Improved Content-Based Image Retrieval via Discriminant Analysis

Smarajit Bose, Amita Pal, Disha Chakrabarti, and Taranga Mukherjee

Abstract—The objective of Content-Based Image Retrieval (CBIR) methods is essentially to extract, from large (image) databases, a specified number of images similar in visual and semantic content to a so- called *query* image. To bridge the *semantic gap* that exists between the representation of an image by low-level features (namely, colour, shape, texture) and its high-level semantic content as perceived by humans, CBIR systems typically make use of the *relevance feedback* (RF) mechanism. RF iteratively incorporates user-given inputs regarding the relevance of retrieved images, to improve retrieval efficiency. One approach is to vary the weights of the features dynamically via *feature reweighting*. In this work, a novel approach has been proposed for improving the retrieval accuracy of CBIR system which incorporates RF based on feature reweighting using discriminant analysis. Results of a number of experiments have been presented to illustrate the significant improvement is retrieval accuracy with the proposed approach.

Index Terms—Content-based image retrieval, Discriminant Analysis, Feature reweighting, Relevance feedback.

I. INTRODUCTION

Due to rapid advances in digital imaging technology, countless images are being generated everyday by innumerable sources like defense and civilian satellites, military reconnaissance and surveillance flights, fingerprinting and facial-image-capturing devices for security and forensic purposes, scientific experiments, biomedical imaging and home entertainment systems. Large repositories of images have become ubiquitous due to the availability of cheaper digital storage devices and the internet. However, maintaining such repositories is meaningless in the absence of methodologies that can enable a user to extract or retrieve information (in the form of images) of interest as and when required.

CBIR methods look for images in large databases that are very similar to a supplied query image, where the search is based on the contents of the image rather than metadata. The term content in this context might refer to colour, shape, texture or any other higher-level descriptor(s) that can be derived from the image itself.

In a typical CBIR system, features are extracted from each image in the database and stored in the feature database. The same features are extracted from the query image as well. The system computes the distance or the similarity between the feature vectors for the query image and that of each image in the database, and retrieves images (usually a fixed number, specified by the user, known as the scope of the system) closest to the query image [1], [2]. A lot of research has been taking place in this area, using a variety of paradigms, though the objective is the same. Instances of some state-of-the-art CBIR systems can be found in [3], [4].

The low-level features used to represent an image do not necessarily capture adequately the high-level semantics and human perception of that image. This leads to the so-called semantic gap in the CBIR context. A solution to this problem is user intervention in the form of Relevance Feedback (RF) [5]-[10]. For a given query, the system first retrieves a set of images ranked in order of their similarity to the query image, in terms of a similarity metric, which represents the distance between the feature vector of the query image and that of each image in the database. Then the user is asked to identify images that are relevant or irrelevant (or non-relevant) to his/her query. The system extracts information from these samples and uses that information to improve retrieval results, and a revised ranked list of images is presented to the user. This process continues until there is no further improvement in the result or the user is satisfied with the result. One way of attaining this objective is feature reweighting, which essentially assigns greater weights to features that discriminate well between relevant and non-relevant images, thus enhancing retrieval, and smaller weights to those features that do not.

In this work, a novel approach has been proposed for improving the retrieval accuracy of any CBIR system that uses relevance feedback based on feature reweighting through the application of the well-known statistical methodology known as discriminant analysis [11]. This was motivated by the observation that, with respect to any query image, every RF iteration generates two categories of images– relevant and non-relevant. To implement the next RF iteration, in which another set of relevant images are to be looked for among the remaining database images, a ranking procedure based on discriminant analysis is readily applicable, and actually leads to dramatic improvement in the retrieval accuracy with every RF iteration.

Organization of this paper is as follows. Section II provides an overview of the classical CBIR paradigm based on relevance feedback (RF). Section III presents the proposed approach. Results are presented in Section IV, summarizing

Smarajit Bose and Amita Pal are with the Interdisciplinary Statistical Research Unit, Applied Statistics Division, Indian Statistical Institute, Kolkata, India (e-mail: {smarajit,pamita}@isical.ac.in).

Disha Chakrabarti is with Tata Consultancy Services, Kolkata, India (e-mail: disha.chakrabarti@gmail.com).

Taranga Mukherjee is with the Department of Statistics, University of Calcutta, Kolkata, India (e-mail: tm.custat@gmail.com).

the effectiveness of the contribution made by this work to CBIR.

II. CLASSICAL APPROACH TO CBIR

The user of a typical CBIR system supplies a query image to it and expects it to extract similar images from a large database. An important component of the system is a feature extraction algorithm which is used to process each image in the database and extract a set of features from it. For an image I, let $\mathbf{f}_I = (f_{I_1}, f_{I_2}, \ldots, f_{I_d})'$, a d $\times$ 1 vector in $\mathbb{R}^d$, be the d features extracted. For a database with N images, the d$\times$N matrix $\mathbf{F} = (\mathbf{f}_{I_1}, \mathbf{f}_{I_2}, \ldots, \mathbf{f}_{I_N})'$, whose j-th column is the d$\times$1 feature vector of the j-th image in the database, represents the entire collection of feature vectors that are extracted and stored. The same feature extraction algorithm is used to process the query image Q too, and the query feature vector is obtained, say, $\mathbf{f}_Q = (f_{Q_1}, f_{Q_2}, \ldots, f_{Q_d})'$. The system subsequently uses an appropriate measure to compute the similarity between the query image and each image of the database, and retrieves a fixed number (specified by the user, known as the Scope) of images most similar or closest to the query image.

The inadequacy of the features to represent the perceived content of an image leads to a *semantic gap*, which is bridged through a *relevance feedback* technique (Section II-B).

Details of the basic components of a typical CBIR system are discussed briefly in the following sections.

A. Similarity Measures

The similarity between the query image *Q* and any other image I is inversely proportional to the distance between their respective feature vectors. Popular choices of distance measures in CBIR literature are

$$d_1(Q, I) = \sum_{j=1}^{d} w_j \left| f_{Q_j} - f_{I_j} \right|, \qquad (1)$$

$$d_2(Q, I) = \sqrt{\sum_{j=1}^{d} w_j (f_{Q_j} - f_{I_j})^2}, \qquad (2)$$

based on the L1- and L2-norms, respectively, which are commonly referred to as the Manhattan distance and the Euclidean distance. The usual practice is to initialize the weights as $w_j = 1/d$. In this work, both distance measures have been used. The generic distance measure has been denoted by $d(Q, I)$ in the following sections.

B. Relevance Feedback (RF)

As mentioned earlier, the relatively low-level features used to represent an image are generally not able to capture adequately its semantic content as perceived by human beings. This creates the so-called semantic gap in the CBIR context. Relevance Feedback (RF) is a commonly-used mechanism which aims to bridge this gap through user intervention

[5]-[10]. For any given query, the system first retrieves a set of images from the database, ranked in order of their similarity to the query image. The user is then asked to identify images that are relevant or irrelevant (or non-relevant) to his/her query. The system extracts information from these samples, uses that information to improve retrieval results, and presents a revised ranked list of images to the user. This process is repeated until there is no further improvement in the result or the user is satisfied with the result.

One popular method for providing this feedback is feature reweighting, which is described below.

C. Feature Reweighting

This widely-used method for implementing relevance feedback assigns different weights to different features [2], [12]. These weights are modified in each iteration of the relevance feedback. Larger weights are given to those features that discriminate well between relevant and non-relevant images and thus enhance retrieval accuracy. A choice of weights used by Das [12] is based on the ratio of feature variability over all retrieved to the relevant images that are retrieved. Let $\sigma_j^{(t)}$ and $\sigma_{rel,j}^{(t)}$, respectively, denote the standard deviations of f_j over the sets $R_t \cup N_t$ and R_t, where R_t and N_t represent the sets of relevant and non-relevant images at the *t*-th RF iteration. A very obvious choice of the weight for the feature f_j at the (*t*+1)-th RF iteration is

$$w_j^{(t+1)} = \frac{\sigma_j^{(t)}}{\sigma_{rel,j}^{(t)}}. \qquad (3)$$

When no relevant image (other than the query itself) is retrieved, the denominator is assigned a small positive value ε to avoid the computational problem arising due to the denominator in (3) becoming zero. The value of ε is chosen such that the weights do not change significantly.

An efficient way of using both positive and negative samples has been proposed by Wu and Zhang [2]. They used a discriminant ratio to determine the ability of a feature to separate relevant images from the non-relevant ones. If $F_{rel,j}^{(t)} = \left\{ f_{I_j}, I \in R_t \right\}$, the collection of the *j*-th features of all images in R_t, then the dominant range over relevant images at the *t*-th iteration for the *j*-th feature component is defined as:

$$D_j^{(t)} = \left\lfloor \min F_{rel,j}^{(t)}, \max F_{rel,j}^{(t)} \right\rfloor \qquad (4)$$

A discriminant ratio (as in [13]) can be used to determine the ability of a feature component to separate the relevant images from the non-relevant ones:

$$\delta_j^{(t)} = 1 - \frac{\text{Number of non-relevant images in } D_j^{(t)}}{|N_t|} \qquad (5)$$

The value of $\delta_j^{(t)}$ lies between 0 and 1. It is 0 when all

non-relevant images are within the dominant range and thus, no weight should be given for that feature component. On the other hand, when there is not a single non-relevant image lying within the dominant range, maximum weight should be given to that feature component. Das [12] proposed a new weight, combining the two weights defined in (3) and (5) as

$$w_j^{(t+1)} = \delta_j^{(t)} * \frac{\sigma_j^{(t)}}{\sigma_{rel,j}^{(t)}}. \qquad (6)$$

III. THE PROPOSED DISCRIMINANT ANALYSIS-BASED APPROACH

Here we propose a discriminant analysis-based approach that not only gives importance to the relevant images but also extracts valuable information from non-relevant images, thereby further enhancing the image retrieval performance.

A. Discriminant Analysis [11]

Consider the problem of discriminating between two groups of objects C_1 and C_2, for each of which (training) observations are available on a *p*-variate feature vector **X**. To classify an object with feature vector **x**, its distance from the mean vectors, $\mathbf{m}_1$ and $\mathbf{m}_2$, of the two classes is computed. Typically, the distance used is the Mahalanobis distance whose square is $(\mathbf{x}-\mathbf{m}_i)' \mathbf{S}_i^{-1} (\mathbf{x}-\mathbf{m}_i)$, $\mathbf{S}_i$ being the dispersion matrix for the observations from the *i*-th class. If $d_1(\mathbf{x})$ and $d_2(\mathbf{x})$ are the distances of the object from C_1 and C_2 respectively, then it is classified into C_1 if $d_1(\mathbf{x}) - d_2(\mathbf{x}) > 0$, and into C_2 otherwise, ties being resolved arbitrarily.

It is to be noted that if the dispersion matrices $\mathbf{S}_i$ are diagonal, then the aforementioned Mahalanobis distance reduces to the weighted Euclidean distance $\sqrt{\sum_j w_j (x_j - m_j)^2}$ with $w_j = 1/\sigma_{jj}$.

B. Application to CBIR

With respect to a particular query image *Q*, let the class of relevant and non-relevant images in an image database *S* be represented by *R* and *NR* respectively. Then $F = R \cup NR$ represents the set of all images retrieved. Once all images in *F* have been retrieved from *S*, the problem of labeling the remaining images in $S \setminus F$ can be viewed as a discrimination problem which can be solved by the discriminant analysis rule given above, based on the difference $D_R - D_{NR}$, where D_R and D_{NR} are respectively the distances of any arbitrary image *I* in $S \setminus F$. If $|R| < k$ (the scope) then, to get more relevant images in the next RF iteration, the values of D_R / D_{NR} are arranged in increasing order and the images in $S \setminus F$ which correspond to the $k - |R|$ smallest distances are retrieved.

However, ranking the images by the value of $D_R - D_{NR}$

does not take into account the absolute values of the distances, that is, D_R and D_{NR} however small or large, are considered to be equivalent as long as their difference is the same. This may lead to error in retrieval. Therefore, one must make the relevance score relative to the values of D_R and D_{NR}. One possible choice of the relevance score of an image is

$$\frac{D_R - D_{NR}}{(D_R + D_{NR})/2}. \qquad (7)$$

Denoting the set of all relevant and non-relevant images, as labeled by the user till the *t*-th RF iteration, by $R^{(t)}$ and $NR^{(t)}$ respectively, it is noted that $R^{(t)} \subset R$ and $NR^{(t)} \subset NR$. Further, let us denote the set of all images retrieved till the *t*-th iteration by $F^{(t)}$, that is, $F^{(t)} = R^{(t)} \cup NR^{(t)}$.

The proposed method can be described succinctly by the following algorithm, where *S* denotes set of all database images, *k* = scope, and *Q* denotes the query image:

Step 1.   *Initialize t = 0, k_t, the current scope by k, and A_t, the set of available images at iteration no. t, by S.*

Step 2.   *For each $I \in A_t$, compute the distance (Euclidean/Manhattan) between I and Q and retrieve top k_t images for which distance is minimum.*

Step 3.   *Categorize retrieved images into classes R and NR, creating $R^{(t)}$ and $NR^{(t)}$. $F^{(t)} = R^{(t)} \cup NR^{(t)}$.*

Step 4. $t \leftarrow t+1$, $A_t \leftarrow A^t \setminus F^{(t)}$ *and* $k_t \leftarrow k_t - |R^{(t)}|$.

Step 5.   *If $k_t > 0$, for each $I \in A_t$, compute relevance score (7), substituting $R^{(t)}$ and $NR^{(t)}$ respectively for R and NR. Retrieve top k_t images for which score is maximum, and go to Step 3; else stop.*

IV. IMPLEMENTATION AND RESULTS

A. Image Databases Used

To demonstrate the effectiveness of the proposed approach, three benchmark databases were used, which are listed and briefly described in Table I.

TABLE I: IMAGE DATABASES USED

Name	Size	No. of Categories	Size per Category
DB2000	2000	10	200
CALTECH	8365	93	26-871
COREL	10800	80	>100

B. Features Used

Standard features like those obtained from the Colour Co-occurrence Matrix (CCM) [13]-[15] and some visual descriptors from the MPEG-7 standard [16], [17] have been used in this work. Predictably, they are based on properties like colour, shape, edge and texture, which are key descriptors

of image content. These features are described below in brief.

TABLE II: RETRIEVAL EFFICIENCY (%) OF PROPOSED METHOD ON THE DB2000 DATABASE

Features	Method	Distance	RF Iteration no.						
			0	1	2	3	4	5	6
CCM	Baseline	Euclidean	52.74	71.49	79.54	84.35	87.29	89.52	91.14
		Manhattan	56.41	73.69	82.22	86.83	89.68	91.73	93.18
	Proposed	Euclidean	52.74	77.70	87.90	92.69	95.03	96.35	97.18
		Manhattan	56.41	80.04	89.18	93.52	95.65	96.94	97.74
CLD	Baseline	Euclidean	57.14	73.97	81.86	86.28	89.14	91.05	92.41
		Manhattan	54.56	73.54	82.57	87.32	90.24	92.15	93.64
	Proposed	Euclidean	57.14	80.80	90.51	94.68	96.52	97.40	97.91
		Manhattan	54.56	78.03	87.95	93.03	95.72	97.13	98.05
CSD	Baseline	Euclidean	62.14	78.82	86.49	90.49	92.54	93.88	94.82
		Manhattan	63.34	79.24	87.46	91.42	93.66	95.03	95.88
	Proposed	Euclidean	62.14	83.79	91.80	95.28	97.07	98.03	**98.66**
		Manhattan	63.34	84.51	92.29	95.52	97.11	98.04	98.46
EHD	Baseline	Euclidean	60.10	75.54	83.68	88.01	90.60	92.51	93.90
		Manhattan	59.51	74.76	83.45	87.86	90.79	92.74	94.13
	Proposed	Euclidean	60.10	80.72	89.56	93.79	96.07	97.15	97.93
		Manhattan	59.51	79.27	88.53	93.20	95.61	96.97	97.77

1) Features from the colour co-occurrence matrix (CCM)

An important representation of the spatial distribution of colour in an image is provided by the colour co-occurrence matrix (CCM) [13]-[15]. The $L \times L$ CCM of an image having colour levels in any one of the dimensions of the *HSV* (Hue, Saturation, Value) colour space, denoted by $\mathbf{P} = \left(\left(p_{ij} \right) \right)$, is such that p_{ij} represents the proportion of pixels with colour level i co-occurring with other pixels with colour level j, at a relative position, say, d. The diagonal elements of the CCM give the colour distribution in the image, while the non-diagonal elements convey shape information, since colour changes between adjacent pixels indicates the possible existence of an object edge. The feature vector used consists of all L diagonal elements of the CCM as well as a single number to represent the information contained in its non-diagonal elements, defined as

$$ave\_ndiag = \sum_{i=1}^{L-1} \sum_{j=i+1}^{L} (i+j) p_{ij}, \tag{8}$$

where i and j are respectively the row and column indices.

It has been observed by researchers that $L_H = 16$ and $L_S = L_v = 3$ are good choices for number of quantization levels of H, S and V for specifying co-occurrence matrices. A co-occurrence distance $d = 1$ has been used in this work and pixel pairs in both vertical and horizontal directions have been considered, leading to symmetric co-occurrence matrices. Thus only upper diagonal elements of the CCMs needed to be considered.

Consequently, D = (16 + 1 + 3 + 1 + 3 + 1) = 25 features were used in this work, following [12].

2) Features from MPEG 7

MPEG-7 is a Multimedia Content Description standard, belonging to the MPEG (Moving Picture Experts Group) series [16], [17]. It specifies a number of descriptors, out of which this work uses

1) the *Color Structure Descriptor* (CSD) which describes the local spatial distribution of colors,
2) the *Color Layout Descriptor* (CLD) which describes the global spatial distribution of colors, and
3) the *Edge Histogram Descriptor* (EHD) which specifies the spatial distribution of five types of edges in the image.

C. Implementation

TABLE IIII: RETRIEVAL EFFICIENCY (%) OF PROPOSED METHOD ON THE CALTECH DATABASE

	Method	Distance	RF Iteration no.						
			0	1	2	3	4	5	6
CCM	Baseline	Euclidean	25.46	33.35	37.58	40.47	42.60	44.39	45.97
		Manhattan	26.80	34.29	38.69	41.73	44.02	45.95	47.56
	Proposed	Euclidean	25.46	36.68	41.89	45.34	48.00	50.11	51.99
		Manhattan	26.80	37.31	42.44	45.77	48.23	50.39	52.32
CLD	Baseline	Euclidean	36.11	43.75	47.58	50.15	52.09	53.78	55.25
		Manhattan	36.13	44.31	48.77	51.56	53.66	55.43	56.98
	Proposed	Euclidean	36.11	48.51	54.69	58.67	61.51	63.78	65.68
		Manhattan	36.13	47.83	53.70	57.72	60.60	62.86	64.71
CSD	Baseline	Euclidean	31.17	38.13	42.19	44.88	46.92	48.57	50.11
		Manhattan	30.88	38.07	42.41	45.12	47.21	48.94	50.47
	Proposed	Euclidean	31.17	40.24	44.79	47.97	50.57	52.70	54.66
		Manhattan	30.88	39.81	44.58	47.91	50.60	52.84	54.69
EHD	Baseline	Euclidean	42.51	50.09	55.02	58.72	67.93	63.35	65.06
		Manhattan	42.87	49.55	54.74	58.14	60.76	62.87	64.64
	Proposed	Euclidean	42.51	53.59	60.00	64.37	67.62	70.28	**72.56**
		Manhattan	42.87	53.33	59.44	63.72	67.03	69.61	71.78

TABLE IV: RETRIEVAL EFFICIENCY (%) OF PROPOSED METHOD ON THE COREL DATABASE

	Method	Distance	RF Iteration no.						
			0	1	2	3	4	5	6
CCM	Baseline	Euclidean	18.70	26.87	32.35	36.52	39.81	42.67	45.13
		Manhattan	20.21	28.60	34.64	39.16	42.76	45.82	48.45
	Proposed	Euclidean	18.70	30.29	38.52	44.64	49.41	53.31	56.47
		Manhattan	20.21	31.69	39.76	45.50	49.97	53.67	56.68
CLD	Baseline	Euclidean	26.30	37.67	44.57	49.37	52.89	55.73	58.25
		Manhattan	23.35	36.56	44.54	49.86	53.71	56.83	59.37
	Proposed	Euclidean	57.14	80.80	90.51	94.68	96.52	97.40	97.91
		Manhattan	54.56	78.03	87.95	93.03	95.72	97.13	**98.05**
CSD	Baseline	Euclidean	36.42	49.29	57.30	62.57	66.46	69.46	71.94
		Manhattan	36.91	49.39	57.71	63.28	67.28	70.41	72.90
	Proposed	Euclidean	36.42	50.16	59.17	65.97	71.10	74.90	78.03
		Manhattan	36.91	54.42	64.96	71.58	76.15	79.46	81.99
EHD	Baseline	Euclidean	25.54	34.72	41.78	46.92	50.79	53.96	56.63
		Manhattan	25.96	34.60	41.90	47.18	51.19	54.48	57.27
	Proposed	Euclidean	25.53	37.35	45.76	52.14	57.27	61.53	65.02
		Manhattan	25.96	36.95	45.02	51.36	56.57	60.85	64.35

The CBIR system which incorporates RF via feature reweighting using the weight defined in (6) is used as the baseline against which the proposed method is compared.

To evaluate retrieval performance, the following measure, called Retrieval Efficiency (RE), which is inspired by the standard measure *Precision* but is more appropriate when RF

is used [18], has been computed:

$$\text{Retrieval Efficiency } (RE)$$
$$= \frac{\text{Number of relevant images retrieved}}{\text{Scope}}. \qquad (9)$$

Tables II-IV provide details of the steady increase in RE over 6 RF iterations on the three databases listed above, using both the baseline and the proposed methods with the Manhattan distance as well as the Euclidean distance. It is immediately evident that the proposed method yields significant increase in RE at each iteration. The quantum of increase varies with the feature and/or the distance measure used. The RE corresponding to the most remarkable retrieval performance for each database is highlighted in bold font in the corresponding table.

V. CONCLUSION

The proposed method has been found to consistently and remarkably improve retrieval efficiency with every RF iteration, irrespective of the distance measure or the features used. Of course, the degree of improvement depends upon the complexity of the database and the optimal choice of features and/or distance measure is dependent on the database.

REFERENCES

[1] Y. Rui, T. Huang, M. Ortega, and S. Mehrotra, "Relevance feedback: a power tool for interactive content-based image retrieval," *IEEE Transactions on Circuits and Video Technology*, 1998.

[2] Y. Wu and A. Zhang, "A feature reweighting approach for relevance feedback in image retrieval," in *Proc. IEEE International Conference on Image Processing* (ICIP 2002), (Rochester, New York), September 2002.

[3] M. Chowdhury and M. K. Kundu, "Comparative assessment of efficiency for content based image retrieval systems using different wavelet features and pre-classifier, *Multimedia Tools and Applications*, vol. 74, no. 24, pp. 11595–11630, 2015.

[4] M. K. Kundu, M. Chowdhury, and S. Rota Bulo, "A graph-based relevance feedback mechanism in content-based image retrieval," *Knowledge-Based Systems*, vol. 73, pp. 254–264, 2015.

[5] H. Muller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content-based image retrieval systems in medical applications–clinical benefits and future directions," *International Journal of Medical Applications*, 2004.

[6] M. Ortega-Binderberger and S. Mehrotra, "Relevance feedback in multimedia databases," *Handbook of Video Databases: Design and Applications*, ch. 1, pp. 23–28, CRC Press, 2003.

[7] Y. Rui, T. S. Huang, M. Ortega, and S.-F. Chang, "Image retrieval: current techniques, promising directions and open issues," *Journal of Visual Communication and Image Presentation*, vol. 10, no. 4, 1999.

[8] A. Yoshitaka and T. Ichikawa, "A survey on content-based retrieval for multimedia databases," *IEEE Transactions on Knowledge and Data Engineering*, vol. 11, no. 1, pp. 81–93, 1991.

[9] H. Zhang, "Relevance feedback in content-based image retrieval," in *Multimedia Information Retrieval and Management–Technological Fundamentals and Applications*, D. D. Feng, W. C. Siu, and H. Zhang, Eds., ch. 3, pp. 57–74, Germany: Springer-Verlag, 2003.

[10] X. S. Zhou and T. S. Huang, "Relevance feedback in image retrieval: a comprehensive review," *ACM Multimedia Systems Journal*, vol. 8, pp. 536–544, April 2003.

[11] C. R. Rao, *Linear Statistical Inference and Its Applications*, Wiley, 2nd ed., 2001.

[12] G. Das, "Reduction of Semantic Gap in Content-based Image Retrieval," PhD thesis, Clayton School of Information Technology, Monash University, Melbourne, Australia, 2007.

[13] S. Shin and T. Choi, "Image indexing in modified color co-occurrence matrix," in *Proc. International Conference on Image Processing*, September 2003.

[14] J. Huang, "Color-spatial image indexing and applications," PhD thesis, Cornell University, 1998.

[15] T. Ojala, M. Rautiainen, E. Matinmikko, and M. Aittola, "Semantic image retrieval with HSV correlograms," in *Proc. 12th Scandinavian Conference on Image Analysis*, (Bergen, Norway), 2001, pp. 621–627.

[16] H. Kosch, *Distributed Multimedia Database Technologies Supported by MPEG-7 and MPEG-21*, CRC Press, 2004.

[17] B. S. Manjunath, J.-R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 703–715, 2001.

[18] S. Bose, A. Pal, J. Mallick, and S. Kumar, "A hybrid approach to content-based image retrieval," in *Proc. the 1st International Conference on Information, Operations Management and Statistics*, September 2013.

Smarajit Bose received B. Stat. and M. Stat. degrees from the Indian Statistical Institute, Calcutta in 1984 and 1986 respectively, a Ph. D. degree from the University of California, Berkeley in 1992, all in statistics.

Between 1991 and 1992, he was a visiting scientist in IBM Almaden Research Center, San Jose where he worked with the Advance Process Monitoring Group. He visited the Statistics Departments of the University of Washington, Seattle during 1992–1993 and the Ohio State University, Columbus during 1993–1996 where he worked on classification and clustering problems and their applications in medical imaging and ergonomics. In 1996, he joined the Indian Statistical Institute, Calcutta and now is a Professor in the Applied Statistics Division of the Institute. He has also visited the University of California, Santa Barbara in 2001, and also in 2002–2003. His current research interests are pattern recognition and its applications in image processing and speaker identification.

Amita Pal obtained a B.Sc. degree with Honours and a M.Sc. degree in statistics from the University of Calcutta in 1979 and 1981 respectively, and a Ph.D. degree in statistics from the Indian Statistical Institute, Kolkata in 1991.

She joined the Indian Statistical Institute as a Lecturer in 1994, and is currently working as an associate professor in the Interdisciplinary Statistical Research Unit of the same Institute. She visited the Imperial College of Science, Technology and Medicine, London in 1994 on a six-month UNDP fellowship. Her research interests included pattern recognition, image processing and statistical machine learning.

Disha Chakrabarti received a bachelor of technology in computer science engineering from the RCC Institute of Information Technology, Kolkata, in 2014.

She is currently working as a systems engineer in Tata Consultancy Services, Kolkata, India.

Taranga Mukherjee received a bachelor's degree in Science with Honours in statistics from Bidhannagar College, Kolkata, under the West Bengal State University in 2011, and a Master of Science (M.Sc.) degree in statistics from the University of Calcutta in 2013.

He is currently pursuing a Ph.D. degree in statistics as a research fellow at the University of Calcutta.