This is the published version:


Available from Deakin Research Online:

http://hdl.handle.net/10536/DRO/DU:30039515

©2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Copyright : 2010, IEEE
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 on 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 health 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...
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 recent 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 important 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] and [18], 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 denoted as

$$F^I = \{f^I_1, \cdots, f^I_i, \cdots, f^I_m\}$$

(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

$$d_i(I, J) = D_i(f_i^I, f_i^J)$$

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

$$\Omega' = Q \cup \Omega$$

Given a image example $Q_i$ in the query $Q$, the distances to each image in the image collection $\Omega$ using the feature descriptor $f_j$ can be represented as

$$D_j(Q_i) = \{d_j(Q_i, Q_1), \ldots, d_j(Q_i, Q_s), d_j(Q_i, I_1), \ldots, d_j(Q_i, I_N)\}$$

where $D_j(Q_i)$ denotes the distances set for example image $Q_i$ on image set $\Omega$ 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

$$\tilde{d}_j = \frac{d_j - d_{j_{\text{min}}}}{d_{j_{\text{max}}} - d_{j_{\text{min}}}}$$

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

$$s_j = 1 - \tilde{d}_j$$

The similarities between the image example $Q_i$ and the image $I_j$ in image collection $\Omega'$ with $m$ different feature descriptors can be represent as a similarities vector

$$S(Q_i, I_j) = (s_1(Q_i, I_j), \ldots, s_m(Q_i, I_j))$$

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

$$\begin{bmatrix}
\begin{array}{ccc}
  s_1(Q_i, I_1) & \cdots & s_1(Q_i, I_q) \\
  \vdots & \ddots & \vdots \\
  s_1(Q_i, I_1) & \cdots & s_1(Q_i, I_q) \\
  \end{array}
\end{bmatrix}
\begin{bmatrix}
\begin{array}{ccc}
  \cdots & \cdots & \cdots \\
  \cdots & \cdots & \cdots \\
  \cdots & \cdots & \cdots \\
\end{array}
\end{bmatrix}
\begin{bmatrix}
\begin{array}{ccc}
  s_m(Q_i, I_1) & \cdots & s_m(Q_i, I_q) \\
  \vdots & \ddots & \vdots \\
  s_m(Q_i, I_1) & \cdots & s_m(Q_i, I_q) \\
  \end{array}
\end{bmatrix}$$

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

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

where $w = (w_1, \ldots, w_m)$ 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 $\Theta$. Thus the optimal Query-Dependent feature fusion for the example image $Q_i$ is to find appropriate feature weight vector $w = (w_1, \ldots, w_m)$ that can separate the relevant image set $\Theta$ from the image collection $\Omega'$ in the similarity space $\varphi(Q_i)$ as

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

where $\rho$ is the similarity threshold to separate the relevant image set $\Theta$ from the image collection $\Omega'$.

Since each example image $Q_i$ 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 $w = (w_1, \ldots, w_m)$ that can separate the relevant image set $\Theta$ from the image collection $\Omega'$ in the similarity spaces $\varphi(Q_1), \ldots, \varphi(Q_q)$ as

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

which is equally to find appropriate feature weight vector $w = (w_1, \ldots, w_m)$ that can separate the relevant image set $\Theta$ from the image collection $\Omega'$ in the combined similarity space $\varphi$ as

$$\begin{bmatrix}
\begin{array}{ccc}
  s_1(Q_1, I_1') & \cdots & s_1(Q_q, I_1') \\
  \vdots & \ddots & \vdots \\
  s_1(Q_1, I_1') & \cdots & s_1(Q_q, I_1') \\
  \end{array}
\end{bmatrix}
\begin{bmatrix}
\begin{array}{ccc}
  \cdots & \cdots & \cdots \\
  \cdots & \cdots & \cdots \\
  \cdots & \cdots & \cdots \\
\end{array}
\end{bmatrix}
\begin{bmatrix}
\begin{array}{ccc}
  \cdots & \cdots & \cdots \\
  \cdots & \cdots & \cdots \\
  \cdots & \cdots & \cdots \\
\end{array}
\end{bmatrix}
\begin{bmatrix}
\begin{array}{ccc}
  s_m(Q_1, I_1') & \cdots & s_m(Q_q, I_1') \\
  \vdots & \ddots & \vdots \\
  s_m(Q_1, I_1') & \cdots & s_m(Q_q, I_1') \\
  \end{array}
\end{bmatrix}$$

Notice that each image in image collection $\Omega'$ 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 $\varphi$ can be obtained by simply combing the similarities spaces

$$\varphi = \varphi(Q_1) \cup \cdots \cup \varphi(Q_q)$$

Consider that the size of $\Theta$ is much smaller compared to the size of the image collection $\Omega'$ as

$$|\Theta| \ll |\Omega'|$$

Thus the query dependent feature fusion fusion problem for the query $Q$ can be regards as a typical one class classification problem in the combined similarity space $\varphi$ with the training
data as
\[
\{(\text{Sim}(Q_i, I_j), L_{ij})| 1 \leq j \leq (N + q), 1 \leq i \leq q\}
\]
and
\[
L_{ij} = \begin{cases} 
1 & I_j \in Q \\
0 & I_j \notin Q 
\end{cases}
\]  
(15)

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

Treating the example image in the query \( Q \) equally, the similarities between the image \( I_i \) and the query \( Q \) can be computed as
\[
S(Q, I_i) = \sum_{i=1}^{q} S(Q_i, I_i)
\]
and the combined similarities between the query \( Q \) and the images in \( \Omega^* \) can be obtained as
\[
\text{Sim}(Q, \Omega^*) = \begin{bmatrix} 
\text{Sim}(Q_1, I_1) \\
\vdots \\
\text{Sim}(Q_q, I_1) \\
\vdots \\
\text{Sim}(Q_1, I_{N+q}) 
\end{bmatrix} = \sum_{i=1}^{q} \varphi(Q_i) \cdot w^T
\]
(17)

In summary, the query dependent feature fusion problem for the query \( Q \) is to find appropriate feature weight vector \( w = (w_1, \cdots, w_m) \) 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 \( \varphi \). 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 class classification problem in the combined similarity space \( \varphi \) with the positive examples \( \varphi^+ \)
\[
\varphi^+ = \{s_1, s_2, \cdots, s_l\} \subset \varphi
\]
(18)
where \( l = p \times p \) 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
\[
f(s) = w \cdot s - \rho
\]
(19)
where \( w \) 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 \( \varphi^+ \) as
\[
f(s_i) > 0, \quad i = 1, 2, \cdots, t
\]
(20)
and with the largest margin to the origin in such hyperplanes as Fig.1 presents.
\[
\max \frac{\rho}{\|w\|}
\]
(21)
By properly chosen nonlinear function \( \phi \), 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
\[
\phi: \varphi \rightarrow F
\]
(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 \( \frac{\rho}{\|w\|} \) and satisfy \( f(s_i) > 0 \) for all the positive examples \( s_i \) in \( \varphi^+ \) as
\[
f(s) = w \cdot \phi(s) - \rho
\]
(23)
The linear One-Class SVM can be regards as a typical nonlinear One-Class SVM with the mapping function
\[
\phi(s) = s
\]
(24)

With the training data \( \phi(s_1), \phi(s_2), \cdots, \phi(s_l) \), the optimal hyperplane \( w \) can be found by solving the following quadratic programming problem [19]
\[
\begin{cases}
\min & \frac{1}{2} \|w\|^2 - \rho \\
\text{s.t.} & w \cdot \phi(s_i) \geq \rho \quad i = 1, 2, \cdots, t
\end{cases}
\]
(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 \( \varphi^+ \) from \( \varphi \). 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.
\[
\begin{cases}
\min & \frac{1}{2} \|w\|^2 - \rho + \frac{1}{\nu} \sum_{i=1}^{l} \xi_i \\
\text{s.t.} & w \cdot \phi(s_i) \geq \rho - \xi_i, \quad \xi_i \geq 0, \quad i = 1, 2, \cdots, t
\end{cases}
\]
(26)
trade-off between maximizing the distance from the origin and separating most of the relevant samples. After introducing Lagrange multipliers $\alpha_i$ for each training sample, the dual problem of the optimization problem can be obtained as

$$\begin{align*}
\max & \quad \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j \phi(s_i) \cdot \phi(s_j) \\
\text{s.t.} & \quad 0 \leq \alpha_i \leq \frac{1}{n} \\
& \quad \sum_{i=1}^{l} \alpha_i = 1
\end{align*}$$ (27)

Solving the dual problem leads to

$$w = \sum_{i=1}^{l} \alpha_i \phi(s_i), 0 \leq \alpha_i \leq \frac{1}{n}$$ (28)

and the corresponding decision function becomes

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

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

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

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

$$Sim(Q_i, I_j) = S(Q_i, I_j)^T \cdot w^T$$ (31)

Thus the combined similarity between the image example $Q_i$ and the image $I_j$ in $\Omega$ with the decision function in the combined similarity space $\varphi$ can be represented as

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

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

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

where $\mu$ and $\sigma$ are the average value and the standard deviation of the similarities obtained with example image $Q_i$ as

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

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

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

$$Sim'(Q, I_j) = \sum_{i=1}^{q} \exp(Sim'(Q_i, I_j))$$ (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 $8 \times 8$ 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:

$$D_{cl}(Q, I) = \sqrt{\sum_{i} (Y_{Qi} - Y_{Ii})^2 + \sum_{i} (Cb_{Qi} - Cb_{Ii})^2} + \sum_{i} (Cr_{Qi} - Cr_{Ii})^2$$ (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:

$$D_{tamura}(Q, I) = \sqrt{\sum_{i} (T_{Qi} - T_{Ii})^2}$$ (38)

**Edge Feature:** The Edge Histogram Descriptor (EHD) [20] is used to represent the global edge feature in this paper. The EHD represents local edge distribution in an image by dividing the image into $4 \times 4$ 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 $16 \times 5 = 80$ bins is obtained, corresponding to a 80-dimensional feature vector. The distance between two EHD vectors is calculated as shown below:

$$D_{ehd} = \sum_{i} |H_{Qi} - H_{Ii}|$$ (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
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 experts 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 experts.

TABLE I: Final parameter set for SVMs

<table>
<thead>
<tr>
<th>Kernel</th>
<th>$C$</th>
<th>$\gamma$</th>
<th>$r$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Polynomial</td>
<td>256</td>
<td>0.0625</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Sigmoid</td>
<td>32</td>
<td>0.0313</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

TABLE II: Final parameter set for One-Class SVMs

<table>
<thead>
<tr>
<th>Kernel</th>
<th>$\nu$</th>
<th>$\gamma$</th>
<th>$r$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>0.05</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Polynomial</td>
<td>0.0625</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Sigmoid</td>
<td>0.5</td>
<td>0.0313</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

TABLE III: Mean average precision of different feature fusion methods

<table>
<thead>
<tr>
<th>Examples</th>
<th>CombSum Score</th>
<th>CombMax Score</th>
<th>CombSum Rank</th>
<th>CombMax Rank</th>
<th>Linear LSVMC</th>
<th>Polynomial LSVMC</th>
<th>Sigmoid LSVMC</th>
<th>Linear OSVM-QDFF</th>
<th>Polynomial OSVM-QDFF</th>
<th>Sigmoid OSVM-QDFF</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 Examples</td>
<td>0.4128</td>
<td>0.2279</td>
<td>0.2806</td>
<td>0.3518</td>
<td>0.4298</td>
<td>0.4448</td>
<td>0.4206</td>
<td>0.5199</td>
<td>0.5462</td>
<td>0.5468</td>
</tr>
<tr>
<td>6 Examples</td>
<td>0.4244</td>
<td>0.263</td>
<td>0.2857</td>
<td>0.4048</td>
<td>0.4509</td>
<td>0.4763</td>
<td>0.4507</td>
<td>0.5199</td>
<td>0.5342</td>
<td>0.5254</td>
</tr>
<tr>
<td>8 Examples</td>
<td>0.4351</td>
<td>0.2868</td>
<td>0.2969</td>
<td>0.4291</td>
<td>0.4721</td>
<td>0.5024</td>
<td>0.4719</td>
<td>0.564</td>
<td>0.5801</td>
<td>0.5685</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Kernel</th>
<th>$C$</th>
<th>$\gamma$</th>
<th>$r$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Polynomial</td>
<td>256</td>
<td>0.0625</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Sigmoid</td>
<td>32</td>
<td>0.0313</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

TABLE III: Mean average precision of different feature fusion methods

<table>
<thead>
<tr>
<th>Examples</th>
<th>CombSum Score</th>
<th>CombMax Score</th>
<th>CombSum Rank</th>
<th>CombMax Rank</th>
<th>Linear LSVMC</th>
<th>Polynomial LSVMC</th>
<th>Sigmoid LSVMC</th>
<th>Linear OSVM-QDFF</th>
<th>Polynomial OSVM-QDFF</th>
<th>Sigmoid OSVM-QDFF</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 Examples</td>
<td>0.4128</td>
<td>0.2279</td>
<td>0.2806</td>
<td>0.3518</td>
<td>0.4298</td>
<td>0.4448</td>
<td>0.4206</td>
<td>0.5199</td>
<td>0.5462</td>
<td>0.5468</td>
</tr>
<tr>
<td>6 Examples</td>
<td>0.4244</td>
<td>0.263</td>
<td>0.2857</td>
<td>0.4048</td>
<td>0.4509</td>
<td>0.4763</td>
<td>0.4507</td>
<td>0.5199</td>
<td>0.5342</td>
<td>0.5254</td>
</tr>
<tr>
<td>8 Examples</td>
<td>0.4351</td>
<td>0.2868</td>
<td>0.2969</td>
<td>0.4291</td>
<td>0.4721</td>
<td>0.5024</td>
<td>0.4719</td>
<td>0.564</td>
<td>0.5801</td>
<td>0.5685</td>
</tr>
</tbody>
</table>

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

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 experts 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 experts.

TABLE I: Final parameter set for SVMs

<table>
<thead>
<tr>
<th>Kernel</th>
<th>$C$</th>
<th>$\gamma$</th>
<th>$r$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Polynomial</td>
<td>256</td>
<td>0.0625</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Sigmoid</td>
<td>32</td>
<td>0.0313</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

TABLE II: Final parameter set for One-Class SVMs

<table>
<thead>
<tr>
<th>Kernel</th>
<th>$\nu$</th>
<th>$\gamma$</th>
<th>$r$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>0.05</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Polynomial</td>
<td>0.0625</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Sigmoid</td>
<td>0.5</td>
<td>0.0313</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

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
  \[ K(s_i, s_j) = s_i^T s_j \]  

- the polynomial machines with kernel function
  \[ K(s_i, s_j) = (\gamma s_i^T s_j + r)^d, \gamma > 0 \]

where $d$ is the degree of the polynomial kernel.

- the sigmoid machines with kernel function
  \[ K(s_i, s_j) = \tan(\gamma s_i^T s_j + r) \]

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 $\gamma$, $r$ 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 $\gamma$ 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,
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


