Abstract—In this paper we study the problem of content-based image retrieval. In this problem, the most popular performance measure is the top precision measure, and the most important component of a retrieval system is the similarity function used to compare a query image against a database image. However, up to now, there is no existing similarity learning method proposed to optimize the top precision measure. To fill this gap, in this paper, we propose a novel similarity learning method to maximize the top precision measure. We model this problem as a minimization problem with an objective function as the combination of the losses of the relevant images ranked behind the top-ranked irrelevant image, and the squared Frobenius norm of the similarity function parameter. This minimization problem is solved as a quadratic programming problem. The experiments over two benchmark data sets show the advantages of the proposed method over other similarity learning methods when the top precision is used as the performance measure.

I. INTRODUCTION

In the problem of content-based image retrieval, the two basic steps of the content-based image retrieval are image visual feature extraction, and the similarity calculation. Recently, the similarity learning problem has been a hot topic in the field of machine learning. This problem tries to learn a parameterized similarity function from a training set, so that a targeted performance can be optimized [1]. To evaluate the performance of a retrieval system, common measures for retrieval performance are taken, including the area under the receiver operating characteristic (ROC) curve, and the top precision value [2], [3], [4]. The top precision measure is defined as the portion of the top-ranked relevant images out of all the relevant database images. The top-ranked relevant images are further defined as the database images which are ranked before the first ranked irrelevant database image. Top precision is an important performance measure for retrieval problems, however, all the existing similarity learning methods ignore the performance measures, and the resulted similarity function cannot give an optimal performance measure of top precision.

To fill the gap between similarity learning and top precision optimization in the problem of content-based image retrieval, we propose a problem of top precision-optimal similarity learning, together with a corresponding solution using a novel algorithm. Our contributions in this paper are divided into two parts, which are given as followed. The proposed problem is to learn a similarity function to compare a query image against a database image. With the similarity scores between query and database images, when retrieval task is performed, we can obtain a maximum top precision measure. The problem is to adjust the parameters of the similarity function, so that the top precision can be maximized. The proposed novel similarity function learning algorithm is to maximize the top precision over a training set of query images and a database of images. We design the similarity function as a linear function of the two feature vectors of a query image and a database image. To learn the parameter matrix of the similarity function, we propose to maximize the top precision for each query, and parallelly minimize the squared Frobenius norm of the parameter matrix in order to prevent the emergence of the over-fitting problem. The optimization problem is modeled by a hinge loss and solved as a quadratic programming.
The rest parts of this paper are organized as followed. In section II, we introduce the proposed similarity learning algorithm. In section III, we give the experimental results of the proposed method. In section IV, we give the conclusion of this paper with future works.

II. PROPOSED SIMILARITY LEARNING METHOD

We assume we have an image database of \( m \) images, \( \mathcal{D} = \{x_1, \cdots, x_m\} \), where \( x_j \in \mathbb{R}^d \) is the feature vector of the \( j \)-th database image, a set of query images \( \mathcal{Q} = \{z_1, \cdots, z_n\} \), where \( z_i \in \mathbb{R}^d \) is the feature vector of the \( i \)-th query image, and a relevance matrix \( Y = \{y_{ij}\} \in \{0, 1\}^{n \times m} \), where \( y_{ij} = 1 \) if \( z_i \) is relevant to \( x_j \), and 0 otherwise. We define a similarity function for a query \( z \) and a database image \( x \) parameterized by a matrix \( W \in \mathbb{R}^{d \times d} \),

\[
s(z, x) = z^\top W x.
\]  

(1)

The first ranked irrelevant database image of a query \( z_i \), denoted as top irrelevant, is defined as the irrelevant database image with the largest similarity to the query, \( x_{\phi_i}, \) where \( \phi_i = \arg \max_{k,y_{ik}=0} s(z_i, x_k) \).

To obtain a good performance measure of top precision, we hope that for any pair of query and relevant database image, \((z_i, x_j)_{i,j:y_{ij}=1}\), their similarity \((z_i, x_j)\) is larger than the similarity to the top irrelevant database image plus a margin, \( s(z_i, x_{\phi_i}) + 1 \),

\[
s(z_i, x_j) > \max_{k:y_{ik}=0} s(z_i, x_k) + 1,
\]  

(2)

\[
\Rightarrow \max_{k:y_{ik}=0} s(z_i, x_k) - s(z_i, x_j) + 1 \geq 0, \quad \forall i, j : y_{ij} = 1.
\]

(4)

To maximize the top precision, we propose to learn the similarity function parameter matrix \( W \) by penalizing the case which doesn’t satisfies this condition,

\[
0 \leq \max_{k:y_{ik}=0} s(z_i, x_k) - s(z_i, x_j) + 1 \geq 0, \quad \forall i, j : y_{ij} = 1.
\]

(5)

Thus the loss function for each \((z_i, x_j)_{i,j:y_{ij}=1}\) is,

\[
\ell(z_i, x_j; W) = \max_{k:y_{ik}=0} s(z_i, x_k) - s(z_i, x_j) + 1.
\]

(6)

The optimization problem is modeled as,

\[
\min_{W, \xi_{ij}|i,j:y_{ij}=1} \left\{ \frac{1}{2} \|W\|_F^2 + C \sum_{i,j:y_{ij}=1} \xi_{ij} \right\},
\]

subject to \(\forall i, j : y_{ij} = 1, \xi_{ij} \geq 0, \xi_{ij} \geq z_i^\top W (x_k - x_j) + 1\),

(7)

where \(\xi_{ij}\) is slack variable of \(6\), \(\frac{1}{2} \|W\|_F^2\) is the squared Frobenius norm of \(W\) to prevent the over-fitting problem, and \(C\) is the tradeoff parameter.

The Lagrange function of \(7\) is

\[
\mathcal{L} = \frac{1}{2} \|W\|_F^2 + C \sum_{i,j:y_{ij}=1} \xi_{ij}
\]

\[
- \sum_{i,j:y_{ij}=1} \alpha_{ij} \xi_{ij}
\]

\[
- \sum_{i,j,y_{ij}=1} \beta_{ijk} (\xi_{ij} - z_i^\top W (x_k - x_j) - 1)
\]

(8)

where \(\alpha_{ij} \geq 0\) is the Lagrange multiplier of the constraint \(\xi_{ij} \geq 0\), and \(\beta_{ijk} \geq 0\) is the Lagrange multiplier of the constraint \(\xi_{ij} \geq z_i^\top W (x_k - x_j) + 1\). The dual form of the optimization problem of \(7\) is given as follows,

\[
\max_{\alpha_{ij}|i,j:y_{ij}=1;\beta_{ijk}|i,k:y_{ik}=1;\beta_{ijk}|j,k:y_{jk}=1} \min_{W, \xi_{ij}|i,j:y_{ij}=1} \mathcal{L}
\]

subject to \(\alpha_{ij} \geq 0, \forall i, j : y_{ij} = 1, \beta_{ijk} \geq 0, \forall i, j, k : y_{ij} = 1, y_{jk} = 0\).

(9)

By setting the derivative of \(\mathcal{L}\) with regard to \(\xi_{ij}\) to zero, we have

\[
\frac{\partial \mathcal{L}}{\partial \xi_{ij}} = \frac{\partial \mathcal{L}}{\partial \alpha_{ij}} \Rightarrow \alpha_{ij} = C - \sum_{k:y_{ik}=0} \beta_{ijk}\]

(10)

\[
\Rightarrow \alpha_{ij} \geq \sum_{k:y_{ik}=0} \beta_{ijk}.
\]

(11)

By substituting \(10\) to \(\mathcal{L}\), we have

\[
\mathcal{L} = \frac{1}{2} \|W\|_F^2 - \sum_{i,j,k:y_{ik}=1; y_{ijk}=1} \beta_{ijk} z_i^\top W (x_j - x_k)
\]

\[
+ \sum_{i,j,k:y_{ik}=1; y_{ijk}=1} \beta_{ijk}.
\]

(11)

By setting the derivative of \(\mathcal{L}\) with regard to \(W\) to zero, we have

\[
\frac{\partial \mathcal{L}}{\partial W} = W - \sum_{i,j,k:y_{ik}=1; y_{ijk}=1} \beta_{ijk} z_i (x_j - x_k)^\top \Rightarrow W = \sum_{i,j,k:y_{ik}=1; y_{ijk}=1} \beta_{ijk} z_i (x_j - x_k)^\top.
\]

(12)

By substituting \(W\) of \(12\) to \(\mathcal{L}\) of \(11\), we have
This is a quadratic function of the multiplier variables, \( \beta_{ijk} | i,j,k:y_{ij}=1, y_{ik}=0 \). The optimization problem of (9) is transferred to

\[
\begin{align*}
\max_{\beta_{ijk} | i,j,k:y_{ij}=1, y_{ik}=0} & \left\{ -\frac{1}{2} \sum_{i,j,k:y_{ij}=1,y_{ik}=0} \beta_{ijk} \beta_{i'j'k'} z_{i'}^T z_{j'} (x_{i'} - x_{k'}) + \sum_{i,j,k:y_{ij}=1,y_{ik}=0} \beta_{ijk} \right\} \\
\text{s.t.} & \sum_{k:y_{ik}=0} \beta_{ijk}, \text{ and } \beta_{ijk} \geq 0, \\
& \forall i,j,k : y_{ij} = 1, y_{ik} = 0.
\end{align*}
\]

This is a quadratic programming (QP) problem, and we can use a standard active set algorithm to solve this problem. After the \( \beta_{ijk} | i,j,k:y_{ij}=1, y_{ik}=0 \) are solved, we can recover \( W \) from (12).

### III. Experiments

#### A. Experimental protocol

In this experiment, we used two benchmark image databases, which are Indoor image database [5], and Caltech256 image database [6]. Then total number of images/number classes of these two databases are 15,620/67 and 30,670/256. Given a database, we first split it to a set of queries and a set of database images, and these queries are further split into a training set and a test set. We apply the proposed similarity learning algorithm to the training set of queries and the database images to learn the similarity function. Then we use the similarity function to conduct the retrieval task for the test set of queries. The performance is evaluated by top precision measure.

#### B. Experimental results

We compare our proposed maximum top precision similarity learning method (MTPS), to several state-of-the-art similarity learning methods, online algorithm for scalable image similarity (OASIS) [7], online multiple kernel similarity (OMKS) [8], Boosted distance (BD) [9], stochastic intersection kernel machine (SIKMA) [10], and visuality-preserving distance metric (VPDM) [11]. The results of the comparison over the four benchmark databases are given in Figure 1. It is obvious that the proposed algorithm, MTPS, outperforms the compared similarity learning algorithms over both the two benchmark databases. In the results over the Caltech256, the top precisions of almost all the compared algorithms are below 0.16, while our algorithms, MTPS, have top precisions higher than 0.18.

In addition, we test the running time of the proposed algorithm over the two benchmark databases, displayed in Figure 2. Our proposed algorithm MTPS exhibits shorter running time than others, except for the OASIS and OMKS, suggesting that MTPS provides a comparably efficient way against the problem, and shows a promising improvement space for the future work.

### IV. Conclusion and future works

In this paper we study the problem of similarity learning for content-based image retrieval. The critical novelty of this work is it learns the similarity to maximize the top precision measure over the training set. We model the learning problem as a support top irrelevant database image weighting problem to obtain the optimal similarity function. The experiments over some benchmark databases show its advantages over other similarity learning methods. In the future, we will investigate...
to use some other similarity function as similarity measure instead of linear function, such as Bayesian network \[12\], \[13\], \[14\], and also to develop novel algorithms of other machine learning problems and applications besides similarity learning, to maximize top precision measure, such as importance sampling \[15\], \[16\], \[17\], portfolio choices \[18\], \[19\], multimedia technology \[20\], \[21\], \[22\], \[23\], \[24\], \[25\], \[26\], \[27\], \[28\], \[29\], computational biology \[30\], \[31\], \[32\], \[33\], \[34\], big data processing \[35\], \[36\], \[37\], \[38\], \[39\], computer vision \[40\], \[41\], \[42\], \[43\], \[44\], \[45\], \[46\], \[47\], \[48\], \[49\], \[50\], \[51\], \[52\], \[53\], information security \[54\], \[55\], \[56\], \[57\], \[58\], \[59\], and medical imaging \[60\], \[61\], \[62\], etc.

ACKNOWLEDGEMENTS

The study is supported by a grant from Provincial Key Laboratory for Computer Information Processing Technology, Soochow University, China.

REFERENCES


