If a proper partition function is chosen it will wholly define the behaviour of clouds to thermodynamic parameters. We quote them as follows[12]

Entropy $S = kN \log Z + \frac{3}{2} Nk,\ kN \log z + \frac{3}{2} Nk,$

Helmholtz

free energy $F = -kT \log Z,$

Total energy $E = NkT^2 \left[\frac{\delta}{\delta T} (\log z) \right]_V,$

Enthalpy $H = NkT^2 \left[\frac{\delta}{\delta t} (\log Z) \right]_V + RT,$

Gibbs potential $G = RT - NkT \log z$

Pressure $P = NkT \left[\frac{\delta}{\delta v} (\log z) + T^2 \right]_V,$

Specific heat

$$C_V = Nk \left[2T \frac{\delta}{\delta T} (\log Z) + T^2 \left(\frac{\delta^2 \log Z}{\delta T^2} \right) \right]_V.$$

1. Chaiken, P. M. and Lubernsky, T. C., *Principles of Condensed Matter Physics*, Cambridge University Press, 1999.
2. http://earth observatory.nasa.gov/fa/clouds/printall.php
3. Earth radiation B data on the Earth radiation budget satellite and NOAA-9 processed at NASA.
4. Mani, A. and Rangarajan, S., *Solar Radiations over India*, Allied Publishers, India, 1982.
5. Gufee, K. and Henderson-Sellors, A., *A Climate Modeling Primer*, John Wiley, 1997.
6. Ema, K. A., *Annu. Rev. Fluid Mech.*, 1991, **23**, 179–196.
7. Laurence, M. G., The relationship between relative humidity and the dew point temperature in moist air: a simple conversion and applications. *Bull. Am. Meteorol. Soc.*, 2005, **86**, 225–233.
8. Kondratyev, K. K., *Radiations in the Atmosphere*, The International Geophysics Series, Academic Press, 1969, vol. 12.
9. Dufour, L. and Defay, R., *Thermodynamics of Clouds*, International Geophysics Series, Academic Press, 1963.
10. Landalu, L. D. and Lifshitz, E. M., *Statistical Physics*, Pergamon Press, 1975, vol. 5.
11. Agrawal, B. K. and Eisner, M., *Statistical Mechanics*, New Age International (P) Limited, 1994.
12. Brij, L., Subramaniam, N. and Hemne, P. S., *Heat, Thermodynamics and Statistical Physics*, S. Chand Publisher, 2008.

Received 22 March 2013; revised accepted 18 September 2013

Improved content-based classification and retrieval of images using support vector machine

V. Karpagam[1,]* and R. Rangarajan[2]

[1]Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore 641 022, India
[2]Indus College of Engineering, Coimbatore 641 101, India

Content-based image retrieval (CBIR) entails probing for similar images for a query image in an image database and returning the most relevant images. The proposed methodology aims at improving the classification and retrieval accuracy of images. Wavelet histograms are used to design a simple and efficient CBIR system with good performance and without using any intensive image-processing feature extraction technique. The unique indexed colour histogram and wavelet decomposition-based horizontal, vertical and diagonal image attributes serve as the main features for the retrieval system. Support vector machine is used for classification and thereby to improve retrieval accuracy of the system. The performance of the proposed content-based image classification and retrieval system is evaluated with the standard SIMPLIcity dataset. Precision is used as a metric to measure the performance of the system. The system is validated with holdout and *k*-fold cross-validation techniques. The proposed system performs better than SIMPLIcity and all the other compared methods.

Keywords: Colour image representation, discrete wavelet decomposition, image classification, image feature extraction.

CONTENT-based image retrieval (CBIR) finds application in a number of areas like video surveillance, medicine and geographic information system (GIS). All these applications require a high degree of accuracy with minimal user involvement. A myriad of CBIR engines have been proposed in the literature. Though most of the methods perform significantly well, the semantic gap remains to be bridged. Most of the popular methods involve region-based techniques which are computationally intensive and the success of the methods depends on the segmentation techniques used. Many relevance-based techniques have also been proposed, but the retrieved results may depend on individual perception of relevance. This spawns the need for a simple and efficient retrieval system with no user involvement.

There are several methods being used for the retrieval of images based on visual features such as colour, texture and shape. Most of the successful methods use sophisticated, time-consuming image processing techniques to

*For correspondence. (e-mail: karpagam11@yahoo.com)

learn the semantic content of the image. For example, if we want to study the separate regions of the image, then suitable colour or texture segmentation algorithms should be used to separate the homogeneous regions for further analysis to classify them based on the features. Even after such sophisticated semantic analysis, the improvements in the results were not so significant. Further, simple image-matching policies often lead to poor accuracy in image retrieval. So, a CBIR system with model-based classification technique may lead to better results. Lu and Weng[1] present a comprehensive survey of advanced image classification techniques. The image processing procedure and selection of suitable classification method has been stated to play a significant role in improving the classification accuracy.

Here, a simple CBIR system is modelled which will use features that can be acquired from the image in a fast mode. An attempt is made to show that the accuracy of a simple CBIR system can be made competitively equal to that of a sophisticated CBIR system, if the simple and more significant features of the image are chosen for coding in the image feature dataset. To improve the image retrieval accuracy, a support vector machine (SVM)-based classification model is used. The development process for this project involves three phases – image feature extraction, training the SVM network and matching the query image with database images using the previously trained network. For the first phase, wavelet histograms (WH) are used[2]. For the second phase, SVM neural network is used and for the retrieval part, the previously trained network as well as simple distance measure are used.

A CBIR system based on a multiscale geometric analysis (MGA) tool, called ripplet transform type-I (RT) has been presented by Chowdhury *et al.*[3]. The scheme utilizes a neural network-based pre-classifier for the images to improve the retrieval accuracy. Similarity is measured using Manhattan distance and fuzzy entropy-based feature is used for relevance feedback. The performance of the CBIR system has been evaluated using a 2×5-fold cross validation followed by a statistical analysis.

Deselaers *et al.*[4] have conducted experiments on five different publicly available image databases and have analysed the retrieval performance of the features. The correlation of the features has also been analysed and it has been concluded that the frequently used, but simple, colour histogram performs well in many applications.

An image retrieval scheme which makes use of visually significant point features has been presented earlier[5]. The clusters of points around significant curvature regions are extracted using a fuzzy set theoretic approach. Some invariant colour features are computed from these points to evaluate the similarity between images. A set of relevant and non-redundant features is selected using the mutual information-based minimum redundancy–maximum relevance framework. The relative importance of each feature is evaluated using a fuzzy entropy-based measure. Salient points determination is based on colour saliency[6]. The combination of the local colour, texture and the global shape features have been used to provide a robust feature set for image retrieval. SIMPLIcity makes use of a wavelet-based approach for feature extraction, and integrated region matching based upon image segmentation[7]. The image is represented by a set of regions, roughly corresponding to objects characterized by colour, texture, shape and location. The system classifies images into semantic categories enhancing the retrieval. The system is fairly robust to image alterations.

Fuzzy SVM network makes use of a SVM with a kernel comprised of fuzzy basis functions[8]. This network has been evaluated using images from the aceMedia Repository1, specifically for the beach/urban scenes classification problem. A good classification rate has been achieved while fusing edge histogram, colour layout and scalable colour MPEG-7 descriptors. The performances of SVM, KNN and logistic regression classifiers on different categories of images have been compared[9]. While both KNN and SVM perform better than logistic regression, the choice of k is essential for the success of the algorithm. Another major problem with KNN is that the class with the more frequent training samples would dominate the prediction of the new vector. An integrated many artificial neural networks (ANNs) and one SVM for classification and recognition method for Roman numeral images has been proposed[10]. Though the precision of the two-layer classifier is high, the processing time needs to be improved for complex image classification applications.

In image recognition or pattern recognition in general, the two major issues are feature extraction and distance measure definition. Failure in either of the two issues will lead to poor performance of the recognition system. There is no exception to the CBIR system. In this work, we propose the use of a unique colour histogram and wavelet decomposition-based horizontal, vertical and diagonal image attributes as the main features and design the retrieval system.

Generally, the three colour layers information of a typical RGB image is handled separately in the feature extraction process. This may lead to inaccurate representation of the colour features. For example, the same level of one particular layer colour may create different colours in different parts of the image, since the colours of the other two layers will decide the colour of the pixel. Several previous works handle the three layers separately and use separate histograms to measure the colour features.

In the study by Nandagopalan *et al.*[11], a single feature vector represents the colour, texture and edge information. The image is segmented. The mean, median and standard deviation of the red, green and blue channels of the colour histograms are computed for the dominant segments. Texture co-occurrence matrix-based entropy and energy are calculated and for edge density the edge histogram descriptor (EHD) is found.

Laplacian transform of the sharpened greyscale image is statistically quantized into colour histogram bins in Malik and Baharudin[12]. It has been shown that the category-wise precision for 32 bins was the best value achieved using the method. Patil and Kolhe[13] have segmented every image using *k*-means clustering to identify regions. Then, the features of every region are extracted to form the feature vector. KNN multiple instance learning is carried out to classify and annotate the unlabelled image object.

In the proposed method, the RGB image is converted to an indexed image with low level of colour detail. All the 16,777,216 possible colours of the RGB space are not perceivable by the human eye. Therefore, the 24 bit colour image is quantized to a 256-colour-indexed colour image. A common operation that reduces the size of large 24 bit bitmaps is to convert them to indexed colour with an optimized palette, that is, a palette which best represents the colours available in the bitmap. The colour map of only one image of the whole dataset is stored separately to decompose the remaining images of the dataset. A colour approximation method is used to do the colour mapping and the images will be almost in their original colour.

Indexed colour is an economical way of storing colour bitmaps without using 3 bytes per pixel. As with 8-bit grey bitmaps, each pixel has a byte associated with it. Only now the value in the byte is no longer a colour value, but an index into a table of colours called a palette or colour table. Figure 1 shows the colour map.

There are a number of interesting attributes of such a colour indexing system. If there are less than 256 colours in the image then this bitmap will have the same quality as a 24 bit bitmap, but it can be stored with one-third the data. Interesting colouring and animation effects can be achieved by simply modifying the palette. This immediately changes the appearance of the bitmap and with careful design can lead to intentional changes in the visual appearance of the bitmap.

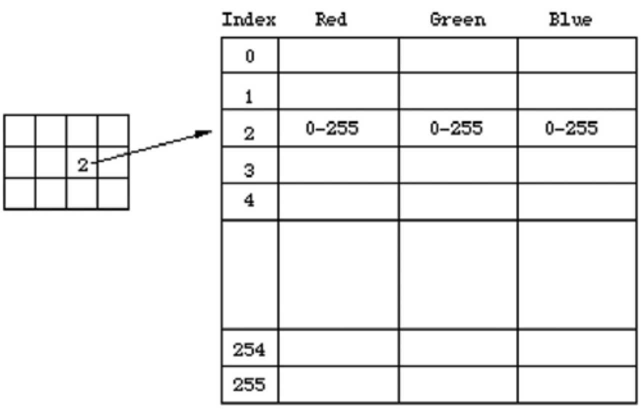

Figure 1. Colour map.

After this decomposition, each indexed 8-bit pixel will represent a particular colour which is stored separately as a map. For example, if the colour red (255, 0, 0) is indexed with a number 78, then all the indexed pixels with value 78 will represent the same red colour (255, 0, 0). So, now the value 78 indicates red and the value 78 in all the indexed images of the whole dataset will represent red and red only. Now it becomes possible to represent the colour distribution of the image with a single histogram in which the bin 78 will just represent the count of the red colour indexed pixels.

The colour indexed histogram represents the colour distribution of the image, but it excludes the shape information. Two images with the same colour distribution may be semantically different, for example, a red colour car and a red colour ball. Therefore, there is a necessity to combine the power of colour indexed histograms with a method which can extract the shape and texture features from an image.

Wavelet transform has become popular in different fields and is often used for analysis, de-noising and compression of signals and images. The resultant images of single-level two-dimensional wavelet decomposition have several interesting characteristics. Generally, the 2D wavelet decomposition will produce four output images, L1, H1, V1 and D1. The MATLAB implementation of two-dimensional dwt function (dwt2) computes the approximation coefficients matrix L1 and details coefficients matrices H1, V1, D1, obtained by a wavelet decomposition of the input image matrix.

The feature dataset is nothing but a feature-based index that will represent the whole image data in a most simplified form. These features will reflect the content of the image. A recent trend for image search is to fuse two basic modalities of the images. The key issue is how to fuse the two modalities to represent the image. Information from multiple sources is expected to increase the robustness and retrieval accuracy of the system through redundancy. Different levels of fusion have been identified, namely pixel, fusion and decision. Feature-level fusion is performed here. Fusion can also be categorized as early fusion or delayed fusion. Numerous methods of classification use an early fusion of features by concatenating the descriptors extracted from one image in one single vector. An advantage of this approach is the use of a simple fusion model involving low-cost computing compared to a late fusion model. Early fusion of the features is proposed in this work. The final feature set obtained is a simple concatenation of the colour features and texture features. The feature set is a vector of 263 dimensions. The first 256 values represent the frequency of the histogram bins, for the 256 colours present in the colour map. The wavelet coefficients form the next four features in the vector. The next three values represent the traditional RGB colour histogram. The structure of the final feature vector is shown in Figure 2.

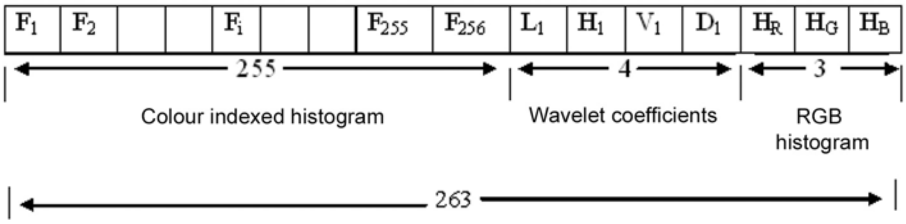

Figure 2. Structure of the wavelet histogram feature vector.

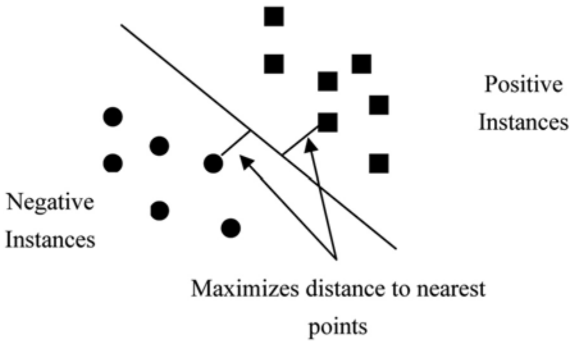

Figure 3. Data classification using support vector machine.

For matching the input image features with the stored features of the image dataset, the simple Euclidean distance is used as a distance metric. The rank of the matching images was calculated based on the Euclidean distance with the query image. The image with minimum Euclidean distance is started with rank 1 and all the images were ranked in an increasing order based on the Euclidean distance. In our evaluations, we only considered the top 50 ranked matching images and calculated the precision by taking the average of precision of several runs with same category input query images.

The SVM approach is considered a good candidate for classification because of its high generalization performance without the need to add a priori knowledge, even when the dimension of the input space is very high[14,15]. Although the SVM can be applied to various optimization problems such as regression, the classic problem is that of data classification. SVM has been largely used in CBIR as a learning technique based on relevance feedback. All these methods pose the limitation that it requires the users' involvement. The decisions made by the system will be tailored to the needs of individuals and may not be applicable generally.

The basic idea of SVM is shown in Figure 3. The data points are identified as being positive or negative, and the problem is to find a hyper-plane that separates the data points by a maximal margin. Figure 3 only shows the two-dimensional case where the data points are linearly separable.

The mathematics of the problem to be solved is the following

$$\min_{\vec{w},b} \frac{1}{2} \| w \|,$$

$$\text{s.t. } y_i = +1 \Rightarrow \vec{w}\cdot\vec{x}_i + b \geq +1,$$

$$\text{s.t. } y_i = -1 \Rightarrow \vec{w}\cdot\vec{x}_i - b \leq -1,$$

$$\text{s.t. } y_i (\vec{w}\cdot\vec{x}_i + b) \geq 1, \ \forall i. \tag{1}$$

The identification of each data point x_i is y_i, which can take a value of +1 or –1 (representing positive or negative respectively). The solution hyper-plane is the following

$$u = \vec{w} \cdot \vec{x} + b. \tag{2}$$

The scalar b is also termed the bias. A standard method to solve this problem is to apply the theory of Lagrange to convert it to a dual Lagrangian problem. The dual problem is the following

$$\min_{\alpha} \Psi(\vec{\alpha}) = \min_{\alpha} \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} y_i y_j (\vec{x}_i \cdot \vec{x}_j) \alpha_i \alpha_j - \sum_{i=1}^{N} \alpha_i,$$

$$\sum_{i=1}^{N} \alpha_i y_i = 0, \quad \alpha_i \geq 0, \ \forall i. \tag{3}$$

The variables α_i are the Lagrangian multipliers for corresponding data point x_i.

The proposed CBIR system is implemented and evaluated in LIBSVM toolbox[16] for MATLAB.

SVM solves non-separable feature vectors by relaxing the constraints of the hyper-plane. A cost function is added into the separating margin regions. In most practical applications where classification is required, the data to be separated are not linearly separable. A solution that ignores the few weird instances would be preferable. Therefore, an instance variable x_i is allowed to exist without meeting the margin requirement, at a cost proportional to a slack variable ζ_i.

The steps of the proposed CBIR system using SVM include:

- Constructing a SVM model of a classification network using the colour indexed image histogram features

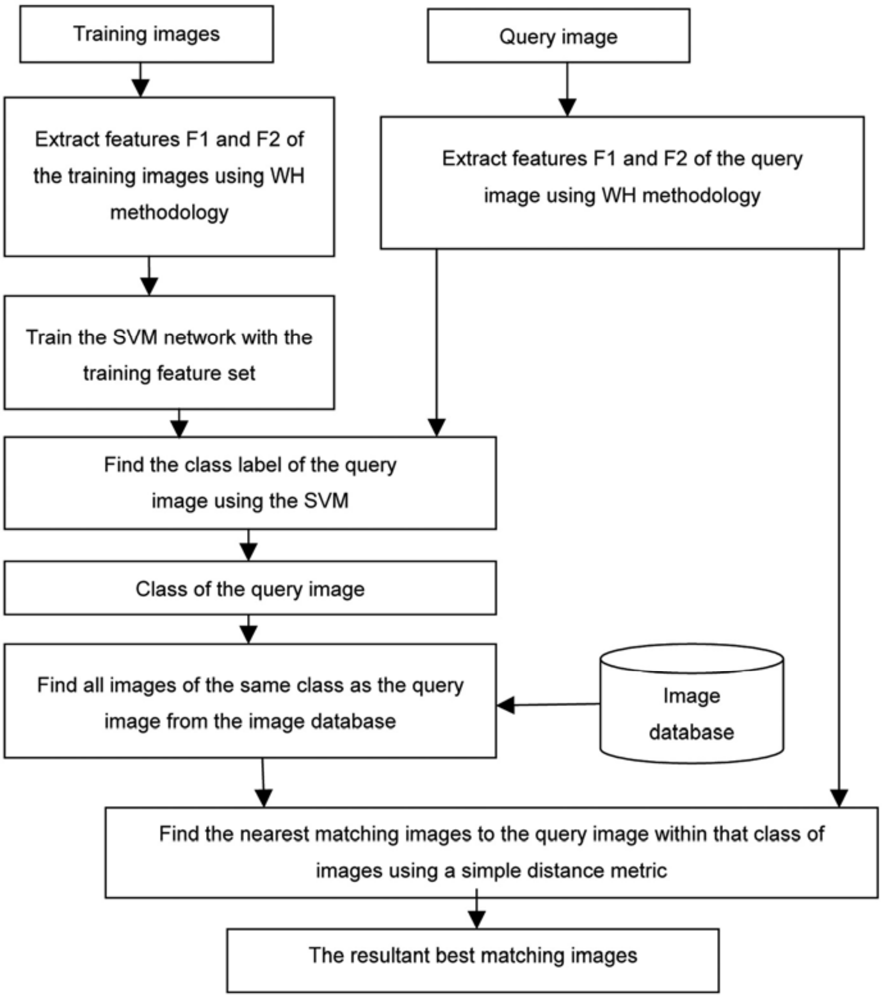

Figure 4. Model of the proposed SVM-based CBIR system.

and discrete wavelet decomposition of the training images.

- Classifying the input image using the trained model.
- Retrieving all the best matching images from the matching class of the input image using a simple distance metric.

Figure 4 explains the proposed SVM-based CBIR system.

Supervised classification has been used in this research to categorize the images as the SIMPLIcity dataset contains images with well-defined labels. Input images to the supervised classifier are labelled as Africa, beach, buildings, buses, dinosaurs, elephants, flowers, horses, mountains and food. The feature descriptors of the images, in the WH format are given to the supervised classifier to infer a rule that assigns a label to each image. The classifier assigns a class label to the output value y, that best matches the given input pattern and is denoted by Ck, $k = 1, 2, \ldots, K$, where K is the number of classes. Here, $K = 10$. The input data source is the feature descriptor obtained by fusing the colour indexed histogram and the output of the wavelet transformation. A vector of 263 real numbers forms the input vector denoted by $\overline{x}$.

The vector format of the image features is suitable for an SVM classifier. The ability of a supervised classifier to map an input vector $\overline{x}$ to the desired output class y is based on the performance of the learning algorithm. Image classification is cast as learning a prediction

$$y \approx H(\overline{x}), \tag{4}$$

that maps the space of images X to the space of classes Y, based on a training set of input/output pairs.

There are two approaches to minimize the mapping error in eq. (4), namely empirical risk minimization and structural risk minimization. The appeal of SVM is that it uses structural risk minimization. SVM attempts to map the correct intensities of the colour indexed histogram and the energy output of the wavelet transformation that will prove to be the most probable match between the input vector and the correct class.

SVM is susceptible to scaling of data. Therefore, it is necessary to normalize the input vector. Min–max normalization is carried out on the WH feature vector. This preserves the relationships among the original data values. Min–max normalization is the process of transforming data to a value between 0.0 and 1.0. The minimum value is set to 0.0 and the maximum is set to 1.0. This provides an easy way to compare values that are measured using different scales. Normalization is given by eq. (5)

$$\text{Normalized value} = \text{Minimum value}/\text{Range of values}. \tag{5}$$

SVM which was originally designed for binary classification has been extended to support multi-class classification through one-against-all (1AA) and one-against-one (1A1) strategies. The 1A1 strategy decomposes the multiclass problem into a set of binary classifiers. For n number of output classes, $n*(n-1)/2$ classifiers are constructed and each one is trained with data from two classes i and j. A separate decision boundary is independently created between every pair of classes. Equation (1) is transformed to the form given in eq. (6):

$$\text{sign}((w^{ij})^T \phi(x) + b^{ij}), \tag{6}$$

where $\phi(x)$ represents the images mapped in high-dimensional feature space, and i and j represent the corresponding pairs of classes.

The voting strategy called as max–wins approach is followed to predict the class of the WH feature vector. If eq. (6) predicts x to belong to the ith class, the vote for i is incremented, else the vote for j is incremented. Then, x is predicted as belonging to the class with the maximum number of votes.

The default parameters of LIBSVM are used. Experiments conducted with different parameters show that the performance of SVM with the default parameters is found to be good. Only the following parameters are changed.

$-s$ 1 $-t$ 3 $-g$ 0.1 $-e$ 0.01 $-p$ 0.0000000001.

Choice of parameter-s indicates that nu-SVC is used. The range of nu in nu-SVC is 0 to 1, upper bounded by the fraction of outliers and lower bounded by the fraction of support vectors. This value decides whether a hard margin or soft margin separation is performed. The default value of 0.5 is chosen, leading to soft margin SVM.

The choice of the kernel function is a decisive factor in the performance of the SVM classifier. Kernels provide a simple bridge from linearity to nonlinearity in algorithms. Sigmoid kernel is used in this research. The hyperbolic tangent kernel is also known as the sigmoid kernel and as the MLP kernel. The sigmoid kernel comes from the neural networks field, where the bipolar sigmoid function is often used as activation for artificial neurons.

$$k(x_i, x_j) = \tan h(\text{gamma}*u'*v + \text{coef0}), \tag{7}$$

where –g 0.1 – gamma. coef0 is the coefficient for the Kernel and the default value is 0.

An SVM model using a sigmoid kernel function is equivalent to a two-layer perceptron neural network. It has been found to perform well in practice.

There are different metrics for evaluating the performance of a typical CBIR system. We selected precision as the main evaluation metric because it was used in several previous works[2,5–7].

The quantitative measure defined is average precision and is explained as

$$p(i) = \frac{1}{n}\left(\sum_{l \le j \le N, r(i, j) \le n, C(i)=C(j)} 1\right). \tag{8}$$

This value is the percentile of images belonging to the category of image i in the first n retrieved images, where $p(i)$ is precision of query image i, n the number of images in each category. In this case, there are 100 images in each category and so $n = 100$. N is the total number of images in the database. In this case, $N = 1000$. $C(i)$ and $C(j)$ are category ID of images i and j respectively. In this experiment, the category ID will be in the range 1–10. $r(i, j)$ is the rank of image j (i.e. position of image j in the retrieved images for query image i, an integer between 1 and N).

The average precision p_c for category $C(1 \le c \le n)$ is given by

$$p_c = \frac{1}{n}\left(\sum_{l \le j \le N, C(i)=c} 1\right). \tag{9}$$

The holdout method for cross-validation divides the data into two disjoint subsets, namely training and test sets. The training set is used to train the classifier. The test set is used to estimate the error rate of the trained classifier. Holdout validation is performed by randomly selecting 90% of images for training the classification model and the remaining 10% of the images for testing the model. The validation is repeated five times and the average is calculated.

In this work k-fold cross-validation is chosen as the main metric for evaluating the performance of the image classification system. Also, ten-fold cross-validation is applied for evaluating the performance of the classifiers.

The k-fold validation creates a k-fold partition of the dataset. The experiment for model selection is repeated k times, with $k-1$ folds forming the training dataset and the remaining one-fold forming the test dataset. k-fold cross-validation has the advantage that all instances in the dataset are eventually used for both training and testing. The ten-fold cross-validation is performed in this work.

The results are benchmarked with standard systems, viz. SIMPLIcity[7], FIRM[17] and other previous works using the same database. We compared our earlier results[2] with those of the proposed method. The performance of the system is measured with more accuracy. Two validation methods, viz. holdout validation and *k*-fold cross-validation are used. The category-wise precision is measured and tabulated. The average of precision in all the categories is considered as the overall precision of the CBIR system. Table 1 shows the results of some earlier CBIR systems which are compared with the proposed CBIR system.

Generally, during *k*-fold validation, the value of *k* is taken as 10. Therefore, 90% of data will be used for the construction of the network model and the remaining 10% for validating the model.

Figure 5 shows the 3D plot of all the images classified using the SVM-based classifier. This virtual image space shows the clusters of images which belong to different categories. Each colour legend in this plot shows one category of image in the database. If we see the class 'dinosaurs' (black dots at the right top), it forms a distinct cluster in this image space. That is why the precision of classification of dinosaurs is high compared to the other classes.

Table 2 shows the results of the proposed CBIR systems. The first column of Table 2 shows the result of WH method[2]. The second and third are the results of the proposed WHSVM model. The precision of the proposed model is presented through holdout and *k*-fold validation.

The performance of the proposed CBIR model in terms of precision is significantly higher compared to the earlier works, as shown in Tables 1 and 2, and Figure 6.

Figure 7 shows the comparison of category-wise performance of the proposed CBIR systems WHSVM and WH[2]. The proposed WHSVM model is proficient in finding matching images from the database with more accuracy in almost all the categories of images. The category-wise precision was found to be good during both holdout validation as well as *k*-fold validation. This proves that there is significant improvement in the accuracy of the proposed SVM-based CBIR system.

Figure 8 shows the category-wise comparison of performance of the earlier methods and the proposed methods.

Table 1. The results of other previous methods

Class	FEI	SIMPLIcity	Simple Hist	FIRM
Africa	0.45	0.48	0.30	0.47
Beach	0.35	0.32	0.30	0.35
Buildings	0.35	0.35	0.25	0.35
Buses	0.60	0.36	0.26	0.60
Dinosaurs	0.95	0.95	0.90	0.95
Elephants	0.60	0.38	0.36	0.25
Flower	0.65	0.42	0.40	0.65
Horses	0.70	0.72	0.38	0.65
Mountains	0.40	0.35	0.25	0.30
Food	0.40	0.38	0.20	0.48
Average	0.545	0.471	0.36	0.505

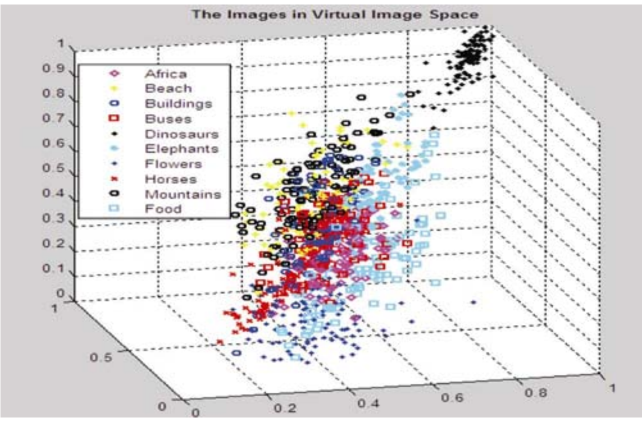

Figure 5. The virtual image space.

Table 2. The results of proposed methods

	Precision at *n* (calculated using *n* topmost results)		
Class	Proposed method	Proposed WHSVM (holdout)	Proposed WHSVM (*k*-fold)
Africa	0.69	0.76	0.82
Beach	0.35	0.73	0.76
Buildings	0.46	0.66	0.74
Buses	0.62	0.86	0.75
Dinosaurs	0.98	1.00	0.98
Elephants	0.38	0.78	0.73
Flower	0.60	0.90	0.98
Horses	0.91	0.88	0.89
Mountains	0.44	0.72	0.88
Food	0.58	0.90	0.69
Average	0.60	0.82	0.82

WHSVM, Wavelet Histogram with Support Vector Machine.

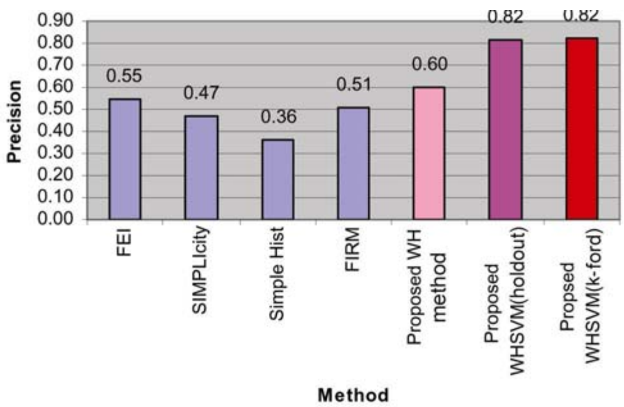

Figure 6. Performance in terms of precision.

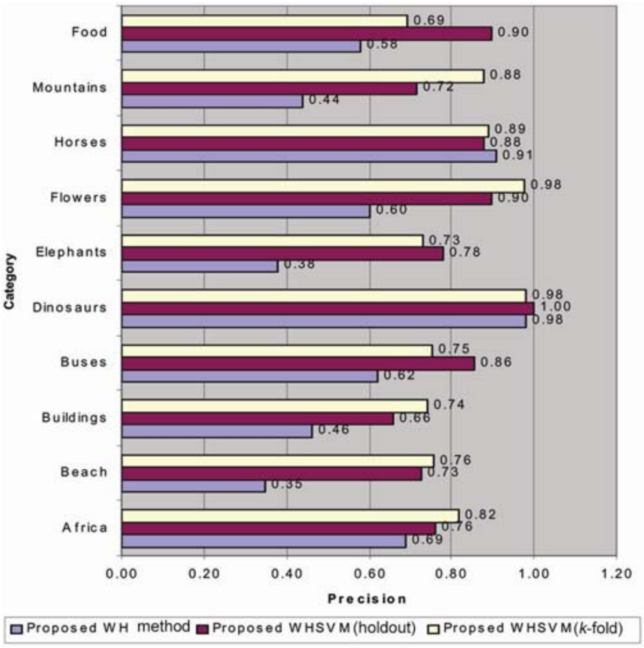

Figure 7. Category-wise performance of proposed method.

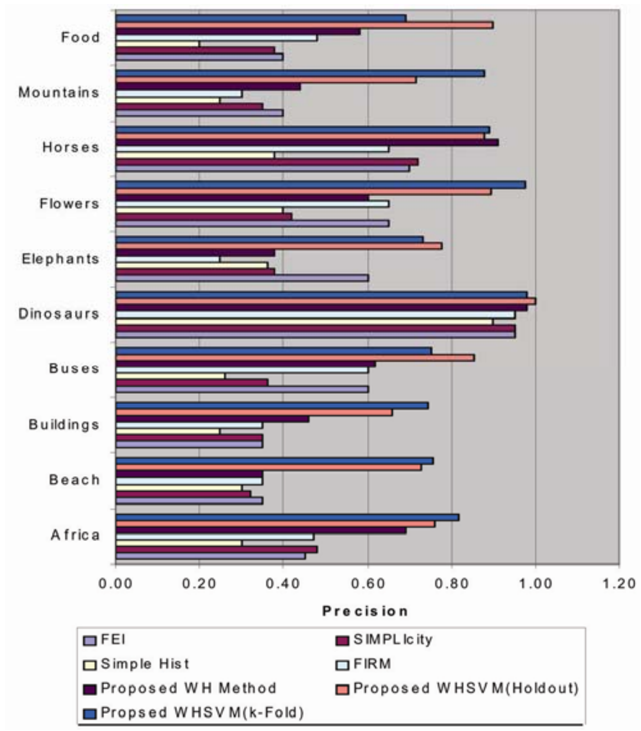

Figure 8. Category-wise comparison of proposed methods with earlier methods.

This fulfills the original research objective. It has been proved that the performance of a simple features-based CBIR system can be made better than most of the sophisticated texture and shape-based CBIR systems, if the simple and distinct features are selected from the image and a suitable classification model is employed.

A CBIR system has been implemented successfully using the SVM neural network-based classification and retrieval model. For constructing the model the simple colour indexed histogram and wavelet features are used[2]. The performance of the proposed CBIR system has been evaluated with standard SIMPLIcity dataset and has been compared with the results of some previous studies. The precision of the proposed system measured was found to be good and the proposed model was able to compete with all the compared models. Most of the earlier models produced almost equal or poor results even with the aid of sophisticated region, shape and texture matching techniques. But the proposed model provided excellent performance with simple features and a simple classification model. Hence we have shown here the possibility of a better CBIR system with more simple and significant feature sets. It has been observed that the cause of error in finding the correct match is the nature of some images which may belong to more than one category. So, our future work will address and solve this problem using a multi-modal system or with a multi-class classification technique. In this study, we have directly used the feature set which has more than 256 attributes. Future studies may address more sophisticated statistical techniques such as PCA and LDA to select principal features/attributes from the whole attributes.

1. Lu, D. and Weng, Q., A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sensing*, 2007, **28**, 823–870.
2. Karpagam, V. and Rangarajan, R., A simple and competent system for content based retrieval of images using color indexed image histogram combined with discrete wavelet decomposition. *Eur. J. Sci. Res.*, 2012, **73**, 278–290.
3. Chowdhury, M., Das, S. and Kundu, M. K., Novel CBIR system based on ripplet transform using interactive neuro-fuzzy technique. *Electron. Lett. Computer Vision Image Anal.*, 2012, **11**, 1–13.
4. Deselaers, T., Keysers, D. and Ney, H., Features for image retrieval: an experimental comparison. *Inf. Retrieval*, 2008, **11**, 77–107.
5. Banerjee, M., Kundua, M. K. and Maji, P., Content-based image retrieval using visually significant point features. *Fuzzy Sets Syst.*, 2009, **160**, 3323–3341.
6. Hiremath, P. S. and Pujari, J., Content based image retrieval using color boosted salient points and shape features of an image. *Int. J. Image Process.*, 2008, **2**, 10–17.
7. Wang, J. Z., Li, J. and Wiederhold, G., SIMPLIcity: semantics-sensitive integrated matching for picture libraries. *IEEE Trans. Pattern Anal. Mach. Intelli.*, 2001, **23**, 947–963.
8. Spyrou, E., Stamou, G., Avrithis, Y. and Kollias, S., Fuzzy support vector machines for image classification fusing MPEG-7 visual descriptors. In Proceeding of Integration of Knowledge, Semantics and Digital Media Technology, 2005. EWIMT 2005. The 2nd European Workshop on the (Ref. No. 2005/11099), ISBN: 0-86341-595-4.
9. Gu, Q. and Song, Z., Image classification using SVM, KNN and performance comparison with logistic regression, www.cs.dartmouth.edu/~afra/courses/44/w09/.../gu-song-report.pdf, 2009.
10. Thai, L. E., Hai, T. S. and Thuy, N. T., Image classification using support vector machine and artificial neural network. *Int. J. Infor. Technol. Comput. Sci.*, 2012, **5**, 32–38.

11. Nandagopalan, S., Adiga, B. S. and Deepak, N., A universal model for content-based image retrieval. *Int. J. Electr. Comput. Eng.*, 2009, **4**.
12. Malik, F. and Baharudin, B. B., Feature analysis of quantized histogram color features for content-based image retrieval based on Laplacian filter. In International Conference on System Engineering and Modeling. IACSIT Press, Singapore, 2012, vol. 34.
13. Patil, M. P. and Kolhe, S. R., Automatic image categorization and annotation using K-NN for corel dataset. *Adv. Comput. Res.*, **4**, 108–112.
14. Chapelle, O., Haffner, P. and Vapnik, V. N., Support vector machines for histogram-based image classification. *IEEE Trans. Neural Networks*, 1999, **10**, 1055–1064.
15. Chen, Y. and Wang, J. Z., A region-based fuzzy feature matching approach to content-based image retrieval. *IEEE Trans. Pattern Anal. Mach. Intelli.*, 2002, **24**, 1252–1267.
16. Chang, C.-C. and Lin, C.-J., LIBSVM: a library for support vector machines, 2001; Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm
17. Li, J., Wang, J. Z. and Wiederhold, G., IRM: integrated region matching for image retrieval. Proceedings of the 8th ACM International Conference on Multimedia, Los Angeles, CA, USA, 2000, pp. 147–156.

Received 18 January 2013; revised accepted 28 August 2013

Application of fractal geometry in determining optimal quadrat size for vegetation sampling

Anubhav Chaturvedi[1,]* and P. Rama Chandra Prasad[2]

[1]Centre for Computational, Natural Sciences and Bioinformatics, and
[2]Lab for Spatial Informatics, International Institute of Information Technology, Gachibowli, Hyderabad 500 032, India

Geometry in ecological patterns of landscape and vegetation is not truly fractal, and varies across a range of scales, whereas fractal geometry provides tools for predicting and describing ecological patterns. In this study, fractal analysis is used to assess presence of pseudo random quadrats or spatial dependence which hamper generality and performance of classical inferential statistics. Fractal dimension (FD) as a function of scale is used to determine quadrat size which eliminates spatial dependence. The semivariograms are plotted with fractograms to correlate structures of spatial dependence of the properties studied. The use of FD as a degree of spatial dependence of variables is the basis of applications of fractals.

Keywords: Ecological patterns, fractal geometry, quadrat size, spatial dependence, vegetation sampling.

A significant challenge encountered in plant ecological studies is vegetation sampling[1,2]. Researchers worldwide have analysed ecological attributes (species diversity, richness, dominance, etc.) of vegetation using random or stratified random sampling or by laying transects across some gradient[3–6]. Ecologists have used larger contiguous area and researchers have also designed certain plots as 'long-term ecological plots' or 'permanent dynamic plots' to monitor variability in species characteristics in spatio-temporal domain[7–15]. Whichever method is adopted, sampling is always a time-consuming process. Also, the size of the plot or quadrat that is used as the basic unit of sampling, varies depending on the type of vegetation and area covered. Though significant variation exists in the ecological patterns captured by random method (usually high and diverse) compared with large-area contiguous plots, these are basically used to understand the behavioural patterns of the species in contiguous scale[11,16–19].

Studying the large-area plots (which range between 1 and 50 ha), researchers have subdivided the entire plot into smaller units for better and quick sampling. The size of the smaller units are 1 m × 1 m, 10 m × 10 m, 30 m × 30 m or sometimes circular plots with varied dimensions[2,19–22]. Within sub-units, ecologists study characteristic features of a species and its population or general diversity patterns, and compare the changing attributes across the quadrats conceptualizing the pattern at higher scale[23]. But when comparisons are made between neighbouring or adjacent quadrats, probability of variation is low as it lies in the same homogenous conditions – may be precipitation, edaphic, sometimes topography. This indicates greater similarity in two closely spaced quadrats compared to those that are separated by larger distances. These samples may be referred to as pseudo replicates, violating the most important assumption of classical inferential statistics that the samples are spatially independent[24].

A homogeneous distribution is one that remains similar on repeated sub-division[25]. The arrangement or ordering of data as a function of location is called spatial autocorrelation of the function and the range of spatial scales in which spatial autocorrelation exists is called spatial dependence[26]. Avoiding spatially dependent quadrats (pseudo replicates), that do not contribute significant changes in any ecological property is necessary, to improve the performance of classical inferential statistics, as the existence of spatial dependence hampers the generality of results and overall performance of classical inferential statistics[27].

The concepts of fractal geometry can suggest a better statistically rectified sampling scheme, which eliminates the problem of spatial dependence among the quadrats[25]. Optimal quadrat sizes for homogeneous or spatially independent distribution can be determined by using the methods of fractal analysis on the data. The quadrats of suggested sizes will be spatially independent and thus independent of the distances by which they are separated.

*For correspondence. (e-mail: anubhav.chaturvedi@students.iiit.ac.in)