class support vector machine based query dependent feature fusion (*OSVM-QDFF*) approach is presented to solve the specific one class classification problem defined in Section III. Section IV discusses the low-level feature extraction processes for the medical image retrieval. The comparison experiments and the analysis of the results are presented in section V, and finally section VI provides our conclusion.

II. RELATED WORK

Since the image retrieval performance can be enhanced when employing multiple features, some research efforts have been reported to enhance the image retrieval performances taking the feature fusion approaches.

There are two main approaches to address the feature fusion problem for image retrieval [11]. One is called as *early fusion*, which perform the feature fusion by stacking the descriptor values as a single, large vector and the images is ranked by calculating the distances between these large vectors in a high dimensional feature space [12]. The *early fusion* approach usually suffers from the dimensionality arising [11]. The other is called as *late fusion*, which obtains image similarity through combining multiple feature similarities.Compared to the *early fusion* approach, the *late fusion* approach alleviates the dimensionality arising and different similarity measures can be used for different features [13]. Since these merits of the *late fusion* approach, the recently research works usually adopt *late fusion* approach for image retrieval.

In [14], the CombSumScore, CombMaxScore, CombSumRank,CombMaxRank fusion models are used to fuse the multiple similarities obtained with multi-feature multi-example queries, which treat different features equally for all the queries and can be called as average fusion models. Obviously, the average fusion models are not optimal as different features usually have different retrieval performances. In literate [15], the genetic algorithm is used to learn the best weights for different features, and then the learned feature fusion model is applied for all the image queries. In literate [16], different features are assigned with different weights according to the average retrieval precision of these features, and then the adjusted feature fusion model is applied for all the image queries. The feature fusion methods presented in [15] and [16] can enhance the retrieval performance to some extent as the different retrieval performances of different features are considered. However, firstly, a certain amount of training data in needed in [15] and [16], secondly, the learned fusion models are not optimal for each image query as a special feature is not equally importance for different image queries. In summary, all these feature fusion methods for image retrieval apply a single feature fusion model for all the image queries and do not consider that a special feature is not equally important for different image queries.

In [17] and [18], the combined similarity between images is measured using one of the features selected by a feature fusion model expressed with logic operation based on Boolean model. To overcome the limitation of traditional Boolean model, [12] introduced a hierarchical decision fusion framework formulated based on fuzzy logic to extend AND and OR operations in Boolean logic. In [17] [18] [12], the feature fusion models for different image queries are presented with logic-based expressions and usually require the users to tune appropriate parameters for the fusion models, which could only be successful in specific field(for example, the art image [12]) since they require the user having a good understanding of the low level feature of the query images.

In literate [10], the author proposed a query dependent feature fusion method for image retrieval (which is called as local aggregation function in [10]) based on support vector machine (*LSVMC*). Regarding the multiply image examples provided by the user as positive examples and the randomly selected image examples from the image collection as negative examples, the author in [10] formulate the query dependent feature fusion problem as a strict two class classification problem and solved it by support vector machines, with equal treatments on both positive and negative examples. However, the strict two class classification based approach is not always reasonable since the negative examples randomly selected from the image collection can belong to any class and they usually do not cluster.

III. PROBLEM DEFINITION FOR THE QUERY-DEPENDENT FEATURE FUSION

In this section, we investigate the query dependent feature fusion problem for the query by example search paradigm when the user provides multiply example images as a query, and we formulate the query dependent feature fusion problem as a one class classification problem.

Let us consider an medical image collection $\Omega = \{I_1, \cdots, I_i, \cdots, I_N\}$ which contains N images that we are interested in retrieval.Suppose m low level feature descriptors are available $F = \{f_1, \cdots, f_i, \cdots, f_m\}$. The low level feature representation for image I with the feature descriptors set F can be denotes as

$$F^I = \{f_1^I, \cdots, f_i^I, \cdots, f_m^I\} \tag{1}$$

where f_i^I denotes the feature vector for image I using the feature descriptor f_i, and F^I denotes the feature vectors set for image I.

Let $D_i(.,.)$ denotes the distance metric for the i_{th} feature descriptor f_i, thus the distance between image I and image J

when using the i_{th} feature descriptor f_i can be represented as

$$d_i(I,J) = D_i(f_i^I, f_i^J) \tag{2}$$

Suppose the user provides multiply image examples as a query $Q = \{Q_1, \cdots, Q_i, \cdots, Q_q\}$. The combined image collection of the query and the image collection that we are interested in retrieval can be represented as

$$\Omega' = Q \cup \Omega \tag{3}$$

Given a image example Q_i in the query Q, the distances to each image in the image collection Ω' using the feature descriptor f_i can be represented as.

$$D_j(Q_i) = \{d_j(Q_i,Q_1), \cdots, d_j(Q_i,Q_q), d_j(Q_i,I_1), \cdots, d_j(Q_i,I_N)\} \tag{4}$$

where $D_j(Q_i)$ denotes the distances set for example image Q_i on image set Ω' with the feature descriptor f_j. In order to make the distances obtained with different feature descriptor be comparable, the distances with feature descriptor f_j are normalized as

$$\bar{d_j} = \frac{d_j - d_j^{min}}{d_j^{max} - d_j^{min}} \tag{5}$$

where d_j^{max} and d_j^{min} denotes the maximum and minimum distance in the distances set $D_j(Q_i)$. The normalized distances can be converted to the similarity as

$$s_j = 1 - \bar{d_j} \tag{6}$$

The similarities between the image example Q_i and the image I_i' in image collection Ω' with m different feature descriptors can be represent as a similarities vector

$$S(Q_i, I_j') = (s_1(Q_i, I_j'), \cdots, s_m(Q_i, I_j')) \tag{7}$$

and the similarities between the example image Q_i and all the images in image collection Ω' can be represented as a similarity space $\varphi(Q_i)$ with the size of $(N+q) \times m$

$$\begin{bmatrix} s_1(Q_i,I_1') & \cdots & s_i(Q_i,I_1') & \cdots & s_m(Q_i,I_1') \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s_1(Q_i,I_j') & \cdots & s_i(Q_i,I_j') & \cdots & s_m(Q_i,I_j') \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s_1(Q_i,I_{N+q}') & \cdots & s_i(Q_i,I_{N+q}') & \cdots & s_m(Q_i,I_{N+q}') \end{bmatrix} \tag{8}$$

By considering a linear fusion solution , the combined similarity between the image example Q_i and the image I_j' in Ω' can be represented as.

$$Sim(Q_i, I_j') = S(Q_i, I_j') \cdot \mathbf{w}^T \tag{9}$$

where $\mathbf{w} = (w_1, \cdots, w_m)$ is feature weight vector and w_i denotes the weight assigned for feature f_i which reflect the feature importance for the query with the query set Q.

Suppose that the relevant image set for the query Q is Θ. Thus the optimal *Query-Dependent* feature fusion for the example image Q_i is to find appropriate feature weight vector $\mathbf{w} = (w_1, \cdots, w_m)$ that can separate the relevant image set Θ from the image collection Ω' in the similarity space $\varphi(Q_i)$ as

$$\begin{cases} Sim(Q_i, I_j') = \mathbf{w} \cdot S(Q_i, I_j') > \rho & if\ I_j' \in \Theta \\ Sim(Q_i, I_j') = \mathbf{w} \cdot S(Q_i, I_j') < \rho & if\ I_j' \notin \Theta \end{cases} \tag{10}$$

where ρ is the similarity threshold to separate the relevant image set Θ from the image collection Ω'.

Since each image example Q_i in query Q expresses the users' retrieval purpose equally and the feature fusion model for all the image examples Q_i in query Q should be the same (which is also the feature fusion model for the query Q). Therefore the optimal *Query-Dependent* feature fusion for the query Q is to find appropriate feature weight vector $\mathbf{w} = (w_1, \cdots, w_m)$ that can separate the relevant image set Θ from the image collection Ω' in the similarity spaces $\varphi(Q_1), \cdots, \varphi(Q_q)$ as

$$\begin{cases} Sim(Q_i, I_j') = \mathbf{w} \cdot S(Q_i, I_j') > \rho & if\ I_j' \in \Theta & i = 1, 2 \cdots, q \\ Sim(Q_i, I_j') = \mathbf{w} \cdot S(Q_i, I_j') < \rho & if\ I_j' \notin \Theta & i = 1, 2 \cdots, q \end{cases} \tag{11}$$

which is equally to find appropriate feature weight vector $\mathbf{w} = (w_1, \cdots, w_m)$ that can separate the relevant image set Θ from the image collection Ω' in the combined similarity space φ as

$$\begin{bmatrix} s_1(Q_1,I_1') & \cdots & s_i(Q_1,I_1') & \cdots & s_m(Q_1,I_1') \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s_1(Q_1,I_j') & \cdots & s_i(Q_1,I_j') & \cdots & s_m(Q_1,I_j') \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s_1(Q_1,I_{N+q}') & \cdots & s_i(Q_1,I_{N+q}') & \cdots & s_m(Q_1,I_{N+q}') \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s_1(Q_q,I_1') & \cdots & s_i(Q_q,I_1') & \cdots & s_m(Q_q,I_1') \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s_1(Q_q,I_j') & \cdots & s_i(Q_q,I_j') & \cdots & s_m(Q_q,I_j') \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s_1(Q_q,I_{N+q}') & \cdots & s_i(Q_q,I_{N+q}') & \cdots & s_m(Q_q,I_{N+q}') \end{bmatrix} \tag{12}$$

Notice that each image in image collection Ω' is represented with q similarities vectors, each of which represents the similarities to one example image in Q with m different feature descriptors. The combined similarity space φ can be obtained by simply combing the similarities spaces

$$\varphi = \varphi(Q_1) \cup \cdots \cup \varphi(Q_q) \tag{13}$$

Consider that the size of Θ is much smaller compared to the size of the image collection Ω' as

$$|\Theta| \ll \left|\Omega'\right| \tag{14}$$

Thus the query dependent feature fusion problem for the query Q can be regards as a typical one class classification problem in the combined similarity space φ with the training data as

$$\left\{(Sim(Q_i, I_j^{'}), L_{ij})|1 \leq j \leq (N+q), 1 \leq i \leq q\right\} and$$
$$L_{ij} = \begin{cases} 1 & I_j^{'} \in Q \\ 0 & I_j^{'} \notin Q \end{cases} \tag{15}$$

where 1 indicates a positive sample and 0 indicates a unlabeled sample, since the example images in the query Q are relevant to the query Q.

Treating the example image in the query Q equally, the similarities between the image $I_i^{'}$ in $\Omega^{'}$ and the query Q can be computed as

$$S(Q, I_j^{'}) = \sum_{i=1}^{q} S(Q_i, I_j^{'}) \tag{16}$$

and the combined similarities between the query Q and the images in $\Omega^{'}$ can be obtained as

$$Sim(Q, \Omega^{'}) = \begin{bmatrix} Sim(Q, I_1^{'}) \\ \cdots \\ Sim(Q, I_j^{'}) \\ \cdots \\ Sim(Q, I_{N+q}^{'}) \end{bmatrix} = \sum_{i=1}^{q} \varphi\left(Q_i\right) \cdot \mathbf{w}^T \tag{17}$$

In summary, the query dependent feature fusion problem for the query Q is to find appropriate feature weight vector $\mathbf{w} = (w_1, \cdots, w_m)$ for formulation 17 through solving the one class classification problem defined in formulation 11 and 15.

IV. ONE CLASS SUPPORT VECTOR MACHINE BASED *Query-Dependent* FEATURE FUSION (*OSVM-QDFF*)

In section III, the query dependent feature fusion problem has been formulated as a one class classification problem in the combined similarity space φ. In this section, the One-class support vector machine (One-Class SVM) [19] is selected to solve the specific one class classification problem because of the good generalization ability. The algorithm is named One-class SVM since only positive examples are used in training and testing.

Considering a linear one classification problem in the combined similarity space φ with the positive examples φ^+

$$\varphi^+ = \{s_1, s_2, \cdots, s_l\} \subset \varphi \tag{18}$$

where $l = p * p$ indicates the number of the positive examples in the combined similarity space. The goal of training of a linear One-Class SVM is find a separating hyperplane in the combined similarity space

$$f(s) = \mathbf{w} \cdot s \ - \rho \tag{19}$$

where $\mathbf{w}$ is the adaptive feature weight vector in this paper. The separating hyperplane stratifies that it is closer to the origin than all the examples in φ^+ as

$$f(s_i) > 0, \ i = 1, 2, \cdots, t \tag{20}$$

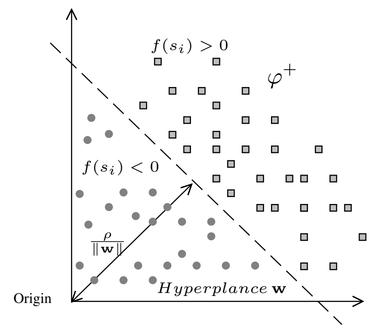

Fig. 1: The goal of training of a linear One-Class SVM is to find a separating hyperplane in the combined similarity space with the largest margin

and with the largest margin to the origin in such hyperplanes as Fig.1 presents.

$$max \ \frac{\rho}{\|\mathbf{w}\|} \tag{21}$$

By properly chosen nonlinear function ϕ, the combined similarities space can be mapped to a high dimensional feature space F to get a potentially better representation of the data point and achieve a better classification as

$$\phi: \quad \varphi \rightarrow F \tag{22}$$

and the output of the nonlinear One-Class SVM is a separating hyperplane in the high dimensional feature space F with the largest margin to the origin $\frac{\rho}{\|\mathbf{w}\|}$ and satisfy $f(s_i) > 0$ for all the positive examples s_i in φ^+ as

$$f(s) = \mathbf{w} \cdot \phi(s) - \rho \tag{23}$$

The linear One-Class SVM can be regards as a typical non-linear One-Class SVM with the mapping function

$$\phi(s) = s \tag{24}$$

With the training data $\phi(s_1), \phi(s_2), \cdots, \phi(s_l)$, the optimal hyperplane $\mathbf{w}$ can be found by solving the following quadratic programming problem [19]

$$\begin{cases} min & \frac{1}{2}\|\mathbf{w}\|^2 - \rho \\ s.t & \mathbf{w} \cdot \phi(s_i) \geq \rho \quad i = 1, 2, \cdots, t \end{cases} \tag{25}$$

Considering that the sample points in F are not always linearly separable and it is too difficult to find a canonical hyperplane quickly in this case. There may be no hyperplane that separate φ^+ from φ in F. Therefore, the slack parameters, denoted by $\xi_i \geq 0$, is associated with each training samples. It allows for some training samples to be within the margin. The optimization is to find maximize margin and at the same time to minimize the average slack.

$$\begin{cases} min & \frac{1}{2}\|\mathbf{w}\|^2 - \rho + \frac{1}{\nu l}\sum_{i=1}^{l} \xi_i \\ s.t & \mathbf{w} \cdot s_i \geq \rho - \xi_i, \ \xi_i \geq 0, \ i = 1, 2, \cdots, t \end{cases} \tag{26}$$

where ξ_i are slack variables, l is the number of training samples, and $\nu \in (0, 1]$ is a parameter that controls the trade-off between maximizing the distance from the origin and separating most of the relevant samples. After introducing Lagrange multipliers α_i for each training samples , the dual problem of the optimization problem can be obtained as

$$\begin{cases} max & \frac{1}{2}\sum_{i,j=1}^{l}\alpha_i\alpha_j\phi(s_i)\cdot\phi(s_j) \\ s.t. & 0\le\alpha_i\le\frac{1}{\nu l} \\ & \sum_{i=1}^{l}\alpha_i=1 \end{cases} \tag{27}$$

Solving the dual problem leads to

$$\mathbf{w}=\sum_{i=1}^{l}\alpha_i\phi(s_i), 0\le\alpha_i\le\frac{1}{\nu l} \tag{28}$$

and the corresponding decision function becomes

$$f(s)=\sum_{i=1}^{l}\alpha_i\phi(s_i)\cdot\phi(s)-\rho \tag{29}$$

with the kernel function $K(s_i,s_j)=\phi(s_i)\cdot\phi(s_j)$ the decision function can be rewritten as

$$f(s)=\sum_{i=1}^{l}\alpha_i K(s_i,s)-\rho \tag{30}$$

Since the combined similarity between the image example Q_i and the image $I_j^{'}$ in $\Omega^{'}$ is obtained as

$$Sim(Q_i,I_j^{'})=S(Q_i,I_j^{'})\cdot\mathbf{w}^T \tag{31}$$

Thus the combined similarity between the image example Q_i and the image I_j in Ω with the decision function in the combined similarity space φ can be represented as

$$Sim(Q_i,I_j)=f(S(Q_i,I_j))+\rho \tag{32}$$

In order to obtained the combined similarity between the query Q and the image I_j in Ω, the gauss normalization is firstly used to make the similarities obtained with with different example image Q_i be comparable as

$$Sim^{'}(Q_i,I_j)=\frac{Sim(Q_i,I_j)-\mu}{3\sigma+1} \tag{33}$$

where μ and σ are the average value and the standard deviation of the similarities obtained with example image Q_i as

$$\mu=\frac{\sum_{j=1}^{N}Sim(Q_i,I_j)}{N} \tag{34}$$

$$\sigma=\sqrt{\frac{\sum_{j=1}^{N}(Sim(Q_i,I_j)-\mu)^2}{N-1}} \tag{35}$$

The final similarity between the query Q and the image I_j in Ω is obtained as the sum of the normalized similarities with convert using the exponential function

$$Sim^{'}(Q,I_j)=\sum_{i=1}^{q}\exp(Sim^{'}(Q_i,I_j)) \tag{36}$$

V. LOW-LEVEL MEDICAL IMAGE FEATURE REPRESENTATION

In order to efficiently retrieve images relevant to a query, a CBIR system usually extracts low-level image features to represent an image in an off-line preprocessing stage. Image features can be categorized into color, shape, texture and spatial relationships. In this paper, we extract the low-level feature representation for medical image retrieval as follows:

Color Feature: we utilize the Color Layout Descriptor (CLD) [20] to represent spatial color distribution within the medical image.Although CLD is created for color images, it equally suitable for gray-level images with proper choice of coefficients [4]. It is obtained by applying the discrete cosine transformation (DCT) on the 2-D array of local representative colors in the $YCbCr$ color space where Y is the luma component and Cb and Cr are the blue and red chroma components. Each channel is represented by 8 bits and each of the 3 channels is averaged separately for the 8×8 image blocks. In our work, a CLD with 64 Y , 3 Cb, and 3 Cr, is extracted to form 70-dimensional feature vector.The distance between two CLD vectors is calculated as :

$$D_{cld}(Q,I)=\sqrt{\sum_i(Y_{Q_i}-Y_{I_i})^2}+\sqrt{\sum_i(Cb_{Q_i}-Cb_{I_i})^2}+\sqrt{\sum_i(Cr_{Q_i}-Cr_{I_i})^2} \tag{37}$$

Texture Feature: In [21], *Tamura* propose six texture features corresponding to human visual perception: *coarseness, contrast, directionality, line-likeness, regularity, and roughness*.The first three features are very important from experiments testing, thus in this paper 1 *coarseness*, 1 *contrast* and 16 *directionality* from 16 directions is extracted to form 18-dimensional feature vector in order to represent the texture feature of medical images.The distance between two *Tamura* feature vector is calculated as:

$$D_{tamura}(Q,I)=\sqrt{\sum_i(T_{Q_i}-T_{I_i})^2} \tag{38}$$

Edge Feature: The Edge Histogram Descriptor (EHD) [20] is used to represented the global edge feature in this paper. The EHD represents local edge distribution in an image by dividing the image into 4×4 sub-images and generating a histogram from the edges present in each of these sub-images. Edges in the image are categorized into five types, namely vertical, horizontal, 45° diagonal, 135° diagonal and non-directional edges. Finally, a histogram with $16\times5=80$ bins is obtained, corresponding to a 80-dimensional feature vector.The distance between two EHD vectors is calculated as shown below:

$$D_{ehd}=\sum_i|H_{Q_i}-H_{I_i}| \tag{39}$$

VI. EXPERIMENTS AND RESULT

To evaluate the effectiveness of the proposed One-Class SVM based query dependent feature fusion method for medical image retrieval, exhaustive experiments were performed

Examples	*CombSum Score*	*CombMax Score*	*CombSum Rank*	*CombMax Rank*	*Linear LSVMC*	*Polynomial LSVMC*	*Sigmoid LSVMC*	*Linear OSVM-QDFF*	*Polynomial OSVM-QDFF*	*Sigmoid OSVM-QDFF*
4 Examples	0.4128	0.2279	0.2806	0.3518	0.4298	0.4448	0.4296	0.4619	0.4662	0.4688
6 Examples	0.4244	0.263	0.2857	0.4048	0.4509	0.4763	0.4507	0.5199	0.5342	0.5254
8 Examples	0.4351	0.2868	0.2969	0.4291	0.4721	0.5024	0.4719	0.564	0.5801	0.5685

TABLE III: Mean average precision of different feature fusion methods

Examples	*Linear OSVM-QDFF* **vs** *Best Average Fusion Model*	*Linear OSVM-QDFF* **vs** *Best LSVMC*	*Polynomial OSVM-QDFF* **vs** *Best Average Fusion Model*	*Polynomial OSVM-QDFF* **vs** *Best LSVMC*	*Sigmoid OSVM-QDFF* **vs** *Best Average Fusion Model*	*Sigmoid OSVM-QDFF* **vs** *Best LSVMC*
4 Examples	11.8944	3.8444	12.936	4.8112	**13.5659**	**5.3957**
6 Examples	22.5024	9.1539	25.8718	12.1562	23.7983	10.3086
8 Examples	25.4169	12.2611	**28.9971**	**15.4658**	26.4176	13.1568

TABLE IV: Relative Improvement [%] of *OSVM-QDFF* to Best *Average Fusion Model* and Best *LSVMC*

Kernel	C	γ	r	d
Linear	1	-	-	-
Polynomial	256	0.0625	0	3
Sigmoid	32	0.0313	0	-

TABLE I: Final parameter set for SVMs

Kernel	ν	γ	r	d
Linear	0.05	-	-	-
Polynomial	0.0625	2	0	3
Sigmoid	0.5	0.0313	0	-

TABLE II: Final parameter set for One-Class SVMs

on the IRMA medical image collection . The IRMA medical image collection contains 9000 radio graphs taken randomly from medical routine at the RWTH Aachen University Hospital which are subdivided into 57 classes [22]. The images in the collection are in grey level and in PNG (Portable Network Graphics) format. All the images are classified manually by reference coding with respect to a mono-hierarchical coding scheme [22] which describe the imaging modality , the body orientation, the body region examined and the biological system examined.The images have a high intra-class variability and inter-class similarity, which make the retrieval task much difficult [4]. To evaluate the content based medical image retrieval, the query which contains a small number of example images was randomly selected from each class and the remained images in that class are regarded as the corresponding ground truth set for the query. In this paper, the precision P, the recall R, the average precision AP and the mean average precision MAP proposed in [23] are used to measured the retrieval performance for medical image retrieval.

1) Retrieval experiments: To evaluate the performance of the proposed One-Class SVM based query dependent feature fusion method, the query independent feature fusion methods– the average fusion models (including CombSumScore, CombMaxScore, CombSumRank, CombMaxRank) presented in literate [14] and the query dependent feature fusion method–the local aggregation function based on support vector machines presented in literate [10] are implemented as references. Three sets of experiments are performed with the number of examples image in the query varying from $4, 6, 8$. For each set of experiments, 4 queries with the corresponding number of example images were generated randomly for each class, which resulting $57 * 4 = 228$ queries and their corresponding ground truth sets.

Since One-Class SVM and SVM have a lot of parameters to be set such as regularization parameter, kernel parameters. In order to produce robust retrieval results, it is very important to set these parameters. In this paper, we conducted the experiments with three different kernels such as linear, polynomial, sigmoid for both One Class SVM and SVM as follows:

-the linear machines with kernel function

$$K(s_i, s_i) = s_i^T s_j \tag{40}$$

-the polynomial machines with kernel function

$$K(s_i, s_i) = (\gamma s_i^T s_j + r)^d, \gamma > 0 \tag{41}$$

where d is the degree of the polynomial kernel

-the sigmoid machines with kernel function

$$K(s_i, s_i) = \tan(\gamma s_i^T s_j + r) \tag{42}$$

For the linear kernel machines, there are no parameters to set. For the nonlinear machines including the polynomial and the sigmoid, there are additional parameters such as γ, r and d should be set appropriately. For the kernel parameters r and d of both polynomial and sigmoid, we used the standard values. In order to effectively to decide the regularization parameter and the kernel parameter γ for the polynomial and sigmoid, we apply grid search for optimal parameter set that produces the best retrieval performance. The retrieval performance is measured by the mean average precision of 57 queries, with 1 query of 6 example images were generated randomly for each class. Table I provides the results of final parameters for SVMs with different three kernels. Table II provides the results of final parameters for one-class SVMs with different three kernels.

Additionally, the SVM and One-Class SVM with radial basis kernel function are also experimented in our work,

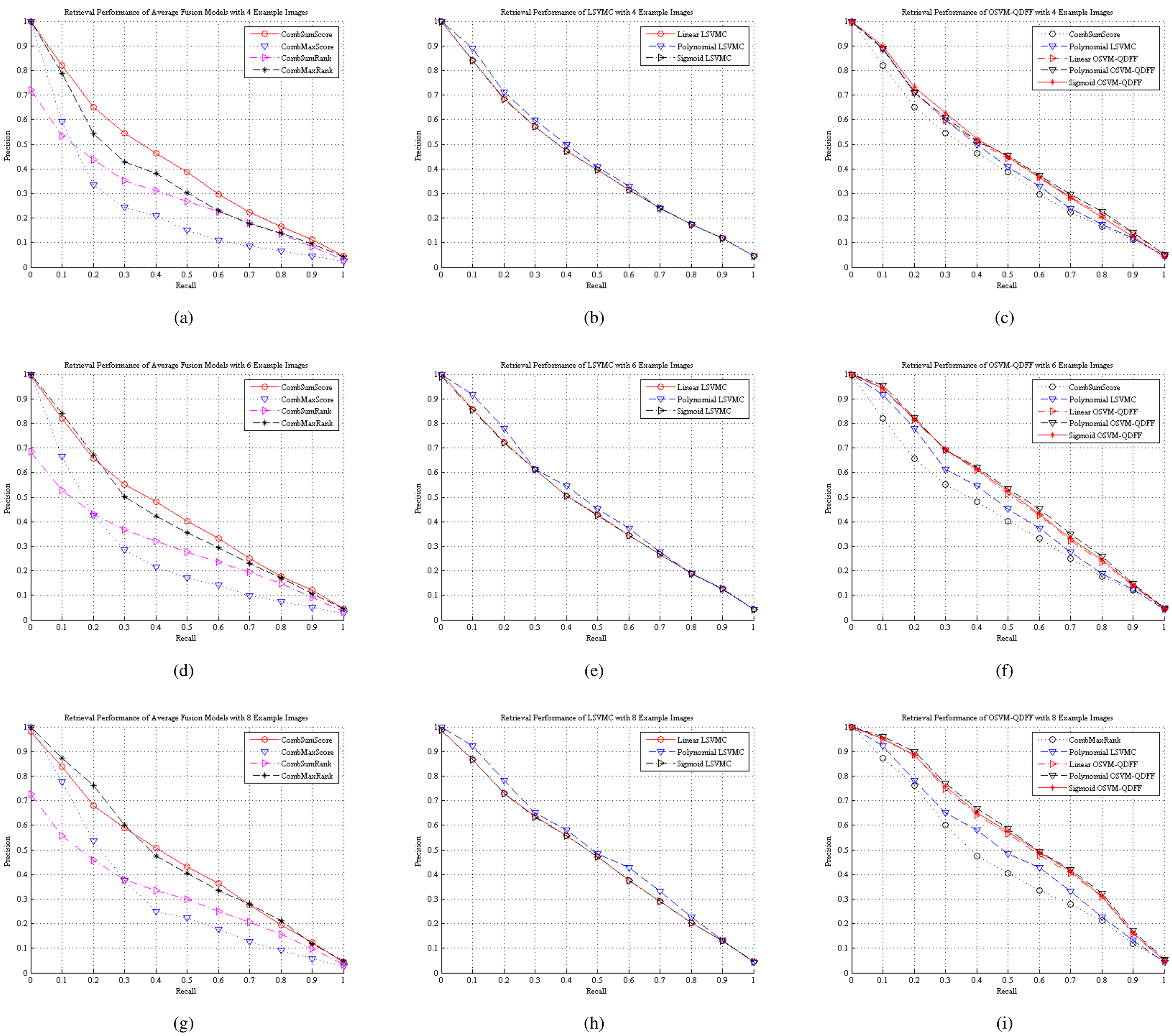

Fig. 2: Retrieval Performance of various feature fusion methods for image retrieval

and their retrieval performances are disappointed on our test dataset.

2) Experimental results and analysis: The plots in Figure 2 depict the average precision-recall graphs over all the 228 queries with different example images for the three comparison feature fusion methods:the *Average Fusion Models* [14] , the local aggregation function based on SVM (*LSVMC*) [10] and the One-Class SVM based query dependent feature fusion method (OSVM-QDFF) proposed in this paper.Table III present the mean average precision over the 228 queries for the three comparison feature fusion methods with different example images. Table IV presents the relative improvement of *OSVM-QDFF* to the best *Average Fusion Model* and the best *LSVMC*. The figure show that the proposed OSVM-QDFF always performances better than average fusion models and *LSVMC*. For the case of four query images, *OSVM-QDFF* improves the retrieval performance over the best average fusion model about $\%$**14** and about $\%$**5** over the the best *LSVMC*.For the case of six query images, *OSVM-QDFF* improves the retrieval performance over the best average fusion model about $\%$**26** and about $\%$**12** over the the best *LSVMC*.For the case of eight query images, *OSVM-QDFF* improves the retrieval performance over the best average fusion model about $\%$**29** and about $\%$**15** over the the best *LSVMC*.

For the *Average Fusion Model*, different features are configured of equal weighting for different queries which does not consider the special feature is not equally important for different queries, thus *Average Fusion Model* do the worst retrieval performance. For the *LSVMC*, the *Query-Independent* feature fusion problem has been regarded as a strict two class classification problem, which is not always reasonable since the negative examples randomly selected from the image collection can belong to any class and they usually do not cluster.

VII. Conclusion

Due to the huge growth of the World Wide Web, medical images are now available in large numbers in online repositories, and there exists the need to retrieval the images based on the modality, anatomic region and different acquisition views through automatically extracting visual information of the medical images, which is commonly known as content-based image retrieval (CBIR).Since each feature extracted from images just characterizes certain aspect of image content, multiple features are necessarily employed to improve the retrieval performance.Meanwhile, a special feature is not equally important for different image queries since a special feature has different importances in reflecting the content of different images.Although some research efforts have been reported to enhance the image retrieval performance taking the feature fusion approaches, most of existed feature fusion methods for image retrieval only utilize query independent feature fusion which usually apply a single feature fusion model for all the image queries and do not consider that a special feature is not equally important for different image queries, the others usually require the users to tune appropriate parameters for the feature fusion models for different image queries. In this paper, with multiply query samples, we formulate the feature fusion problem as a one class classification problem in the combined similarities space and present a query dependent feature fusion method for medical image retrieval based on One-Class support vector machine.The proposed query dependent feature fusion method can learn appropriate feature fusion models for different query based on multiply query samples, and the learned feature fusion models can reflect the different importances of a special feature for different image queries.The experimental results on the IRMA medical image collection demonstrate that the proposed method can improve the retrieval performance effectively and can outperform existed feature fusion methods for image retrieval.

Acknowledgment

This research work was supported by grants from the Chinese Special Project of Science and Technology: Core electronic devices, high-end general chips and infrastructural software (2010ZX01042-002-001) and National Natural Science Foundation of China (NSFC) under Grant No.61003017.

References

[1] T. M. Lehmann, M. O. Güld, C. Thies, B. Fischer, D. Keysers, M. Kohnen, H. Schubert, and B. B. Wein, "Content-based image retrieval in medical applications for picture archiving and communication systems," in *Proceedings SPIE*, vol. 5033, 2003, pp. 109–117.

[2] P. Mildenberger, M. Eichelberg, and E. Martin, "Introduction to the DICOM standard." *European radiology*, vol. 12, no. 4, p. 920, 2002.

[3] F. Florea, H. Müller, A. Rogozan, A. Geissbühler, and S. Darmoni, "Medical image categorization with MedIC and MedGIFT," *Medical Informatics Europe (MIE 2006)*, 2006.

[4] M. Rahman *et al.*, "Medical image retrieval with probabilistic multi-class support vector machine classifiers and adaptive similarity fusion," *Computerized Medical Imaging and Graphics*, vol. 32, no. 2, pp. 95–108, 2008.

[5] A. W. 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 and machine intelligence*, pp. 1349–1380, 2000.

[6] E. A. El-Kwae, H. Xu, and M. R. Kabuka, "Content-based retrieval in picture archiving and communication systems," *Journal of Digital Imaging*, vol. 13, no. 2, pp. 70–81, 2000.

[7] W. W. Chu, C. C. Hsu, A. F. Cardenas, and R. K. Taira, "Knowledge-based image retrieval with spatial and temporal constructs," *IEEE Transactions on Knowledge and Data Engineering*, vol. 10, no. 6, pp. 872–888, 1998.

[8] T. Deselaers, D. Keysers, and H. Ney, "Features for image retrieval: An experimental comparison," *Information Retrieval*, vol. 11, no. 2, pp. 77–107, 2008.

[9] A. Frome, Y. Singer, and J. Malik, "Image retrieval and classification using local distance functions," *Advances in Neural Information Processing Systems*, vol. 19, p. 417, 2007.

[10] J. Zhang and L. Ye, "Local aggregation function learning based on support vector machines," *Signal Processing*, vol. 89, no. 11, pp. 2291–2295, 2009.

[11] C. G. Snoek, M. Worring, and A. W. Smeulders, "Early versus late fusion in semantic video analysis," in *Proceedings of the 13th annual ACM International Conference on Multimedia*, 2005, p. 402.

[12] A. Kushki, P. Androutsos, K. N. Plataniotis, and A. N. Venetsanopoulos, "Retrieval of images from artistic repositories using a decision fusion framework," *IEEE Transactions on Image Processing*, vol. 13, no. 3, p. 277, 2004.

[13] J. Zhang and L. Ye, "Properties of series feature aggregation schemes," *World Review of Science, Technology and Sustainable Development*, vol. 7, no. 1, pp. 100–115, 2010.

[14] K. Donald and A. Smeaton, "A comparison of score, rank and probability-based fusion methods for video shot retrieval," *Image and Video Retrieval*, pp. 61–70, 2005.

[15] H. Shao, J. W. Zhang, W. C. Cui, and H. Zhao, "Automatic feature weight assignment based on genetic algorithm for image retrieval," in *2003 IEEE International Conference on Robotics, Intelligent Systems and Signal Processing, 2003. Proceedings*, 2003, pp. 731–735.

[16] R. Liyun, M. Shaoping, and L. Jing, "Feature fusion based on the average precision in image retrieval," *Journal of Computer Research and Development*, vol. 9, 2005.

[17] M. Ortega, Y. Rui, K. Chakrabarti, K. Porkaew, S. Mehrotra, and T. S. Huang, "Supporting ranked boolean similarity queries in MARS," *IEEE Transactions on Knowledge and Data Engineering*, vol. 10, no. 6, pp. 905–925, 1998.

[18] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: Image segmentation using expectation-maximization and its application to image querying," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1026–1038, 2002.

[19] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural computation*, vol. 13, no. 7, pp. 1443–1471, 2001.

[20] S. F. Chang, T. Sikora, and A. Purl, "Overview of the MPEG-7 standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 688–695, 2001.

[21] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 8, no. 6, pp. 460–473, 1978.

[22] T. M. Lehmann, B. B. Wein, J. Dahmen, J. Bredno, F. Vogelsang, and M. Kohnen, "Content-based image retrieval in medical applications: a novel multistep approach," vol. 3972, Dec. 1999, pp. 312–320.

[23] H. Müller, W. Müller, D. M. Squire, S. Marchand-Maillet, and T. Pun, "Performance evaluation in content-based image retrieval: Overview and proposals," *Pattern Recognition Letters*, vol. 22, no. 5, pp. 593–601, 2001.