Framework for Content-Based Image Identification with Standardized Multiview Features

Rik Das, Sudeep Thepade, and Saurav Ghosh

Information identification with image data by means of low-level visual features has evolved as a challenging research domain. Conventional text-based mapping of image data has been gradually replaced by content-based techniques of image identification. Feature extraction from image content plays a crucial role in facilitating content-based detection processes. In this paper, the authors have proposed four different techniques for multiview feature extraction from images. The efficiency of extracted feature vectors for content-based image classification and retrieval is evaluated by means of fusion-based and data standardization–based techniques. It is observed that the latter surpasses the former. The proposed methods outclass state-of-the-art techniques for content-based image identification and show an average increase in precision of 17.71% and 22.78% for classification and retrieval, respectively. Three public datasets — Wang; Oliva and Torralba (OT-Scene); and Corel — are used for verification purposes. The research findings are statistically validated by conducting a paired *t*-test.

Keywords: Local threshold, partial coefficients, morphological operator, gray-level co-occurrence matrix, feature extraction, classification, retrieval, *t*-test.

Manuscript received Feb. 3, 2015; revised June 13, 2015; accepted June 25, 2015.

Rik Das (corresponding author, rikdas78@gmail.com) is with the Department of Information Technology, Xavier Institute of Social Service, Jharkhand, India.

Sudeep Thepade (sudeepthepade@gmail.com) is with the Department of Information Technology, Pimpri Chinchwad College of Engineering, Maharashtra, India.

Saurav Ghosh (sauravghoshcu@gmail.com) is with the Department of Information Technology, A.K. Choudhury School of Information Technology, West Bengal, India.

I. Introduction

Explosive growth of image data has been envisaged in information and communication media. Visual data archives have proved to be beneficial for diverse application domains including commerce, education, biomedicine, and military services [1]. Consequently, various methods have been formulated for indexing, recognition, and management of digital images [2].

Content-based image identification is acknowledged as an efficient alternative to traditional text-based methods and has an improved accuracy in terms of image searching [3], [4].

Image data collected from satellites and sensors has been observed to be inherently noisy. In such data, noise has also been added for the purposes of preserving privacy, particularly in the case of medical datasets, demographic datasets, and so on. Hence, the information in these datasets has associated items with existential probability. Techniques have been formulated for identification of these datasets, which are uncertain in nature [5]. Association rule mining has been used to discover latent patterns in image data for *uncertain* image identification [6]. Visual semantic algebra has been applied in machine visual and spatial reasoning and computational intelligence system designs [7].

A novel classification technique based on decision tree theory was proposed; the technique can generate rules with the information present in a classifier even if a source image is lost [8]. Traditional content-based recognition processes have been principally dependent upon extraction of a single feature [9], [10]–[12]. However, the rich content of an image can hardly be described with a solo characteristic.

In this paper, the authors have proposed four different techniques for feature extraction — binarization-based,

http://dx.doi.org/10.4218/etrij.16.0115.0102

transform-based, texture-based, and shape-based techniques — for multiview observations; resultant extracted features from these techniques are of considerably small size.

Two different proposed techniques for image identification are also introduced. The first of these involves the fusion of the four different aforementioned techniques for image recognition and the latter standardizes feature vectors for the purposes of effectively identifying images.

Three well-known image datasets were used for evaluation purposes — Wang; Oliva and Torralba (OT-Scene); and Corel. On the whole, 14,488 images are used in the experimentation process. The proposed fusion technique for content-based image identification outclasses state-of-the-art techniques for content-based image identification and demonstrates an average increase of 17.71% and 22.78% in terms of precision for classification and retrieval, respectively. The findings were validated with a paired *t*-test for statistical significance [13].

II. Related Work

Binarization of images has been considered as a popular technique for feature extraction in content-based image identification. The proficiency of a given binarization technique is dependent upon the selection of an appropriate threshold value. However, threshold selection is greatly affected by a large number of variables, including uneven illumination, inadequate contrast, and so on [14]. Three different techniques for threshold selection — mean threshold selection, local threshold selection, and global threshold selection — have been adopted in the contemporary literature to deal with the aforementioned adversities. Selection of both mean threshold and multilevel mean threshold for feature extraction with binarization has revealed improved classification results in six different color spaces [15]. Extraction of features has been carried out using mean threshold for binarization of significant bit planes of images and from even and odd image varieties for better image classification [16], [17]. The problem is that a mean-threshold selection technique considers only the average of the gray values and not the standard deviation. Hence, the spread of data is not well captured in a binarization process for feature extraction. This concern has been well handled by commencing the computation for the standard deviation and variance of gray values for local-threshold selection techniques [18]–[20]. Threshold calculation has been carried out exhaustively in the case of global-threshold selection techniques for feature extraction using binarization [21], [22].

Texture features of images play a major role in the classification of image data. Gray-level co-occurrence matrix (GLCM) texture features calculated for homogeneity, contrast, angular second moment (ASM), and so on have contributed to noteworthy classification results [23], [24].

Multiview feature extraction has been carried out through means of a color layout descriptor (CLD) from MPEG-7, along with mean, variance, skewness, Kurtosis, energy, and entropy for texture description [25]. The retrieval results have surpassed the accuracy of industry standards.

Feature extraction for content-based image retrieval has also been carried out using a hue-saturation-value (HSV) representation and first order statistic (FOS) for improved performance [26].

Cognitive analysis has proven efficient in the analysis of extracted features from image data [27]. Cognitive systems understand the semantic contents of analyzed data with the help of reasoning.

Pattern recognition has been carried out by automatic semantic categorization and image content perception based on linguistic theories of pattern classification. The process of interpretation of a pattern is based on cognitive resonance [28].

Shape feature extraction for efficient image recognition has been carried out with both a gradient and a Sobel operator [29], and with texture content [30].

Transform-domain techniques have readily contributed to rich feature extraction from image content [31]. Recent techniques have shown proficiency in image identification with relevant features [32], [33]. Color and texture were calculated as local descriptors from *color moments* and moments of Gabor filter responses for higher retrieval performance [34]. Improved retrieval results were observed with visually significant point features being chosen from images as feature vectors [35].

A combination of a CLD and a Gabor texture descriptor was considered as a resourceful image signature for a recognition process [36]. Color, texture, and spatial structure descriptors have been used as feature vectors for boosting the retrieval process [37]. Intra-class and inter-class feature extraction from images has been shown to enhance the retrieval process [38]. The fusion of techniques, such as that in the case of an Edge Histogram Descriptor and angular radial transform, has proven to yield increased precision values compared to those obtained through individual techniques [39].

Image uniqueness has been explored through color histograms and spatial orientation trees for feature extraction [40]. The authors identified a disadvantage common to all the techniques of the existing literature. The techniques were shown to have generated feature vectors of large sizes due to their dependency on image dimension. This consequently slowed down the classification process. In contrast, the techniques proposed here produce feature vectors of small size, independent of the dimensions of an image. The novel methods have outperformed existing established techniques and have shown statistically significant improvement in terms of classification and retrieval performance.

III. Proposed Feature Extraction Techniques

Four different feature extraction techniques are proposed to facilitate a content-based image recognition process. Extraction of multiview signatures is aimed at exploring the diverse features of rich image content. Each of the four proposed feature extraction techniques is explained in more detail below.

1. Feature Extraction with Image Binarization

Images from the Wang; Oliva and Torralba (OT-Scene); and Corel datasets are binarized using a sliding window–based local-threshold selection technique, named NICK threshold selection (see Fig. 1) [41]. The threshold selection technique is an improved version of Niblack's local-threshold selection technique and demonstrates an improved level of feature extraction in the case of white- and light-page images.

A threshold calculation at each pixel location has to be carried out in the case of selection of a local threshold. The local threshold is calculated based on local statistics (namely, the range, the variance, and a surface-fitting parameter). A working formula for local threshold selection is given by

$$T_x = m + k\frac{\sqrt{\sum p_i^2 - m^2}}{\mathrm{NP}}, \tag{1}$$

where, k is a Niblack factor, m is the mean grey value, p_i is the pixel value of the greyscale image, and NP is the total number of pixels. The Niblack factor k has a variation range ranging from –0.1 to –0.2 based on the requirements of a given application. Binarization with higher k-values produces slim and broken strokes; whereas, with smaller k-values, binarization creates thick, unclear strokes.

The size of a neighborhood region is considered small enough to reflect a local illumination level and yet is large enough to include both a portion of an object from the image and a portion of the background from the image.

Red, green, and blue color components are separated from each image before the calculation of a local threshold. Each of the color components are binarized by applying NICK local threshold selection. Pixel values that are higher than the local threshold are clustered into what is called a high-intensity group and those that are lower than the local threshold are clustered into a low-intensity group.

The mean and standard deviation of the gray values is calculated for the two intensity groups. Feature vectors are then derived from a summation of the mean and standard deviation,

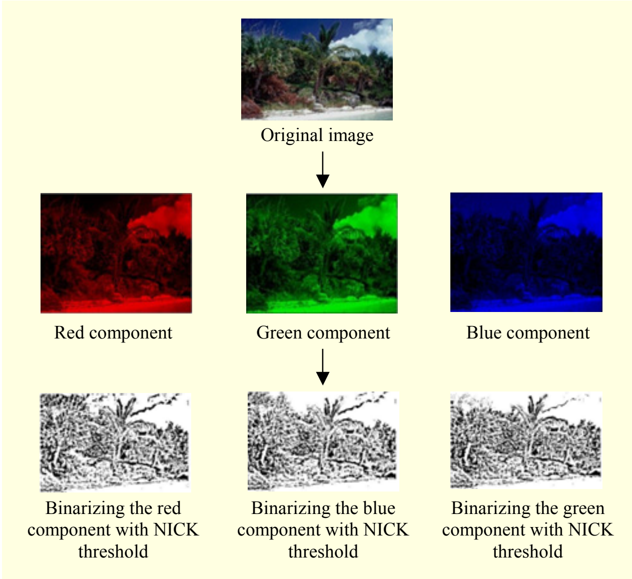


Fig. 1. Effect of binarization with NICK threshold for feature extraction.

for each image in the dataset. The dimension of a feature vector is six; that is, two features are extracted per color component.

$$x_{\mathrm{HIFV}} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} x_{\mathrm{HI}}(i,j) + \sqrt{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(x_{\mathrm{HI}}(i,j) - \mu_{\mathrm{HI}}\right)^2}, \tag{2}$$

$$x_{\mathrm{LOFV}} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} x_{\mathrm{LO}}(i,j) + \sqrt{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(x_{\mathrm{LO}}(i,j) - \mu_{\mathrm{LO}}\right)^2}, \tag{3}$$

where, $\mu = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} x(i,j)$ and $\sigma = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(x(i,j) - \mu\right)^2$.

Here, μ is the mean; σ is the standard deviation; x = R, G, and B for individual components; and T_x is the threshold value for each pixel.

2. Feature Extraction with Image Transform

The intrinsic trait of an image transform is to reposition the component of highest frequency of an image toward the upper end and the component of lowest frequency toward the lower end of an image. The authors have exploited this inherent trait of image transform and have applied Hartley transform to extract transform coefficients. A Hartley transform is a Fourier-related transform that transforms real inputs to real outputs, avoiding the fundamental involvement of complex numbers. It is a linear operator. The working formula for the Hartley transform of a function $f(t)$ can be given by

$$H(\omega) = \{\mathrm{H}f\}(\omega) = \frac{1}{\sqrt{2\Pi}}\int_{-\infty}^{\infty} f(t)\mathrm{cas}(\omega t)\,\mathrm{d}t, \tag{4}$$

where ω is the angular frequency and

ETRI Journal, Volume 38, Number 1, February 2016
http://dx.doi.org/10.4218/etrij.16.0115.0102

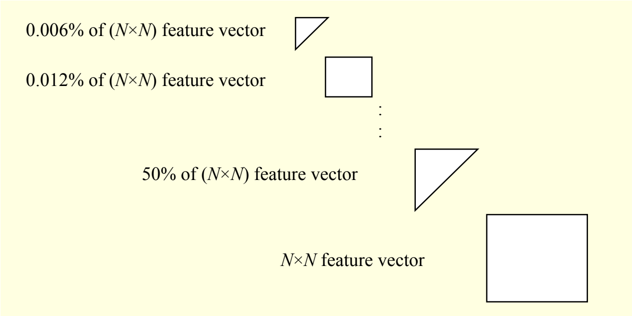


Fig. 2. Extraction of partial Hartley coefficients from transformed image.

$$\operatorname{cas}(t) = \cos(t) + \sin(t) = \sqrt{2}\sin(t + \prod/4) = \sqrt{2}\cos(t - \prod/4)$$

is defined to be the Hartley kernel. Primarily, the three color components (namely, red, green, and blue) are separated from each image. Then, a Hartley transform is applied to each of the color components. The entire set of coefficients is considered as a feature set and the dimension is equivalent to the size of each color component of the image. Further, subsequent partial coefficients are selected ranging from 50% to 0.006% of the full feature-vector size to capture the highest frequency for each of the respective color component (see Fig. 2). The highest recognition rate is observed with 0.012% of the full feature-vector size. The dimension of the feature vector is equal to eight for each color component. On the whole, the dimension of the feature vector is 24 for the three color components in each image in the dataset, irrespective of an image's size.

3. Feature Extraction with Morphological Operator

A *shape feature* plays a vital role in object-based information identification. Shape refers to the contour of a region of interest within an image. Morphological operations have been useful to calculate the value of each pixel in the output image by comparing it to the consequent pixel in the input image. Basic morphological operations comprise of dilation and erosion. The former creates a swelling effect of the shape of an object by adding pixels to the object's boundaries, and the latter participates in object shape shrinking by removing pixels from the object's boundaries. A morphological edge-extraction technique (namely, bottom-hat morphological edge extraction) has been applied to each of the red, green, and blue color component of our images to locate an area of interest for feature extraction (see Fig. 3).

The contour region and background portion of the individual color components of an image are grouped into two different clusters. The mean and standard deviation for the two clusters are derived to compute the feature vectors of the image. The feature-vector size is equal to six for an image.

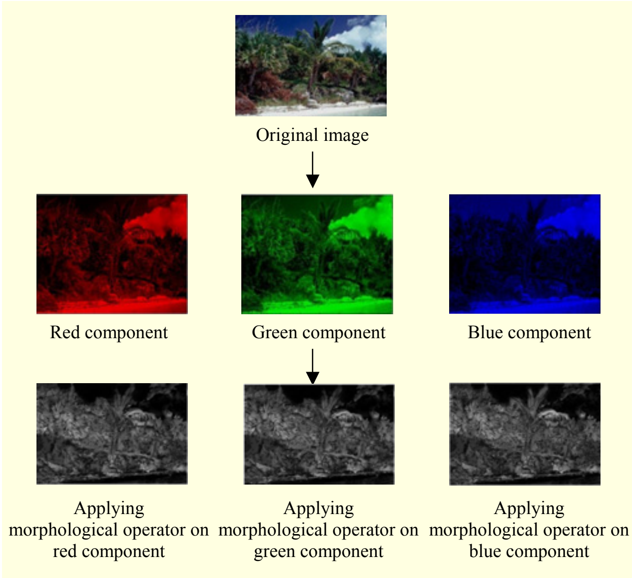


Fig. 3. Effect of applying morphological operator for feature extraction.

4. Feature Extraction with Gray Level Co-occurrence Matrix (GLCM)

The spatial definition of a visual pattern of an image is characterized by texture features. Specific texture detection in an image is performed by representing a texture as a two-dimensional gray-level variation, known as GLCM. GLCM is a statistical method for investigating textures that consider the spatial relationships of pixels. The texture of an image can be represented by a GLCM function by computing the number of existing pixel pairs with specific values and with specific pixel relationships in an image, followed by extraction of statistical measures from the GLCM matrix. The definition for the normalized probability of a GLCM matrix is

$$P_\delta(i,j) = \frac{\#\left\{\left|(x,y),(x+d,y+d)\right| \in S \,\middle|\, f(x,y)=i, f(x+d,y+d)=j\right\}}{\#S}, \quad (5)$$

where, $x, y = 0, 1, \ldots, N-1$ are pixel co-ordinates; $i, j = 0, 1, \ldots, L-1$ are grey levels; S is the set of pixel pairs in an image with a certain relationship; $\#S$ is the number of elements in S; and $P_\delta(i, j)$ is the probability density. Computation of the GLCM matrix is carried out in different directions; namely, $\delta = 0°$, $\delta = 45°$, $\delta = 90°$, and $\delta = 135°$. Four statistical parameters for energy, contrast, entropy, and correlation are calculated based on the GLCM (see (6)–(9)). The repetition of pixel pairs is measured by the energy or ASM. The variance of a gray level is quantified by contrast; the *disorder* of an image is measured by entropy; and correlation returns a measure of how correlated a pixel is to its neighbor over the whole image.

$$\mathrm{ASM} = \sum\sum P^2(i,j), \tag{6}$$

$$\mathrm{CON} = \sum\sum (i-j)^2 P(i,j), \tag{7}$$

$$\mathrm{ENT} = -\sum\sum P(i,j)\log\left(P(i,j)\right), \tag{8}$$

$$\mathrm{COR} = \frac{\sum\sum ijP(i-j) - \mu_x\mu_y}{\sigma_x\sigma_y}, \tag{9}$$

where μ_x and μ_y and σ_x and σ_y denote the mean and standard deviation for P_X and P_Y respectively; P_X is the sum of each row of the GLCM and P_Y is the sum of each column of the GLCM. The mean and standard deviation of all the parameters are considered to compute feature vectors for each image; the dimension of a feature vector in this case is equal to eight.

5. Framework for Classification

Two different techniques are introduced for the purpose of classification. The first technique fuses the classification decision obtained from the four proposed feature extraction techniques (see Fig. 4). Four different similarity measures — Euclidian distance, city block distance, Canberra distance, and mean squared error — are used to classify the four feature extraction techniques, respectively. The working formulae for the similarity measures can be seen in (10)–(13). The fusion process is carried out using *Z*-score normalization, as in (14). *Z*-score normalization is the calculated arithmetic mean and standard deviation of the distance measures in (10)–(13). A final distance is computed by adding the weighted sums of the normalized distances. The weights are calculated from the individual average precision rates of each of the proposed techniques.

$$D_{\text{cityblock}} = \sum_{i-1}^{n} |Q_i - D_i|, \tag{10}$$

$$D_{\text{Euclidian}} = \sqrt{\sum_{i=1}^{n} (Q_i - D_i)^2}, \tag{11}$$

$$D_{\text{Canberra}} = \sum_{i=1}^{n} \frac{|Q_i - D_i|}{|Q_i| + |D_i|}, \tag{12}$$

$$D_{\text{MSE}} = \frac{1}{mn} \sum_{i=1}^{n} (Q_i - D_i)^2, \tag{13}$$

where Q_i is the query image and D_i is the database image.

$$\text{dist}_x = \frac{\text{dist}_x - \mu}{\sigma}, \tag{14}$$

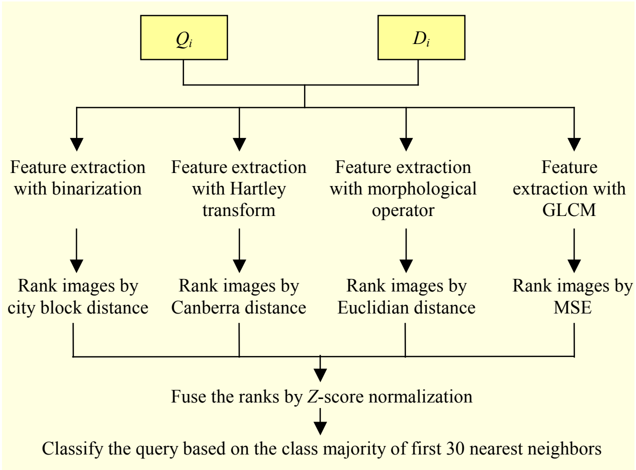


Fig. 4. Illustration for decision fusion with *Z*-score normalization.

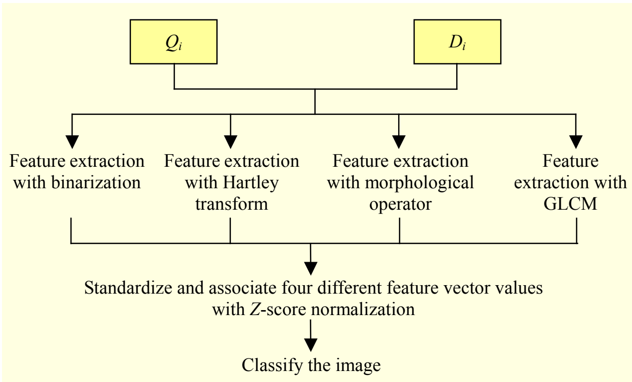


Fig. 5. Illustration for feature-vector standardization with *Z*-score normalization.

where μ is the mean and σ is the standard deviation.

The second technique applies standardization of values achieved from the four different feature-vector sets (see Fig. 5). The feature-vector values obtained from the four different techniques of feature extraction are allied successively and applied with *Z*-score normalization. The normalization process is implemented to avoid dependence on a classification decision (for a feature vector) with high attribute values, which has the potential to have a greater effect or "weight" on the classification results. The process normalizes the data within a common range, such as [–1, 1] or [0, 1]. The normalization process entails the calculation of the arithmetic mean and standard deviation of the feature values extracted from each of the four techniques. The total size of the feature vector of the normalized combined feature values of the four different feature extraction techniques is 44 (6 + 24 + 6 + 8). Further, classification is carried out with a new feature-vector set comprising of elements from the four different feature extraction techniques with the help of the Euclidian distance measure.

Fig. 6. Sample datasets.

6. Framework for Retrieval

A retrieval process is carried out with the *Z*-score-normalized combined feature-vector set. A query image is primarily classified to its nearest category by means of the Euclidian distance similarity measure. Then, a classified query is forwarded to retrieve the top 20 matches only from the class of interest. In contrast to conventional retrieval techniques in terms of searching an entire dataset, the proposed method has a much smaller searching space, which is restricted only within the classified category.

7. Experimental Verification

Experiments are carried out with three widely used public datasets — Wang (10 categories, 1,000 images); Oliva and Torralba (8 categories, 2,688 images); and Corel (80 categories, 10,800 images) [9], [42]. The classification performance is measured by evaluating the entire training set in the case of all three datasets. Retrieval is carried out with the Wang dataset. A sample of the datasets used for testing purposes is given in Fig. 6.

8. Evaluation Metric

The classification results are assessed by an F1 score — the harmonic mean of both the precision and the recall. A high F1 score indicates a superior classification result.

$$\text{F1 score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (15)$$

where "Precision" is the probability that object is classified correctly as per the actual value, and "Recall" is the probability that a classifier produces a true positive result.

The category-wise F1 score for classification by binarization with NICK threshold is given in Fig. 7. The "Horses" category shows the highest classification performance with a maximum F1 score of 0.86, closely followed by the "Roses" and "Bus" categories. The "Gothic Structure" category has the lowest F1 score.

A comparison of average F1 scores for classifications with different partial coefficients extracted from Hartley transforms is given in Fig. 8. The partial coefficient extracted with 0.012% of the full feature-vector size gives the maximum F1 score.

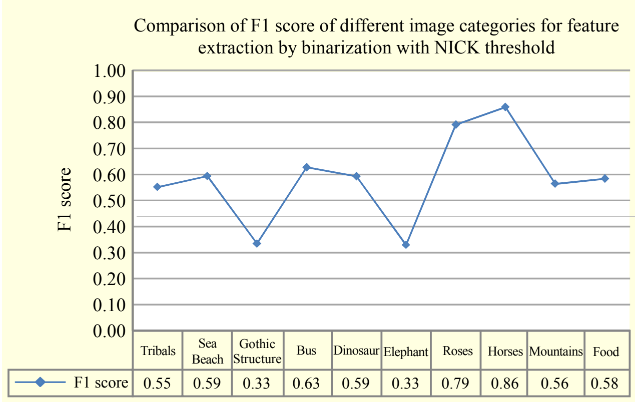


Fig. 7. Category-wise F1 score for classification with feature extraction by binarization.

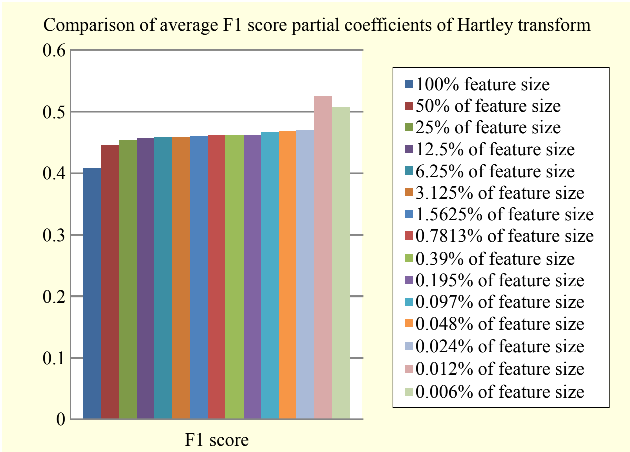


Fig. 8. Comparison of average F1 scores for partial coefficients.

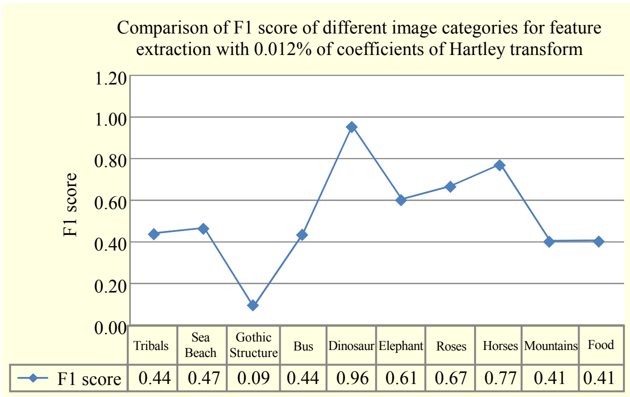


Fig. 9. Category-wise F1 score for classification with full feature-vector extraction by 0.012% of Hartley coefficient.

A further category-wise comparison of the F1 score for classification with 0.012% of the full feature-vector size is given in Fig. 9.

It is observed from Fig. 9 that the "Dinosaur" category gives the maximum F1 score, closely followed by the "Horses" category. The F1 score for the "Gothic Structure" category is

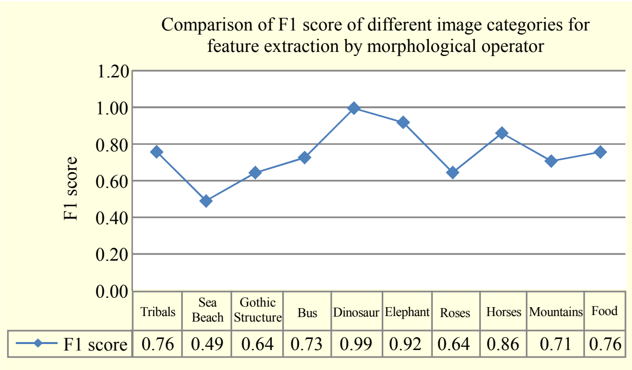


Fig. 10. Category-wise F1 score for classification with feature extraction by morphological operator.

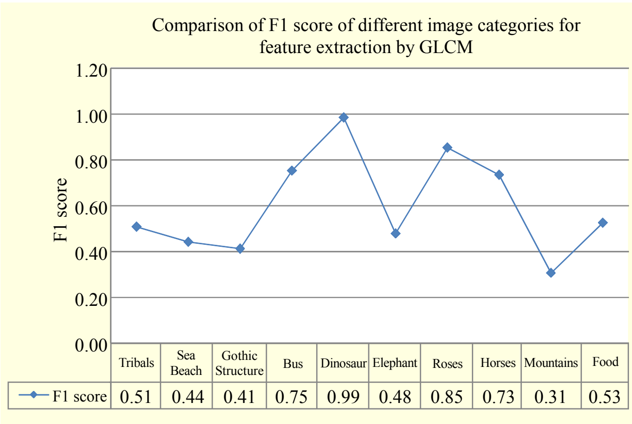


Fig. 11. Category-wise F1 score for classification with feature extraction by GLCM.

noticeably drastically low.

A category-wise comparison of the F1 score for feature extraction under a morphological operator is given in Fig. 10. The results in Fig. 10 reveal equal classification rates for the "Tribals" and "Food" categories. The best performance, with a maximum F1 score of 0.99, is given by the "Dinosaur" category, followed by the "Elephant" category. The minimum F1 score is observed by the "Sea Beach" category.

Finally, the F1 score for each category is given in Fig. 11 for feature extraction using GLCM. Figure 11 reveals the highest F1 score obtained by the "Dinosaur" category, followed by the "Roses" category. The "Mountains" category has the lowest F1 score. The results are compared with respect to the precision and recall values of the individual techniques for classification. The retrieval performance is measured with precision and recall. *Precision* is the ratio of number of relevant images retrieved to the number of total images retrieved, and *recall* is the ratio of number of relevant images retrieved to the total number of images in the database. A derivation of precision and recall for each category of images is carried out, and the average precision and recall values are computed.

9. Results and Discussion

The experiments are performed using a PC with an Intel core i5 processor with 4 GB RAM and using Matlab 7.11.0(R2010b). Initially, classification is carried out for each of the individual techniques for feature extraction; namely, feature extraction by binarization with NICK threshold, feature extraction by partial coefficients of Hartley transform, feature extraction by morphological operator, and feature extraction by GLCM. Subsequently, the classification decisions obtained from the four different feature extraction techniques are fused using *Z*-score normalization, and the result for a combined classification result is thus achieved. Then, *Z*-score normalization is applied to those feature vectors obtained from the four individual techniques (allied consecutively). The normalized data of the four feature extraction techniques forms a single feature vector, which is examined to assess the overall classification performance.

Primarily, classification is performed with the Wang dataset. One thousand queries are evaluated to derive the precision and recall values of classification for each of the four proposed techniques. A comparison of these values for the Wang dataset is given in Table 1.

The precision and recall values in Table 1 clearly establish the superiority of the classification results by decision fusion with *Z*-score normalization and by feature standardization with *Z*-score normalization in comparison to those obtained using individual techniques. Consequently, it is observed that classification by feature standardization with *Z*-score normalization produces higher precision and recall values with respect to classification by decision fusion with *Z*-score normalization. Accordingly, it is inferred that classification by feature standardization with *Z*-score normalization outperforms classification by individual techniques. The feature extraction process with the four individual techniques is further applied to the OT-Scene and Corel datasets, and the classification results are compared to those of classification by feature standardization with *Z*-score

Table 1. Comparison of average precision and recall for classification with Wang dataset.

Wang dataset (classification results)	Avg. precision	Avg. recall
Feature extraction by binarization	0.618	0.595
Feature extraction by partial of Hartley transform coefficients	0.554	0.553
Feature extraction by morphological operator	0.767	0.761
Feature extraction by GLCM	0.615	0.617
Classification by decision fusion	0.779	0.770
Classification by feature standardization	0.844	0.841

http://dx.doi.org/10.4218/etrij.16.0115.0102

Table 2. Comparison of average precision and recall values for OT-Scene dataset.

OT-Scene dataset (classification results)	Avg. precision	Avg. recall
Feature extraction by binarization	0.468	0.432
Feature extraction by partial of Hartley transform coefficients	0.371	0.376
Feature extraction by morphological operator	0.602	0.589
Feature extraction by GLCM	0.528	0.526
Classification by feature standardization	0.651	0.622

Table 3. Comparison of average precision and recall values for Corel dataset.

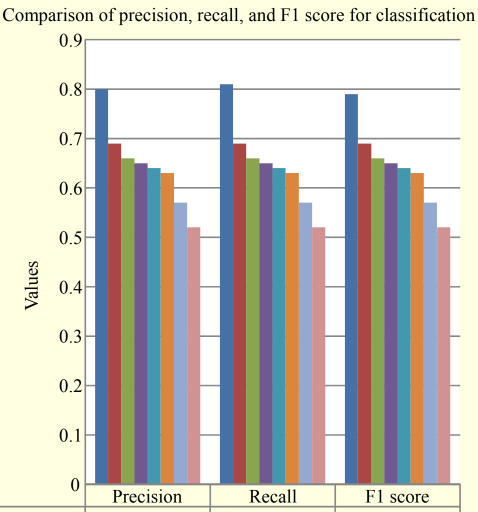
Corel dataset (classification results)	Avg. precision	Avg. recall
Feature extraction by binarization	0.313	0.308
Feature extraction by partial of Hartley transform coefficients	0.171	0.201
Feature extraction by morphological operator	0.287	0.311
Feature extraction by GLCM	0.209	0.245
Classification by feature standardization	0.442	0.429

normalization. The results are shown in Tables 2 and 3.

For both datasets, it is observed that classification by feature standardization with Z-score normalization surpasses the precision and recall values for classification with each of the four feature extraction techniques (see Tables 2 and 3). Further, the precision, recall, and F1-score values for classification by feature standardization with Z-score normalization are compared to state-of-the art classification techniques in Fig. 12. An experiment is carried out on the Wang dataset.

The proposed techniques have the highest precision, recall, and F1-score values for classification (see Fig. 12). A paired t-test (two-tailed) is conducted to find out the p-values of the existing techniques with respect to the proposed techniques (see Table 4). The test produces a comparative measure of the variation in the precision values of the proposed and existing techniques to evaluate the actual difference between the two means appearing in H_0: $\mu d = 0$ and H_1: $\mu d < 0$. The results in Table 4 have revealed significant difference in precision values for classification with the proposed technique compared to the existing techniques.

Then, the proposed techniques for feature standardization with Z-score normalization are tested for retrieval performance with the Wang dataset. Five images are selected randomly from each category in the Wang dataset, and a total of 50 images are selected as query images from 10 different categories. The query images are *fired* one at a time. The retrieval method



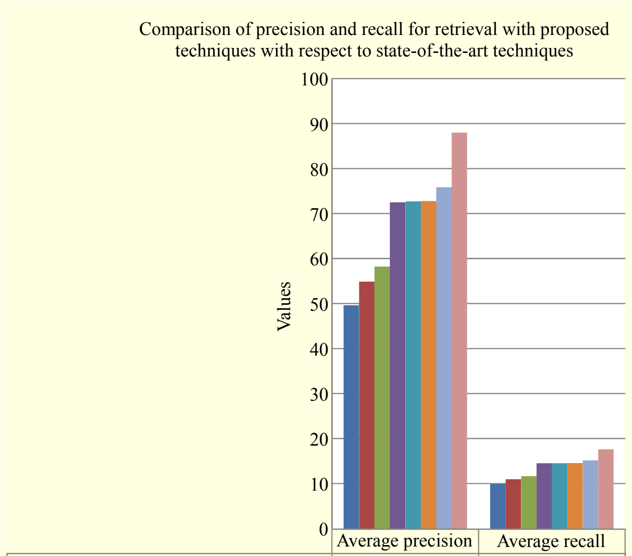
	Precision	Recall	F1 score
Proposed	0.80	0.81	0.79
(S. Thepade et al., 2014)	0.69	0.69	0.69
(H.B. Kekre et al., 2013)	0.66	0.66	0.66
(S. Thepade et al., 2013)	0.65	0.65	0.65
(Y. Yanli & Z. Zhenxing, 2012)	0.64	0.64	0.64
(M. Ramirez-Ortegón & R. Rojas, 2010)	0.63	0.63	0.63
(C. Liu, 2013)	0.57	0.57	0.57
(S.H. Shaikh, 2013)	0.52	0.52	0.52

Fig. 12. Comparison of average precision, recall, and F1-score values of proposed and existing techniques.

Table 4. t-test for significance in precision results.

Wang dataset (Classification results)	t-calc	p-value	Significance of difference in value
Feature extraction using method S. Thepade et al., 2014	3.0833	0.0150	Significant
Feature extraction using method S.B. Kekre et al., 2013	3.4982	0.0081	Significant
Feature extraction using method S.B. Kekre et al., 2012	3.2518	0.0117	Significant
Feature extraction using method S. Thepade et al., 2013	3.2856	0.0111	Significant
Feature extraction using method Y. Yanli and Z. Zhenxing, 2012	3.7700	0.0055	Significant
Feature extraction using method M. Ramirez-Ortegón and R. Rojas, 2010	3.8062	0.0052	Significant
Feature extraction by binarization using method C. Liu, 2013	3.6794	0.0062	Significant
Feature extraction using method, S.H. Shaikh, 2013	4.6918	0.0017	Significant

performs query classification to the nearest class by using the Euclidian distance similarity measure. The process of ranking is followed by retrieval of the top 20 images. The precision results for retrieval with classified query are compared to the



	Average precision	Average recall
M. Rahimi & M.E. Moghaddam (2013)	49.6	9.92
P.S. Hiremath & J. Pujari (2007)	54.9	10.98
H.A. Jalab (2011)	58.2	11.64
M. Subrahmanyam et al. (2012)	72.5	14.50
M. Banerjee et al. (2009)	72.7	14.54
G.L. Shenn & X.J. Wu (2013)	72.8	14.56
E. Walia et al. (2014)	75.8	15.16
Proposed	88.0	17.60

Fig. 13. Comparison of precision and recall for retrieval with proposed techniques with respect to state-of-the-art techniques.

state-of-the-art techniques for retrieval in Fig. 13.

It is observed from Fig. 13 that the retrieval performance of the proposed techniques outclasses those of the existing techniques. Hence, it is established that the proposed methods have made a noteworthy contribution for improvement in retrieval performance compared to the existing techniques.

IV. Conclusion

The authors have proposed four different techniques to facilitate multiview feature extraction from image data for content-based image recognition. The techniques were evaluated for classification and retrieval purposes. The classification process was carried out with two different techniques — classification by decision fusion with *Z*-score normalization and classification by feature standardization with *Z*-score normalization. The second method has surpassed the initial one in terms of classification performance. It has also outperformed the state-of-the-art techniques for classification with statistically significant improvements in precision values. Further, feature standardization with *Z*-score normalization was tested for the purposes of retrieval. A classification of a query image was performed as a precursor to the retrieval technique. The retrieval results with the proposed methods have outclassed all the existing techniques. This work can be applied to image data identification in the field of media, education, entertainment, surveillance, and so on.

References

[1] L. Ai et al., "High-Dimensional Indexing Technologies for Large Scale Content-Based Image Retrieval: A Review," *J. Zhejiang University-Sci. C*, vol. 14, no. 7, July 2013, pp. 505–520.

[2] J. Cao et al., "A Review of Object Representation Based on Local Features," *J. Zhejiang University-Sci. C*, vol. 14, no. 7, July 2013, pp. 495–504.

[3] A. Ahmadian and A. Mostafa, "An Efficient Texture Classification Algorithm Using Gabor Wavelet," *Int. Conf. IEEE Eng. Med. Biol. Soc.*, Cancun, Mexico, Sept. 17–21, 2013, pp. 930–933.

[4] R. Das and S. Bhattacharya, "A Novel Feature Extraction Technique for Content Based Image Classification in Digital Marketing Platform," *American J. Adv. Comput.*, vol. 2, no. 1, Jan. 2015, pp. 17–24.

[5] C.C. Aggarwal and P.S. Yu, "A Survey of Uncertain Data Algorithms and Applications," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 5, Mar. 2009, pp. 609–623.

[6] L. Manikonda, A. Mangalampalli, and V. Pudi, "UACI: Uncertain Associative Classifier for Object Class Identification in Images," *Int. Conf. Image Vis. Comput.*, Queenstown, New Zealand, Nov. 8–9, 2010, pp. 1–8.

[7] Y. Wang, "On Visual Semantic Algebra (VSA) and the Cognitive Process of Pattern Recognition," *IEEE Int. Conf. Cognitive Inf.*, Stanford, CA, USA, Aug. 14–16, 2008, pp. 384–393.

[8] W. Jun et al., "A New Approach for Classification of Fingerprint Image Quality," *IEEE Int. Conf. Cognitive Inf.*, Stanford, CA, USA, Aug. 14–16, 2008, pp. 375–383.

[9] S. Thepade, R. Das, and S. Ghosh, "Feature Extraction with Ordered Mean Values for Content Based Image Classification," *Adv. Comput. Eng.*, vol. 2014, Article ID 454876, Nov. 2014.

[10] S. Thepade, R. Das, and S. Ghosh, "Novel Technique in Block Truncation Coding Based Feature Extraction for Content Based Image Identification," in *Trans. Comput. Sci.*, Germany: Berlin Heidelberg, Springer, vol. 9030, Apr. 2015, pp. 55–76.

[11] S. Thepade, R. Das, and S. Ghosh, "A Novel Feature Extraction Technique with Binarization of Significant Bit Information," *Int. J. Imag. Robot.*, vol. 15, no. 3, 2015, pp. 164–178.

[12] S. Thepade, R. Das, and S. Ghosh, "Content Based Image Classification with Thepade's Static and Dynamic Ternary Block Truncation Coding," *Int. J. Eng. Res.*, vol. 4, no. 1, Jan. 2015, pp. 13–17.

[13] O.T. Yıldız, Ö. Aslan, and E. Alpaydın, "Multivariate Statistical Tests for Comparing Classification Algorithms," in *LNCS Learn. Intell. Optimization*, Germany: Berlin Heidelberg, Springer, vol.

http://dx.doi.org/10.4218/etrij.16.0115.0102

6683, 2011, pp. 1–15.
[14] M. Valizadeh et al., “A Novel Hybrid Algorithm for Binarization of Badly Illuminated Document Images,” *Int. CSI Comput. Conf.*, Tehran, Iran, Oct. 20–21, 2009, pp. 121–126.
[15] H.B. Kekre et al., “Multilevel Block Truncation Coding with Diverse Color Spaces for Image Classification,” *IEEE Int. Conf. Adv. Technol. Eng.*, Mumbai, India, Jan. 23–25, 2013, pp. 1–7.
[16] S. Thepade, R. Das, and S. Ghosh, “Image Classification Using Advanced Block Truncation Coding with Ternary Image Maps,” in *Adv. Comput., Commun. Contr.*, Germany: Berlin Heidelberg, Springer, vol. 361, 2013, pp. 500–509.
[17] S. Thepade, R. Das, and S. Ghosh, “Performance Comparison of Feature Vector Extraction Techniques in RGB Color Space Using Block Truncation Coding for Content Based Image Classification with Discrete Classifiers,” *India Conf.*, Mumbai, India, Dec. 13–15, 2013, pp. 1–6.
[18] C. Liu, “A New Finger Vein Feature Extraction Algorithm,” *IEEE Int. Congress Image Signal Proc.*, Hangzhou, China, Dec. 16–18, 2013, pp. 395–399.
[19] Y. Yanli and Z. Zhenxing, “A Novel Local Threshold Binarization Method for QR Image,” *Int. Conf. Automatic Contr. Artif. Intell.*, Xiamen, China, Mar. 3–5, 2012, pp. 224–227.
[20] M. Ramírez-Ortegón and R. Rojas, “Unsupervised Evaluation Methods Based on Local Gray-Intensity Variances for Binarization of Historical Documents,” *IEEE Int. Conf. Patttern Recogn.*, Istanbul, Turkey, Aug. 23–26, 2010, pp. 2029–2032.
[21] N. Otsu, “A Threshold Selection Method from Gray-Level Histogram,” *IEEE Trans. Syst., Man Cybern.*, vol. 9, no. 1, Jan. 1979, pp. 62–66.
[22] S.H. Shaikh, A.K. Maiti, and N. Chaki, “A New Image Binarization Method Using Iterative Partitioning,” *Mach. Vis. Appl.*, vol. 24, no. 2, Feb. 2013, pp. 337–350.
[23] M. Umaselvi, S.S. Kumar, and M. Athithya, “Color Based Urban and Agricultural Land Classification by GLCM Texture Features,” *IET Int. Conf. Sust. Energy Intell. Syst.*, Tiruchengode, India, Dec. 27–29, 2012, pp. 1–4.
[24] F. Mirzapour and H. Ghassemian, “Using GLCM and Gabor Filters for Classification of PAN Images,” *Iranian Conf. Electr. Eng.*, Mashhad, Iran, May 14–16, 2013, pp. 1–6.
[25] I. Muhammad, H. Rathiah, and A.K. Noor Elaiza, “Content Based Image Retrieval Using MPEG-7 and Histogram,” *Proc. Int. Conf. Soft Comput. Data Mining*, Universiti Tun Hussein Onn, Malaysia, June 16–18, 2014, pp. 453–465.
[26] I. Muhammad, H. Rathiah, and A.K. Noor Elaiza, “Color Histogram and First Order Statistics for Content Based Image Retrieval,” *Proc. Int. Conf. Soft Comp. Data Mining*, Universiti Tun Hussein Onn, Malaysia, June 16–18, 2014, pp. 153–162.
[27] L. Ogiela and M.R. Ogiela, “Semantic Analysis Processes in Advanced Pattern Understanding Systems,” *Proc. Int. Conf., AST*, Seoul, Rep. of Korea, Sept. 27–29, 2011, pp. 26–30.
[28] L. Ogiela, “Cognitive Informatics in Automatic Pattern Understanding and Cognitive Information Systems,” in *Adv. Cognitive Informat. Cognitive Comput.*, Germany: Berlin Heidelberg, Springer, vol. 323, 2010, pp. 209–226.
[29] H.B. Kekre et al., “Image Retrieval with Shape Features Extracted Using Gradient Operators and Slope Magnitude Technique with BTC,” *Int. J. Comput. Appl.*, vol. 6, no. 8, Sept. 2010, pp. 28–33.
[30] H.B. Kekre et al., “Image Retrieval Using Shape Texture Content as Row Mean of Transformed Columns of Morphological Edge Images,” *Int. J. Comput. Sci. Inf. Technol.*, vol. 2, no. 2, 2011, pp. 641–645.
[31] H.B. Kekre et al., “Performance Comparison of Full 2-D DCT, 2-D Walsh and 1-D Transform over Row Mean and Column Mean for Iris Recognition,” *Proc. Int. Conf. Workshop Emerging Trends Technol.*, Mumbai, India, 2010, pp. 202–205.
[32] M.E. El Alami, “A Novel Image Retrieval Model Based on the Most Relevant Features,” *Knowl.-Based Syst.*, vol. 24, no. 1, Feb. 2011, pp. 23–32.
[33] S. Thepade, R. Das, and S. Ghosh, “A Novel Feature Extraction Technique Using Binarization of Bit Planes for Content Based Image Classification,” *J. Eng.*, Article ID 439218, 2014.
[34] P.S. Hiremath and J. Pujari, “Content Based Image Retrieval Using Color, Texture, and Shape Features,” *Int. Conf. Adv. Comput. Commun*, Guwahati, India, Dec. 18–21, 2007, pp. 780–784.
[35] M. Banerjee, M.K. Kundu, and P. Maji, “Content-Based Image Retrieval Using Visually Significant Point Features,” *Fuzzy Sets Syst.*, vol. 160, no. 23, Dec. 2009, pp. 3323–3341.
[36] H.A. Jalab, “Image Retrieval System Based on Color Layout Descriptor and Gabor Filters,” *IEEE Conf. Open Syst.*, Langkawi, Malaysia, Sept. 25–28, 2011, pp. 32–36.
[37] G.L. Shen and X.J. Wu, “Content Based Image Retrieval by Combining Color Texture and CENTRIST,” *IEEE Int. Workshop Signal Process.*, London, UK, vol. 1, Jan. 25, 2013, pp. 1–4.
[38] M. Rahimi and M.E. Moghaddam, “A Content Based Image Retrieval System Based on Color Ton Distributed Descriptors,” *Signal Image Video Process.*, London, UK: Springer, vol. 9, no. 3, Mar. 2015, pp. 691–704.
[39] E. Walia, S. Vesal, and A. Pal, “An Effective and Fast Hybrid Framework for Color Image Retrieval,” *Sensing Img.*, USA: Springer, vol. 15, no. 1, Nov. 2014, pp. 92–116.
[40] M. Subrahmanyam, R.P. Maheshwari, and R. Balasubramanian, “Expert System Design Using Wavelet and Color Vocabulary Trees for Image Retrieval,” *Expert Syst. Appl.*, vol. 39, no. 5, Apr. 2012, pp. 5104–5114.
[41] K. Khurshid et al., “Comparison of Niblack Inspired Binarization Methods for Ancient Documents,” *Proc. SPIE 7247, Document Recogn. Retrieval* XVI, vol. 7247, Jan. 19, 2009.
[42] M. Ortega et al., “Supporting Ranked Boolean Similarity Queries

in MARS," *IEEE Trans. Knowl. Data Eng.*, vol. 10, no. 6, Nov. – Dec. 1998, pp. 905–925.

Rik Das received his MTech degree in Information Technology from the University of Calcutta, India, in 2006 and his BE degree in Information Technology from the University of Burdwan, India, in 2002. He has more than eleven years of teaching experience and one year of industry experience. He is currently pursuing his PhD degree in Information Technology from the University of Calcutta. He has received a certificate of appreciation for being a reviewer at the International Conference on Intelligent Systems and Data Processing (ICISD-2011) under MultiConf-2011. He is a reviewer for research articles with leading publishers such as IEEE Transactions, Springer, Elsevier etc. He is an editorial board member of an International Journal and has more than eleven international publications with leading publishers.

Sudeep Thepade received his PhD degree in computer engineering from SVKM's NMIMS, in 2011; his ME degree in computer engineering from the University of Mumbai, India, in 2008; and his BE degree in computers from North Maharashtra University, India, in 2003. He has eleven years of experience in teaching and in industry. Currently, he is both a dean and a professor of computer engineering at Pimpri Chinchwad College of Engineering, Pune, India. He is a member of the International Advisory Committee for many international conferences, acting as reviewer for many referred international journals/transactions including IEEE and IET. He has more than 135 papers in national/international conferences/journals to his credit.

Saurav Ghosh received his BE degree in instrumentation and electronics from Jadavpur University, Kolkata, India and his ME degree in computer science and engineering from Bengal Engineering College (deemed university), India. He started his academic career as a lecturer with the Department of Information Technology, Kalyani Government Engineering College, Kalyani, India. Later, he joined the University of Calcutta with the Department of A.K. Choudhury School of Information Technology, where he is working at present in the capacity of assistant professor. His research interests include wireless mesh networks, heterogeneous networks, and intrusion detection systems.