Pattern Recognition Letters 34 (2013) 646–654

Contents lists available at SciVerse ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Content-based texture image retrieval using fuzzy class membership

Sudipta Mukhopadhyay [a,*], Jatindra Kumar Dash [a], Rahul Das Gupta [b]

[a] Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur 721 302, India

[b] Center for Educational Technology, Indian Institute of Technology, Kharagpur 721 302, India

ARTICLE INFO

Article history:
Received 8 June 2012
Available online 12 January 2013

Communicated by G. Borgefors

Keywords:
Content-based image retrieval (CBIR)
Texture feature
Wavelet
Neural network
Fuzzy membership
Weighted distance metric

ABSTRACT

There is no single best representation of images that can separate different classes with well defined boundaries in the feature space. Therefore, content-based image retrieval (CBIR) using conventional distance metric is not efficient in the low level image feature space viz. texture. Classifier based retrieval approaches (classification followed by retrieval) classify the query image and retrieve images only from the identified class. The performance of such approaches greatly relies on the performance of classifier. For each correct classification of query image, these systems yield high retrieval accuracy and for each misclassification the systems result in complete failure. It results huge variance in performance. This paper proposes a novel approach to content-based image retrieval called "*Class Membership-based Retrieval*" that addresses the limitations of both conventional distance based and conventional classifier based retrieval approaches. The proposed method consists of two steps. First, the class label and fuzzy class membership of query image is computed using neural network. In the second step, the retrieval is performed using a combination of simple and weighted (class membership based) distance metric in *complete search space* unlike the conventional classifier based retrieval techniques. The proposed technique also provides flexibility in reducing the search space in steps increasing the speed of retrieval at the cost of gradual reduction in accuracy. The performance of the method is evaluated using three texture data sets varying in orientations, complexity and number of classes. Experimental results support the proposed technique favorably when compared with other promising texture retrieval schemes.

1. Introduction

Content based image retrieval (CBIR) is a technique that uses visual contents to index and retrieve similar images from large image database, given a query image. Query formation, visual content extraction (in the form of a feature vector), and checking similarity between the query image and the images stored in the database in feature space are the different steps of a CBIR system. It is an active research area since last two decades and has application in digital libraries and multimedia databases. Early image retrieval techniques were based on textual annotation of images. Later, several CBIR systems have been developed to search through image databases using visual contents like color, texture, and shape attributes e.g. (Bach et al., 1996; Carson et al., 2002; Cox et al., 2000; Flickner et al., 1995; Laaksonen et al., 2000; Ma and Manjunath, 1997; Pentland et al., 1994; Smith and Chang, 1996; Ogle and Stonebraker, 1995). Comprehensive and elaborate literature survey of work on CBIR can be found in (Rui et al., 1999; Smeulders et al., 2000; Veltcamp and Tanase, 2000).

Texture as a primitive visual content has been studied for many years in image analysis and CBIR. Many CBIR systems have been developed using texture features as a visual content. Most of the CBIR systems proposed use image/signal processing technique to calculate a fixed set of features that represents the image and use a distance measure (e.g. Euclidian, city block and Mahalanobis distance) to rank the images in the database based on their similarity with the query image. We call this retrieval approach as the *"Traditional approach"*. For this approach, it is expected that the features used should represent image data effectively. The effectiveness of the representation space (i.e. feature space) is determined by how well patterns/images from different classes are separated (Jain et al., 2000). Several new feature descriptors are proposed in the literature (Guo et al., 2010a,b; Liao et al., 2009; Murala et al., 2012; Pietikainen et al., 2000; Ojala et al., 1996, 2002). But, there is no single best representation of images (Rui et al., 1999) that can separate different classes with well defined boundaries in the feature space, particularly for databases with a large number of texture classes. Therefore, image retrieval using simple distance based criterion in the feature space is inefficient, particularly when the image features correspond to low level attributes such as texture (Guo et al., 2000).

Few researchers have proposed different approaches for CBIR to overcome this drawback of simple distance measure. Ma and Manjunath (1996) proposed a simple hybrid learning based approach to retrieve similar images. They used a trained hybrid

* Corresponding author.
E-mail address: smukho@gmail.com (S. Mukhopadhyay).

0167-8655/$ - see front matter
http://dx.doi.org/10.1016/j.patrec.2013.01.001

neural network to classify the query image into one of the given clusters and then select the n most similar images within that cluster using Euclidian distance. The basic idea is to reduce the search space by prior learning of similarity among images in the database. The method demands many parameters to be adjusted heuristically for optimal result. Satini and Jain (1999) proposed a similarity measure based on Fuzzy logic. Minka and Picard (1996) reported a system, which learns grouping of similar images from positive and negative examples during query session. Guo et al. (2000) used Support Vector Machine (SVM) to measure image similarities and proposed a new metric called 'distance-from-boundary'. The basic idea behind is that a non-linear boundary separates images from the dissimilar ones. Carkacioglu and Vural (2002) developed a configurable approach where the original feature space is non-linearly transformed into a new feature space using artificial neural network and the distance between the feature space of different classes are adjusted by learning. The objective is to obtain maximum inter class separation while maintaining minimum intra class variance in the transformed feature domain. Chen et al. (1998) used L1 distance to organize similar images in 'similarity pyramids' structure. The resulting organization is used for indexing and retrieval. Dy et al. (2003) developed a hierarchical approach called 'Customized-Queries' Approach (CQA) that uses two different feature sets for comparing similarity at each level and for each class. In most of the above literature the performances of the proposed method were compared with simple distance based retrieval approach in the same feature space. The retrieval performance of classifier or learning based retrieval approaches greatly relies on the performance of classifier. For each correct classification of query image, the system yields high retrieval efficiency and for each misclassification the system will result in complete failure. Therefore, although having reasonable overall retrieval efficiency, these systems produce a huge variation in their results.

In this paper, we proposed a novel approach for CBIR called *'Class Membership-based Retrieval'*(CMR) that follows two step approach (computation of class label and fuzzy class membership value of query image followed by retrieval using simple or weighted distance measure) to retrieve similar images against a given query image. Wavelet transform is used to extract statistical texture features. These sets of features are used for computation of class membership value of the query image and distance between the query and database images in feature space. The proposed CMR scheme differs from all other approaches in several ways. First, it does not require any manual regrouping of perceptually similar textures from different classes to reduce the total number of classes for learning. Second, it uses a single feature set at all levels. Third, it makes use of the class label of query image (obtained by the classifier) to adapt the distance metric to be used. Finally, it allows the retrieval to be performed from the complete search space unlike the conventional classifier based approaches.

The retrieval performances of the proposed CMR scheme are compared with the performance of simple distance based retrieval approaches. Two distance metrics (Kokare et al., 2006; Murala et al., 2012) are used to show the consistent behaviour of the proposed paradigm. The performance of the proposed scheme is also compared with the traditional classifier based retrieval method.

The main contributions of this paper are:

(i) Multi layered feed forward neural network with one hidden layer based on scale conjugate gradient back-propagation learning algorithm is used to learn the class label and fuzzy class membership value of the query image.
(ii) A weighted distance metric is proposed to calculate the similarity between the query image and the images in the database during retrieval. The proposed distance metric is inversely proportional to the output fuzzy class membership values of the query image.
(iii) The proposed CMR method significantly improves the average retrieval performance and reduces the fluctuations as it allows search for similar images in the complete search space unlike traditional classifier based approaches those restrict the search space to single output class.
(iv) The search space may be reduced in steps to increase retrieval speed at the cost of gradual reduction in accuracy.

2. Proposed Class Membership-based Retrieval Method

Fig. 1 shows the block diagram of the proposed retrieval scheme. This method comprises of three steps such as feature extraction, computation of class label and fuzzy class membership for the query image using a trained neural network, and retrieval based on class label and fuzzy class membership value of the query image.

2.1. Feature extraction

The image X is decomposed into four sub-bands (viz. LL1, LH1, HL1 and HH1) using discrete wavelet transform (DWT) (Mallat, 1989). The sub-band LL corresponds to approximate wavelet coefficients and the sub-bands labeled LH, HL and HH correspond to detailed wavelet coefficients. To obtain wavelet coefficient of next level of decomposition, LL alone is considered. This provides a second level wavelet decomposition and corresponding sub-bands are LL2, LH2, HL2 and HH2. This process is continued till the third level. The energies (using L_1 norm) and standard deviations of wavelet coefficients for 10 sub-bands (i.e. LHi, HLi, HHi; for $i = 1, 2, 3$ and LLi; for $i = 3$) are calculated.

Let $E_k(X)$ and $\sigma_k(X)$ be the energy and standard deviation of wavelet coefficients for the kth sub-band. Then

$$E_k(X) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}|w_k(i,j)| \tag{1}$$

$$\sigma_k(X) = \sqrt{\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}(w_k(i,j) - \mu_k)^2} \tag{2}$$

where $w_k(i,j)$ is the value of wavelet coefficient at (i,j) for kth sub-band of dimension $N \times N$ and mean value μ_k of the sub-band is given as

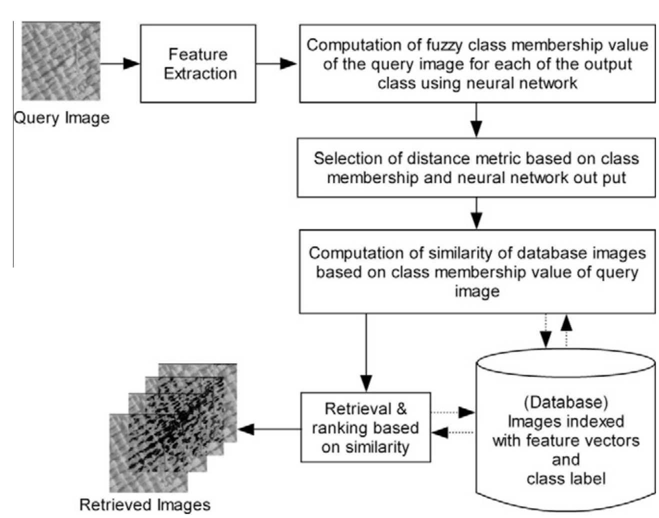

Fig. 1. Class Membership-based Retrieval.

$$\mu_k = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_k(i,j) \tag{3}$$

Each of these features $E_k(X)$ and $\sigma_k(X)$ are then normalised over the entire database as follows

$$E'_k(X) = \frac{E_k(X)}{\alpha(E_k)} \tag{4}$$

$$\sigma'_k(X) = \frac{\sigma_k(X)}{\alpha(\sigma_k)} \tag{5}$$

where, $\alpha(E_k)$ and $\alpha(\sigma_k)$ are the standard deviations of respective features (i.e. energy, $E_k(X)$ and standard deviation, $\sigma_k(X)$ of wavelet coefficients for the kth sub-band) over the entire database.

For a given image X, the 'd'-dimensional ($d = 20$ here) feature vector is now represented as:

$$\vec{f}(X) = [E'_1(X), \sigma'_1(X), E'_2(X), \sigma'_2(X), \ldots, E'_{10}(X), \sigma'_{10}(X)] \tag{6}$$

2.2. Computation of fuzzy class membership for query image

Multi layered feed forward neural network with one hidden layer is trained using a number of labeled image data and the associated feature set. The goal of learning is to achieve maximally nonlinear separations between several texture classes in the feature space. The texture features are extracted from the input query image as described in Section 2.1. This is fed to the network. To classify feature vectors of dimension 20 into a number c of classes the neural network essentially needs $d = 21$ nodes at the input layer (one node at the input layer is used as a biased input) and c (number of output class) nodes at the output layer. The network uses $(2d + 1)$ neurons in the hidden layer as stipulated by Kolmogorov's theorem (Kolmogorov, 1975). The neurons used in the output layer of the network have sigmoidal activation function. The scaled conjugate gradient algorithm (Moller, 1993) is selected for fast and robust training with the back propagation method. A 4-fold cross validation approach is followed as explained in Section 5 to evaluate the classification as well as retrieval performance.

Now for c-class problem with a number n_t of training samples the training set consists of n_t order-pairs of vectors $\{\vec{f}(X_k), \vec{t_k}\}_{k=1}^{n_t}$, where $\vec{f}(X_k)\epsilon R^{20}$ refers the feature vector of kth training image and $\vec{t_k}\epsilon\{0,1\}^c$ denotes the kth target vector. Specifically, if the class label of the kth training sample be l, then the ith component of $\vec{t_k}$ is given by $t_{ik} = 0$ for all $i \neq l$ and $t_{lk} = 1$. For each classification the class corresponding to the maximum output neuron will be chosen. After an extensive training the fuzzy class membership values of the input query image are obtained by normalizing the outputs of the neural network.

Fuzzy class membership value of image X having feature vector $\vec{f}$(X) for jth class is given as:

$$\mu_j(\vec{f}(X)) = \frac{o_j(\vec{f}(X))}{\sum_{i=1}^{c} o_i(\vec{f}(X))} \tag{7}$$

where $o_j(\vec{f}(X))$ is the output of jth neuron at the output layer for input feature vector $\vec{f}$(X).

The defined fuzzy class membership satisfies the following criterions (Lee, 2005):

$$0 \leqslant \mu_j(\vec{f}(X)) \leqslant 1 \forall \quad j \in 1, 2, 3, \ldots, c \tag{8}$$

$$\sum_{j=1}^{c} \mu_j(\vec{f}(X)) = 1 \tag{9}$$

The proposed method calculates the fuzzy membership of query image automatically from the given feature set using a trained neural network. This does not require mapping of the feature sets into set of fuzzy predicates as used by Satini and Jain (1999) to calculate similarity based on fuzzy logic. In addition, unlike Ma and Manjunath, 1996, the proposed technique does not require manual regrouping of perceptually similar textures from different classes to reduce the number of classes for learning.

2.3. Weighted distance metric

For classifier based retrieval approach, if the search space is limited to the class suggested by the classifier, for each correct classification it achieves 100% retrieval performance. However, in case of misclassification, the scheme completely fails to perform. To take care of retrieval efficiency in both correct and misclassification conditions, a weighted distance measure is proposed. The objective is to assign minimum penalty for all the images in the same class and relatively higher penalty to the other class images in the database for each correct classification obtained by the classifier. For each misclassification equal penalty is assigned to all the database images during retrieval.

Let us consider X is the query image and Y_j is the image in database having class label 'j' with their corresponding feature vectors $\vec{f}$(X) and $\vec{f}(Y_j)$ as given in Eq. (6). Then the proposed class weighted distance between two images in the feature space is defined as:

$$d_w(X, Y_j) = \frac{1}{1 + (\xi \times \mu_j(\vec{f}(X)))} \times d(X, Y_j) \tag{10}$$

where 'ξ' is a non-negative number, $\mu_j(\vec{f}(X))$ is the fuzzy membership of image X to output class 'j', and $d(X, Y_j)$ is distance between $\vec{f}(X)$ and $\vec{f}(Y_j)$ calculated using any conventional distance metric chosen suitably.

Two distance metrics such as normalised city block distance (Kokare et al., 2006; Ma and Manjunath, 1996) and d_1 distance (Murala et al., 2012) as given in Eqs. (11) and (12) are used to evaluate the performance of the proposed scheme with their simple distance based counterpart.

$$d_{cb}(X, Y_j) = \sum_k |E'_k(X) - E'_k(Y_j)| + \sum_k |\sigma'_k(X) - \sigma'_k(Y_j)| \tag{11}$$

$$d_1(X, Y_j) = \sum_k \left|\frac{E'_k(X) - E'_k(Y_j)}{1 + E'_k(X) + E'_k(Y_j)}\right| + \sum_k \left|\frac{\sigma'_k(X) - \sigma'_k(Y_j)}{1 + \sigma'_k(X) + \sigma'_k(Y_j)}\right| \tag{12}$$

Unlike (Dy et al., 2003), the proposed method makes use of the same feature set at all levels viz. classification and retrieval. It does not require obtaining class specific features to calculate similarity in the feature space during retrieval.

2.4. Selection of distance metric

For each correct classification the fuzzy membership value of corresponding output neuron will have maximum value. The class weighted distance (Eq. 10) is inversely proportional to the fuzzy membership. For each correct classification it assigns minimum penalty for all the images in the same class and relatively higher penalty to the other class images in the database. However, in case of misclassification this strategy does not work. Therefore it is essential to allow the system to use simple distance metric for similarity measure and retrieval in case of misclassification. To decide whether the network classifies the query image properly or not, the outputs of winning neuron are examined for three sets of training samples belongs to three databases (as defined in Section 4.1) used in the experiment. The probability densities of winning neuron output for three training sample sets are shown in Fig. 2.

From these probability densities it is observed that most of the winning neurons having a value 't' ($t \geqslant 0.3$) result in correct classification of the input query image. Accordingly, the rule for

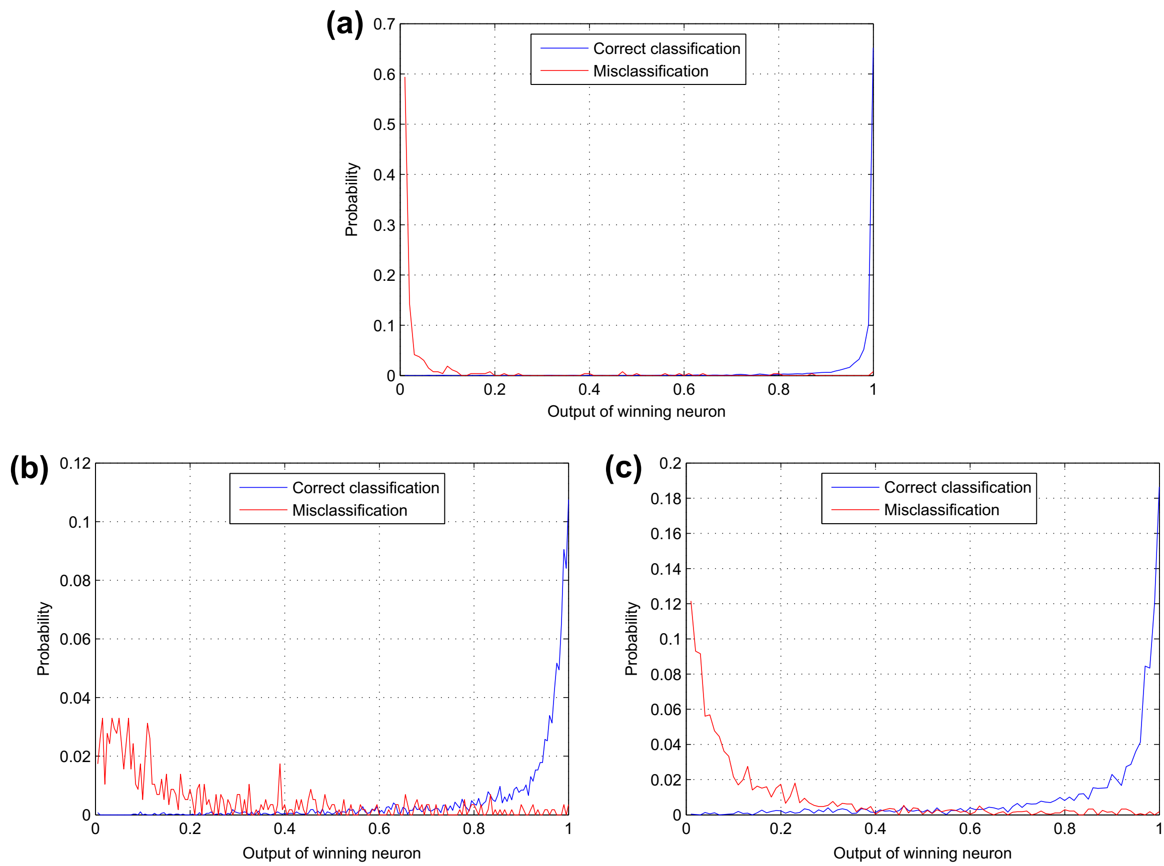

Fig. 2. Probability density of winning neuron outputs for correct and misclassification (a) training sample of D1 (b) training sample of D2 (c) training sample of D3.

selection of distance metric used for retrieval is set as follows Algorithm 1:

Algorithm 1. Takes decision on distance metric to be used

if $o_{j=winning}(\vec{f}(X)) >= t$ **then**
use class weighted distance metric,$d_w(X, Y_j)$
else
use simple distance metric,$d(X, Y_j)$
end if

The distance metric used by the proposed method adapts itself according to the confidence in classification resulting in improved performance as compared to the traditional classifier based approaches.

3. Performance metrics

Percentage precision and percentage recall (Müller et al., 2001) are used to compare the performance of the proposed CMR scheme with other competing methods.

Let Φ_R be the set of relevant images in the database and Φ_T is the set of top retrieved images for a query image. The percentage precision and percentage recall for a query image are given as

$$Precision(\%) = \frac{|\Phi_T \bigcap \Phi_R|}{|\Phi_T|} \times 100\% \tag{13}$$

$$Recall(\%) = \frac{|\Phi_T \bigcap \Phi_R|}{|\Phi_R|} \times 100\% \tag{14}$$

The efficacy of all techniques are also compared using most used measure such as precision vs. recall graph averaged for all query images (Müller et al., 2001). The complexity and computing times of different methods are also examined for practical reasons.

4. Experimental setup

4.1. Database used

Three databases of different sizes in terms of the number of texture classes, orientation and complexity are used in the experiment to evaluate the performance of the texture retrieval methods. These databases are as follows:

Database D1(Small Size Rotated): It contains 1456 images of 13 different texture classes (bark, brick, bubbles, grass, leather, pigskin, raffia, sand, straw, water, weave, wood, and wool) each of size 128×128. From the texture database (University of Southern California), 91 Bordatz Textures each of size 512×512 corresponding to 7 orientations ($0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$ and 200°) are taken. Each image is subdivided into 16 non-overlapping small images each of size 128×128 to prepare this database. There are 112 relevant images for each texture class.

Database D2 (Medium Size non-rotated): This contains 1500 texture images of 60 different texture classes from Brodatz texture photographic album (Broadtz, 1996). One image each of size 640×640 from 60 distinct classes of Brodatz texture are taken. Each image is then subdivided into 25 non-overlapping small images of size 128×128 to produce this database. There are 25 relevant images for each texture class in database D2.

Database D3 (Large Size non-rotated): This contains 2675 texture images of 107 different texture classes from Brodatz texture photographic album (Broadtz, 1996). One image each of size 640 $\times$ 640 from 107 distinct classes of Brodatz texture are taken. Each image is then subdivided into 25 non-overlapping small images of size 128 $\times$ 128 to produce this database. There are 25 relevant images for each texture class in database D3.

4.2. Texture feature sets of competing methods

The proposed CBIR paradigm (CMR) uses wavelet features for texture retrieval. Therefore it is logical to compare its performance with the conventional distance based competing methods using the same feature sets. Few techniques using wavelet features are chosen for the comparative study. All the methods reported in this paper are implemented using two distance metrics used for image retrieval in recent literature (Kokare et al., 2006; Murala et al., 2012) to demonstrate the consistency in behaviour of the proposed method irrespective of distance metric under consideration. Each image in the database is decomposed using DWT (Mallat, 1989), DT-CWT (Kingsbury, 1999; Selesnick et al., 2005) and dual tree rotated complex wavelet filters (DT-RCWF) (Kokare et al., 2006). In each case the decomposition is performed up to the third level. Three different feature sets are computed as described below:

(i) feature set 1: conventional DWT feature as explained in Section 2.1
(ii) feature set 2: rotation invariant DWT feature as proposed by Kokare et al. (2006)
(iii) feature set 3: combination of DT-CWT and DT-RCWF features as proposed by Kokare et al. (2006)

Five methods each with two variants in terms of the use of conventional distance metric (d_{cb} or d_1) are implemented for retrieval of similar images against a given query image. The methods are as follows:

(i) Method M1: It uses feature set 1 and d_{cb} (Kokare et al., 2006; Ma and Manjunath, 1996) or d_1 (Murala et al., 2012) as distance metric.
(ii) Method M2: It uses feature set 2 and d_{cb} or d_1 as distance metric (Kokare et al., 2006; Murala et al., 2012).
(iii) Method M3: It uses feature set 3 and d_{cb} or d_1 as distance metric (Kokare et al., 2006; Murala et al., 2012).
(iv) Method M4: It uses feature set 1 for classification of input image using a trained multi layer perceptron. Search space for retrieval is restricted to a single class to which input image is classified. d_{cb} or d_1 distance is used as distance metric. (Traditional classifier based approach)
(v) Proposed CMR Method: It uses feature set 1 for computation of class label and fuzzy class membership of input image for the output classes (as explained in Section 2.2). Retrieval is performed from all the output classes using Algorithm 1. d_{cb} or d_1 distance is used as simple distance metric.

Here methods M1, M2, and M3 are simple distance based approaches and methods M4 and CMR use classifier prior to retrieval.

5. Results and discussion

Experiments are performed on a Intel (R) Core (TM)2 Due 2.80 GHZ CPU with 8 GB RAM running on a Microsoft Windows 7 Professional (64-bit). For simulation MATLAB 2011b is used. For each of the database used in the experiment the multi layer perceptron with one hidden layer is trained using feature set 1 for methods M4 and CMR as discussed in Section 2.2. In each case 50% of the labeled data are used for training, 25% are used for validation and rest 25% are used for testing. A 4-fold cross validation approach is followed to evaluate the classification as well as retrieval performance. The average classification accuracy achieved using 4-fold cross validation for databases D1, D2 and D3 are 92.65%, 83.93% and 66.24% respectively. For methods M4 and CMR, features of each query image (from the test sets) are first fed to the corresponding neural network to calculate the class label and fuzzy class membership value of the query image. Retrieval is performed from the restricted search space using simple distance measure for method M4. For CMR combination of simple or weighted distance measure are used to retrieve images from the complete search space according to Algorithm 1. For other three methods (M1, M2 and M3) simple distance based retrieval is performed. To obtain the optimum value of the constant 'ξ' for weighted distance used in CMR method (in Eq. 10), the average percentage recalls are examined for different value of 'ξ'. Fig. 3(a) shows the plot for average percentage recall vs. parameter 'ξ' for both variants of CMR.

From the plot it is observed that the performance of CMR remains constant in case of all three databases beyond a value of constant 'ξ' for a particular distance metric.

5.1. Retrieval efficiency

Retrieval performances (average percentage recall) as a function of number of top retrievals considered and the standard deviation of it are listed in Table 1 for databases D1, D2 and D3. The number of relevant images (N_R) in each class for databases D1, D2 and D3 are 112, 25 and 25 respectively. The number of top retrievals considered (N_T) for database D1, D2 and D3 are 140, 32 and 32 respectively. For each database, the value of N_T considered is 25% more

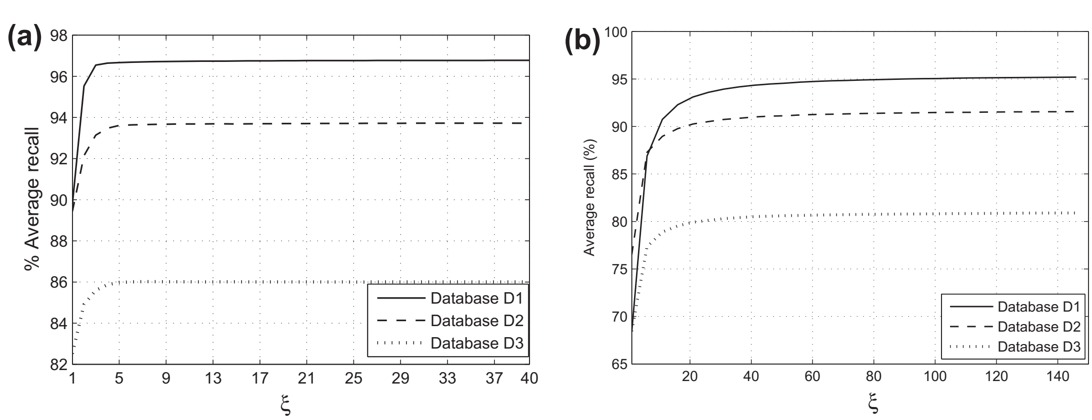

Fig. 3. Average percentage recall vs. parameter 'ξ' (a) $CMR(d_{cb})$ (b) $CMR(d_1)$.

Table 1
Average percentage recall and its standard deviation for databases D1, D2 and D3. The number of relevant images in each class for database D1, D2 and D3 are 112, 25 and 25 respectively. The number of top retrievals considered for database D1, D2 and D3 are 140, 32 and 32 respectively. The best results are highlighted with boldface.

Database	Measures	Distance based						Classifier based		
		M1		M2		M3		M4	CMR	
		d_{cb}	d_1	d_{cb}	d_1	d_{cb}	d_1	d_{cb}/d_1	d_{cb}	d_1
D1	Mean	63.93	49.35	77.61	63.50	76.74	65.82	92.65	**96.72**	95.20
	Std Dev	22.14	23.82	19.69	24.38	22.90	26.83	26.10	**12.54**	15.37
D2	Mean	74.71	59.96	64.63	54.35	57.50	53.13	83.93	**93.68**	91.58
	Std Dev	27.97	29.05	28.12	28.24	27.84	27.18	36.73	**19.38**	21.88
D3	Mean	68.94	54.97	57.15	47.89	51.63	47.84	66.24	**86.01**	80.89
	Std Dev	28.48	28.94	27.25	26.38	27.90	27.78	47.30	**25.42**	29.60

than N_R except for the method M4. For method M4, N_T equals to N_R as the method restricts the search space into single output class. The average percentage recalls of methods M1 through M4 are compared with CMR. The results for the proposed methods $CMR(d_{cb})$ and $CMR(d_1)$ are reported for the value of the constant 'ξ' equals to 10 and 150 respectively in the proposed distance metric (Eq. 10).

Following observations are made:

(i) D1 is a rotated database i.e. it contains similar textures at different orientations within each class. D2 and D3 are non-rotated databases containing similar textures at a fixed orientation within each class. The retrieval method M1 uses DWT to extract texture features which are not rotation invariant whereas the texture features obtained using methods M2 and M3 are invariant to rotation (Kokare et al., 2006). Therefore, methods M2 and M3 outperform method M1 over rotated database (e.g. D1) and underperform over non-rotated database (e.g. D2 and D3).

(ii) The proposed CMR method shows a significant improvement over retrieval accuracy in case of all databases (D1, D2 and D3) irrespective of their nature (rotated or non-rotated), when compared with methods M1, M2, M3 and M4. The proposed method makes use of an adaptive distance metric that takes advantage of both supervised (e.g. classifier based) and unsupervised (e.g. traditional distance based) retrieval methods.

(iii) The proposed CMR method improves retrieval performance from 76.74/65.82% to 96.72/95.20% on rotated image database D1, from 57.50/53.13% to 93.68/91.58% on medium size non-rotated database D2 and from 51.63/47.84% to 86.01/80.89% on very large non-rotated image database D3, compared with retrieval approach using DT-CWT and DT-RCWF features jointly (Method M3).

(iv) The retrieval performance of classifier based retrieval technique (e.g. Method M4) completely relies on the classification performance. It is clear from the Table 1 that, when classification performance is high i.e. above 90%, the conventional classifier based retrieval technique (M4) performs close to the proposed CMR scheme. However, when database contains a large number of output classes (as in case of database D2 and D3), proper classification of input images becomes difficult. Poor classification in turn affects the retrieval efficiency as shown by method M4 in Table 1 for database D2 and D3. Here it is worth noting that the classification accuracy is only 83.93%, 66.24% for database D2 and D3 respectively. For each correct classification, M4 achieves 100% retrieval accuracy as it retrieves all images from the correct class. Here all texture images within a class are similar from the point of class level retrieval accuracy. For each misclassification the class level retrieval accuracy is zero because it retrieves all images from a different class. Therefore, the retrieval accuracy obtained by method M4 equals to the classification accuracy obtained by the classifier irrespective of the type of distance metric used. The effect of distance metric comes into play when the search space is increased. This is demonstrated in Section 5.2.

(v) The retrieval performances for database D2 and D3 are found to be substantially improved by the proposed CMR method. This is achieved by allowing retrieval from complete search space (i.e. all output classes) and using the combination of weighted or simple distance measure based on Algorithm 1. The weighted distance measure depends on the class membership value of the image. For each possible correct classification it assigns minimal penalty to images of that output class for which the input image membership value is maximum and vice versa. For each possible misclassification CMR makes use of simple distance measure and thereby improves the result.

(vi) The retrieval method M4 has the highest standard deviation in percentage recall indicating fluctuation in retrieval accuracy. The standard deviation increases with increase in size of the database. The proposed CMR technique has lowest standard deviation of recall compared to other methods for all the cases.

(vii) Irrespective of the retrieval method under consideration the normalised city block distance metric (d_{cb}) has outperformed the d_1 distance metric.

The one-way analysis of variance (ANOVA) is used to determine the degree of differences between the performance of the proposed CMR method and other competing methods. It is clear from the Table 2 that there are significant differences between the mean performance of the proposed CMR method and that of other

Table 2
Pair-wise one-way ANOVA test results.

Database	Proposed	Other	p value	
			d_{cb}	d_1
D1	CMR	M1	0	0
		M2	6.70E−185	3.23E−301
		M3	1.39E−164	9.28E−238
		M4	8.91E−8	0.0013
D2	CMR	M1	3.67E−96	4.78E−211
		M2	2.06E−203	7.97E−285
		M3	7.21E−296	2.32E−311
		M4	1.81E−19	5.06E−12
D3	CMR	M1	6.93E−113	3.23E−210
		M2	4.45E−307	0
		M3	0	0
		M4	2.88E−78	2.41E−41

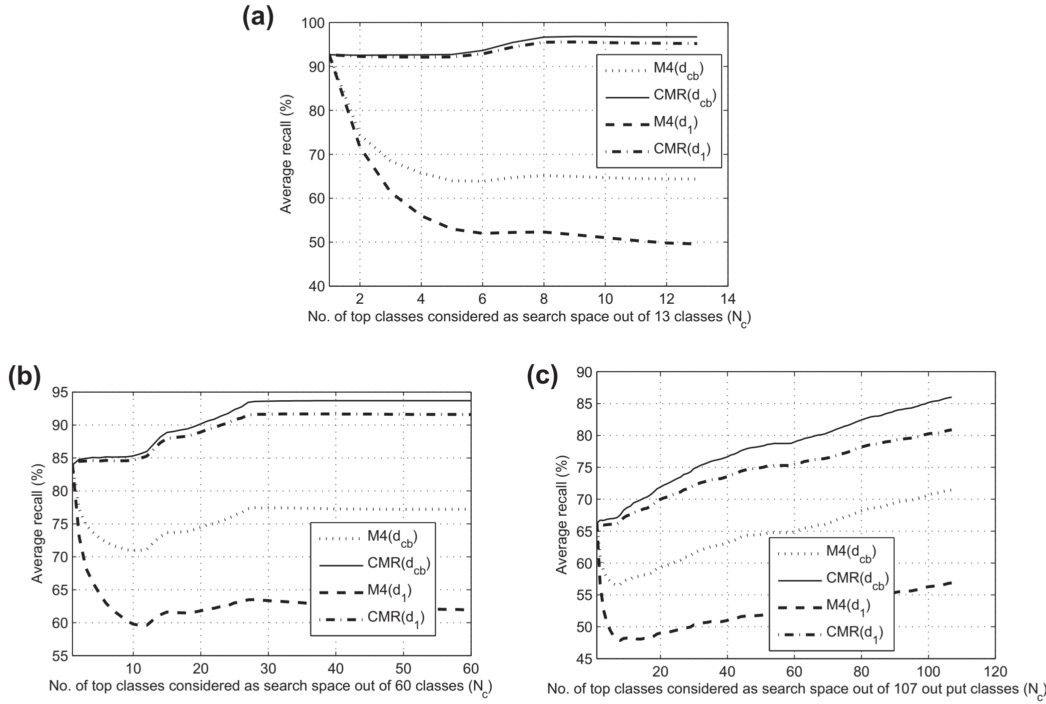

Fig. 4. Conventional classifier based retrieval vs. Class membership-based retrieval (a) database D1 (b) database D2 (c) database D3 (the number of top retrievals considered for database D1, D2 and D3 are 140, 32 and 32 respectively).

competing methods ($p < 0.05$) in all the databases used in this experiment for both the scenarios.

5.2. Retrieval effectiveness in limited search space

Experiments are conducted to demonstrate the impact of increase in the size of search space on retrieval performance. Fig. 4 shows the retrieval performance of method M4 and both the variants of CMR as a function of the number of top fuzzy membership valued classes (N_c) used as search space. These graphs show how the performance of different methods varies with respect to the size of the search space. The number of top retrievals considered for database D1, D2 and D3 are 140, 32 and 32 respectively. In all cases the number of top retrievals considered is 25% more than the number of relevant images in the database. It is clear from the above graphs that the performance of the method M4 deteriorated initially with increase in the search space, N_c. Later it improves and reaches to its local maxima. However, it never exceeds the performance of the proposed CMR method that uses the distance metric as defined in Algorithm 1 for any number of top classes considered, N_c. CMR method shows improvement with each class added to the search space. For CMR method, it is observed that the search space may be reduced in steps to increase retrieval speed at the cost of gradual reduction in accuracy. It is worth noting the point here that when N_c = 1, methods M4 and CMR resembles to each other. When N_c = total number of output classes, the method M4 resembles the method M1. The graphs in Fig. 4 show that the nature of both variants of CMR method are the same for each of the databases used in the experiment. In addition, from the results reported in Table 1, we concluded that for each method the use of normalised city block distance for retrieval achieves the best results. Henceforth, the results are reported for the proposed method and all other competing methods implemented only using normalised city block distance (d_{cb}).

5.3. Retrieval effectiveness for varying number of retrieved image

To evaluate the performance of several methods, Precision vs. Recall plots (Müller et al., 2001) are generated. All the images in each database are considered as query image. The average values obtained for the precision and recall are used to generate the plots. Fig. 5 shows this average precision vs. recall curves. A rule of thumb to read these plots is that the closer the curve to the top, the better the retrieval technique is (Traina et al., 2010). From the Fig. 5 it is clear that the proposed method achieves the best performance among all for all databases used as it pushed the curve nearest to the top.

For conventional classifier based approach (M4), the average precision equals to the classification accuracy of the classifier for any number of top retrievals (N_T) less than or equal to number of relevant images (N_R). Therefore, for M4 the precision vs. recall curve remains constant till N_T reaches N_R. Like M4, the proposed CMR takes advantage of classifier. In the method CMR, the precision vs. recall curve lies above that of M4 with the difference of a higher value of accuracy due to the adaptive nature of the distance metric used as explained in Algorithm 1.

5.4. Retrieval time

Table 3 provides dimension of feature vector used, CPU time for feature extraction, CPU time for searching similar images and computational complexity of searching similar images against a query image for each of the method used in the experiment. It can be seen from the table that, the feature extraction time for methods M1, M4 and CMR are same as they all use the identical feature sets. Method M2 takes a little more feature computing time as it involves 6 additional averaging operations (two at each level of decomposition). Method M3 requires the highest feature computation time as it uses DT-CWT as well as DT-RCWF to extract 8

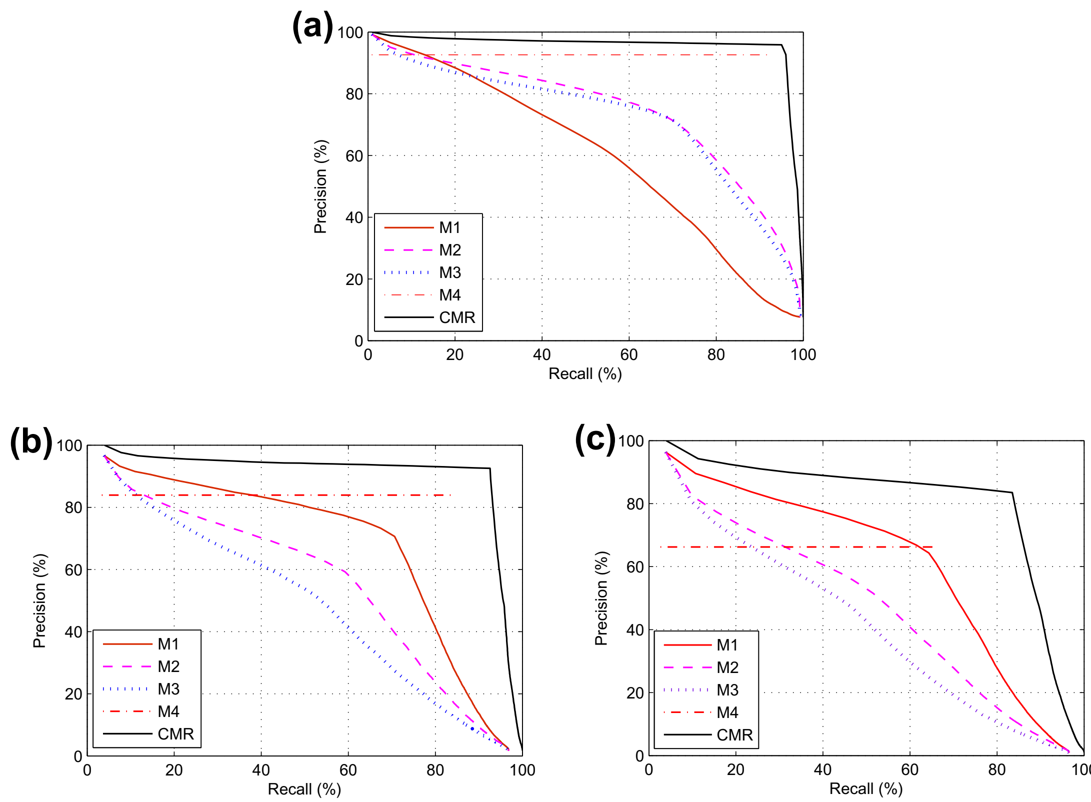

Fig. 5. Average precision vs. recall plot for proposed and competing methods using d_{cb} distance metric (a) database D1 (b) database D2 (c) database D3.

Table 3
Feature vector length, feature extraction time and searching time of query image.

	M1	M2	M3	M4	CMR
(a) Feature vector length	20	8	8	20	20
(b) Feature extraction time (in Secs)	0.007	0.008	0.08	0.007	0.007
(c) Searching time (in Secs)	0.005	0.0045	0.0045	0.0058	0.0123
Total time (b + c) (in Secs)	0.012	0.0125	0.0845	0.0128	0.0193
Computational complexity	$O(N\log N)$	$O(N\log N)$	$O(N\log N)$	$O(N\log N)$	$O(N\log N)$

dimensional feature vectors. For a database with a number N of images and number c of classes, any conventional distance based retrieval method requires a minimum of $(N\log N)$ comparisons using 'Quick Sort' for retrieving similar images against a query image and therefore has a complexity of $O(N\log N)$. The method M4 involves classification followed by retrieval from single class therefore needs $(\frac{N}{c}\log\frac{N}{c})$ number of comparisons and a constant amount of time for classification of input image to any of the output classes. The time complexity of M4 can be given as $O(N\log N)$. CMR follows the same as M4 except it retrieves images from all the out put classes. Therefore it requires $O(N\log N)$ comparisons as in case of M1 through M3. As feature extraction and classification of the query image takes constant amount of time (independent of the number of images in database), the retrieval complexity of CMR can also be given as $O(N\log N)$. However, there is differences in CPU time for searching similar images for different methods due to the use of different sizes of the feature vectors and the number of steps involved for retrieval for each method. It is obvious to observe that the searching time for methods M2 and M3 are the same as they both use 8 dimensional feature vector. Both the methods M4 and CMR use two step approach (viz. classification, searching) for retrieval and therefore found computationally expensive in searching time as compared to other methods. However, CMR takes a little more searching time than M4 as it searches in complete search space. The total retrieval time of such methods can be reduced by using feature extraction methods that are less computationally expensive as used in this experiment.

6. Conclusion

A novel approach is proposed for texture image retrieval. This approach can be used to overcome the bottleneck of simple distance based image retrieval. The approach is tested using three different databases of varying size, orientation, complexity and number of texture class. Performance of this approach is compared with other promising distance based as well as classifier based retrieval approaches. The consistency in behaviour of the proposed scheme is demonstrated using two different distance metrics applied for image retrieval in recent literature. Experimental results obtained show that the proposed CMR method achieves a significant improvement in retrieval accuracy for all databases used (D1, D2 and D3) irrespective of their nature (rotated or non-rotated). The *p values* obtained using pair wise ANOVA test of the proposed CMR method and other methods confirm the improvement. It is observed that the conventional classifier based retrieval method fails to perform when the database consists of a large number of texture classes due to poor classification accuracy. The proposed CMR method overcomes this difficulty by using a combination of simple and weighted distance measure depending on the Neural Network output of the winning neuron. Though the execution time required for the CMR method is little more as compared to the method M4 as it searches the complete search space for retrieval, it has the same computational complexity as other CBIR methods used in the experiment. In case of the proposed CMR method, it is observed that the search space may be reduced in steps at the cost of gradual reduction in accuracy to increase retrieval speed. The graphs of retrieval effectiveness as a function of the number of top classes and the number of top retrieval considered favour the proposed CMR technique.

The proposed CMR technique is evaluated using simple conventional wavelet features and adaptive city block distance or adaptive d_1 distance is taken as the distance metric. In future the performance may be improved using more sophisticated feature extraction techniques and other distance metrics.

Acknowledgement

This work has been supported by Department of Information Technology (DIT), Govt. of India. The authors are thankful to all the members of Computer Vision Laboratory, Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, India for their help.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.patrec.2013.01.001.

References

Bach, J., Fuller, C., Gupta, A., Hampapur, A., Gorowitz, B., Humphrey, R., Jain, R., Shu, C., 1996. Virage image search engine: An open framework for image management. In: Proc. SPIE, Storage and Retrieval for Image and Video Databases IV, San Jose, California, pp. 76–87.

Broadtz, P., 1996. Textures: A Photographic Album for Artists and Designers. Dover, New York.

Carkacioglu, A., Vural, F.-Y., 2002. Learning similarity space. In: Proc. ICIP.

Carson, C., Belongie, S., Greenspan, H., Malik, J., 2002. Blobworld – image segmentation using expectation-maximization and its application to image querying. IEEE Trans. Pattern Anal. Machine Intell. 24 (8), 1026–1038.

Chen, J., Bouman, C.A., Dalton, J.C., 1998. Similarity pyramids for browsing and organization of large image databases. In: Proc. SPIE/IS&T Conf. Human Vision and Electronic Imaging III, vol. 3299.

Cox, I.J., Miller, M.L., Minka, T.P., Papathomas, T.V., 2000. The Bayesian image retrieval system, PicHunter: Theory, implementation, and psychophysical experiments. IEEE Trans. Image Process. 9 (1), 20–37.

Dy, J.G., Brodley, C.E., Kak, A., Broderick, L.S., Aisen, A.M., 2003. Unsupervised feature selection applied to content based retrieval of lung images. IEEE Trans. Pattern Anal. Machine Intell. 25 (3), 373–378.

Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., Yanker, P., 1995. Query by image and video content: The QBIC system. IEEE Comput. 28 (9), 23–32.

Guo, G., Li, S.Z., Chan, K.L., 2008. Learning similarity for texture image retrieval. In: Proc. ECCV, pp. 178–190.

Guo, Z., Zhang, L., Zhang, D., 2010a. A completed modeling of local binary pattern operator for texture classification. IEEE Trans. Image Process. 19 (6), 1657–1663.

Guo, Z., Zhang, L., Zhang, D., 2010b. Rotation invariant texture classification using LBP variance with global matching. Pattern Recognition 43 (3), 706–719.

Jain, A.K., Duin, P.W., Mao, J., 2000. Statistical pattern recognition: A review. IEEE Trans. Pattern Anal. Machine Intell., 22.

Kingsbury, N.G., 1999. Image processing with complex wavelet. Philos. Trans. Roy. Soc. Lond. A. Math. Phys. Sci. 357, 2543–2560.

Kokare, M., Biswas, P.K., Chatterji, B.N., 2006. Rotation-invariant texture image retrieval using rotated complex wavelet filters. IEEE Trans. Systems Man Cybernet. – Part B Cybernet. 36 (6).

Kolmogorov, A.N., 1975. On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. Dokl. Akad. Nauk SSSR 144, 679–681; Amer. Math. Soc. Tansl. 28, 1963 55–59.

Laaksonen, J., Koskela, M., Laakso, S., Oja, E., 2000. PicSOM – content-based image retrieval with self-organizing maps. Pattern Recognition Lett. 21, 1199–1207.

Lee, K.H., 2005. First Course on Fuzzy Theory and Applications. Springer-Verlag, Berlin.

Liao, S., Law, M.W.K., Chung, A.C.S., 2009. Dominant local binary patterns for texture classification. IEEE Trans. Image Process. 18 (5), 1107–1118.

Ma, W.Y., Manjunath, B.S., 1996. Texture features and learning similarity. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition, San Francisco.

Ma, W., Manjunath, B., 1997. NeTra: A toolbox for navigating large image databases. In: Proc. IEEE Internat. Conf. Image Processing, pp. 568–571.

Mallat, S.G., 1989. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. Pattern Anal. Machine Intell. 11 (7), 674–693.

Minka, T.P., Picard, R.W., 1996. Interactive learning using a society of models. In: Proc. IEEE CVPR, pp. 447–452.

Moller, M.F., 1993. A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw. 6 (4), 525–533.

Müller, H., Müller, W., Squire, D.M., Marchand-Maillet, S., Pun, T., 2001. Performance evaluation in content based image retrieval: Overview and proposals. Pattern Recognition Lett. 22 (5), 593–601.

Murala, S., Maheshwari, R.P., Balasubramanian, R., 2012. Local tetra patterns: A new feature descriptor for content-based image retrieval. IEEE Trans. Image Process. 21 (5), 2874–2886.

Ogle, V.E., Stonebraker, M., 1995. Chabot: Retrieval from a relational database of images. IEEE Comput. 28 (9), 40–48.

Ojala, T., Pietikainen, M., Harwood, D., 1996. A comparative study of texture measures with classification based on feature distributions. Pattern Recognition 29 (1), 51–59.

Ojala, T., Pietikainen, M., Maenpaa, T., 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Machine Intell. 24 (7), 971–987.

Pentland, A., Picard, R., Sclaroff, S., 1994. Photobook: Tools for content-based manipulation of image databases. In: Proc. of the SPIE, Storage and Retrieval of Image and Video Databases II, Paper No. 2185–05, San Jose, California, pp. 34–47.

Pietikainen, M., Ojala, T., Scruggs, T., Bowyer, K.W., Jin, C., Hoffman, K., Marques, J., Jacsik, M., Worek, W., 2000. Rotational invariant texture classification using feature distributions. Pattern Recognition. 33 (1), 43–52.

Rui, Y., Huang, T.S., Chang, S.-F., 1999. Image retrieval: Current techniques, promising directions and open issues. J. Visual Comm. Image Represent. 10 (1), 1–23.

Satini, S., Jain, R., 1999. Similarity measures. IEEE Trans. Pattern Anal. Machine Intell. 29 (9).

Selesnick, I., Baraniuk, R., Kingsbury, N., 2005. The dual-tree complex wavelet transform. IEEE Signal Process. Mag. 22 (6), 123–151.

Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R., 2000. Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Machine Intell. 22 (12), 1349–1380.

Smith, J.R., Chang, S.F., 1996. VisualSEEK: A fully automated content-based image 527 query system. In: Proc. of the ACM Multimedia, Boston, MA.

Traina, Agma J.M., Traina, Caetano, Jr., Balan, André G.R., Ribeiro, Marcela X., Bugatti, Pedro H., Watanabe, Carolina Y.V., Azevedo-Marques, Paulo M., 2010. Feature extraction and selection for decision making, In: Deserno, Thomas Martin (Eds.), Biomedical Image Processing. Springer, Berlin Heidelberg, pp. 197–223.

University of Southern California, Signal and Image Processing Institute. Rotated Textures. Available from: <http://sipi.usc.edu/services/database/Databaese.html>.

Veltcamp, R., Tanase, M., 2000. Content-based image retrieval systems: A Survey. Tech. Report UU-CS-2000, 34, Dept. of Computing Science, Utrecht University.