



New shape descriptor in the context of edge continuity

ISSN 2468-2322

Received on 5th January 2019

Revised on 20th April 2019

Accepted on 22nd April 2019

doi: 10.1049/trit.2019.0002

www.ietdl.org

Seba Susan ✉, Prachi Agrawal, Minni Mittal, Srishti Bansal

Department of Information Technology, Delhi Technological University, Bawana Road, Delhi 110042, India

✉ E-mail: seba_406@yahoo.in

Abstract: The object contour is a significant cue for identifying and categorising objects. The current work is motivated by indicative researches that attribute object contours to edge information. The spatial continuity exhibited by the edge pixels belonging to the object contour make these different from the noisy edge pixels belonging to the background clutter. In this study, the authors seek to quantify the object contour from a relative count of the adjacent edge pixels that are oriented in the four possible directions, and measure using exponential functions the continuity of each edge over the next adjacent pixel in that direction. The resulting computationally simple, low-dimensional feature set, called as 'edge continuity features', can successfully distinguish between object contours and at the same time discriminate intra-class contour variations, as proved by the high accuracies of object recognition achieved on a challenging subset of the Caltech-256 dataset. Grey-to-RGB template matching with City-block distance is implemented that makes the object recognition pipeline independent of the actual colour of the object, but at the same time incorporates colour edge information for discrimination. Comparison with the state-of-the-art validates the efficiency of the proposed approach.

1 Introduction

Object recognition has a significant role to play in the current scenario of real-time video surveillance and anomalous object detection. Unsupervised, real-time identification of the primary object in the scene in the midst of background clutter is the requirement [1]. The object recognition methods in literature can be categorised as edge-based or contour-based [2, 3], colour-intensity based [4, 5], local region or patch-based [6, 7], histogram-based [8, 9] and texture-based [10] techniques. Of all these methods, the edge-based and contour-based object recognition gives consistently good results since it relies on discontinuities within an image to distinguish objects, similar to the manner in which the human eye perceives the scene [11]. On the other hand, the texture-based and intensity-based methods seldom cover all possible scenarios due to a wide variety of shapes, colours, surfaces and sizes of the objects being possible that appear different under different lighting and background conditions [12]. Similar constraints are observed for the histogram-based object recognition, wherein it is hard to distinguish the information pertaining to the object alone from the image histogram that includes information on both the foreground and the background.

In this paper, we investigate feature representation of object contours by quantifying the continuity of edge pixels along the four principal directions, and evolve a set of eight 'edge continuity' features that describe effectively an object shape or pattern. The paper is organised as follows. Related work in literature and the motivation for our work are described in Section 2. The problem statement is defined in Section 3 followed by the formulation of the new feature set. The feature extraction process along with the overall classification methodology is explained in Section 4. The experimental results are discussed in Section 5 and the overall conclusions are drawn in Section 6.

2 Related work

Several popular image feature descriptors have been deployed by researchers to define the principal object in the scene in various

computer vision applications. Some examples are the use of local binary patterns (LBP) for defining edge-texture [2], scale-invariant feature transforms (SIFT) and speeded-up robust features (SURF) for localising objects in outdoor environment [13], rotation-invariant version of histogram of oriented gradients (HOG) features for detecting objects in very high-resolution remote-sensing images [14] and so on. In our previous work [15], we introduced a set of (high-dimensional) intensity cooccurrence-based temporal features for dynamic texture recognition from videos. These classic image descriptors generally are computationally costly, such as in the case of the locally defined SIFT where each keypoint is a 128-dimensional vector, or are sensitive to noise as in the case of the globally defined HOG that is sensitive to block noise. The local descriptor SURF is faster than SIFT though still not suitable for real-time applications, a common problem with all local descriptor feature sets. For distinguishing 3D objects from clutter, local surface features and surface matching were found useful in [16]. Local features that are informative on object shape include contour fragments that are processed by local classifiers at the bottom-most level of hierarchical learning models [17]. In another approach, the bag of contour fragments in conjunction with a linear support vector machine (SVM) classifier is used to identify shapes subject to deformation [18]. Histograms of locally detected features give efficient summarisation that can be used for recognition purpose. The HOG features proposed by Dalal and Triggs [19] are an example, being one of the most popular image descriptors that give best results when used in conjunction with the SVM. In HOG computation, the local edge orientations within image cells or windows are detected and a histogram of the edge orientations is computed within each local region for different angular bins. The histograms within a block are normalised and then these are concatenated to form the feature vector. Apart from being computationally heavy, these features are subject to block noise. In another work [20], edge maps are created from the local orientation of each edge pixel for identifying small irregular shaped targets, that coupled with the Hausdorff measure identify possible target locations. A similar approach is seen in [21], wherein the orientation of edges and their correlations between their neighbouring edges are used to describe and identify shapes

for image retrieval. The correlation between neighbouring edges is computed using an edge orientation autocorrelogram that records the number of similar edges at a certain orientation separated by an offset distance of k pixels apart, the k here being variable and the similarity between edges being decided by thresholds. The procedure is thus threshold-dependent, and also dependent on the choice of offset distance. On this line of thought, we propose a new low-dimensional feature set for object recognition that captures the information of edge-edge orientation and how much this orientation changes over the next adjacent edge (NAE), involving only three adjacent pixels in the computation. A summarisation of this locally retrieved information over all the edge pixels in the image would yield the final feature set.

3 New shape descriptor

3.1 Defining object contours by edge continuity

Edges indicate local discontinuities within an image. By linking coherently the adjacent edge (AE) pixels to form a continuous curve, we obtain the object contour [3]. Therefore, edge-based contour-finding techniques integrate the local edge information at a global level, combining the goodness of both aspects. A lot of local information tends to be noise. Distinguishing the useful edges that contribute to the object shape and contour from the more frequent image discontinuities that constitute background noise is the first step towards defining the principal object in the scene. Capturing the contour information from continuous edge pixels is the next step that relies on accurate and well-formed feature representations.

We define the object contour by answering the following two questions for every edge pixel E .

- in what direction does the AE pixel lie?
- by how much, does this direction change for the NAE pixel?

While AE lies among the immediate eight-neighbours of E specifically in the four directions of 0° , 45° , 90° and 135° , similarly NAE is one of the immediate eight-neighbours of AE specifically in the four directions of 0° , 45° , 90° and 135° . Fig. 1 shows this evaluation for an edge pixel E (marked in red). A summary of the answers to the two above questions for all edge pixels, is obtained by averaging in each of the four directions 0° , 45° , 90° , 135° for both AE and NAE. This yields a set of eight features that we call as the ‘edge continuity features’. We expect high edge continuity for zero or negligible direction change as shown in Fig. 2, and higher deviations in edge direction cease to matter since they all contribute to low edge continuity. To implement this non-linear response, we use a non-linear function to map the deviations in direction change for NAE. A decreasing exponential function $f(x) = e^{-x}$ transforms the linear direction change x into a non-linear space (refer Fig. 2). This non-linear transformation ensures that zero direction change results in the highest continuity factor of value 1, while large direction changes are interpreted as low edge continuity

$$F_1 = \frac{\text{COUNT}_{\text{AE}}(0^\circ)}{\text{COUNT}_{\text{AE}}(0^\circ) + \text{COUNT}_{\text{AE}}(45^\circ) + \text{COUNT}_{\text{AE}}(90^\circ) + \text{COUNT}_{\text{AE}}(135^\circ)} \quad (1)$$

$$F_2 = \frac{\text{COUNT}_{\text{AE}}(45^\circ)}{\text{COUNT}_{\text{AE}}(0^\circ) + \text{COUNT}_{\text{AE}}(45^\circ) + \text{COUNT}_{\text{AE}}(90^\circ) + \text{COUNT}_{\text{AE}}(135^\circ)} \quad (2)$$

$$F_3 = \frac{\text{COUNT}_{\text{AE}}(90^\circ)}{\text{COUNT}_{\text{AE}}(0^\circ) + \text{COUNT}_{\text{AE}}(45^\circ) + \text{COUNT}_{\text{AE}}(90^\circ) + \text{COUNT}_{\text{AE}}(135^\circ)} \quad (3)$$

$$F_4 = \frac{\text{COUNT}_{\text{AE}}(135^\circ)}{\text{COUNT}_{\text{AE}}(0^\circ) + \text{COUNT}_{\text{AE}}(45^\circ) + \text{COUNT}_{\text{AE}}(90^\circ) + \text{COUNT}_{\text{AE}}(135^\circ)} \quad (4)$$

3.2 Formulation of the new feature set

The proposed set of eight features that describe the object contour are therefore summarised as the four edge direction features and the four edge direction continuity features. The definition and formulae of the eight ‘edge continuity features’ as we call it, are given below:

(a) *The edge direction features:* The direction histogram for the image gives the count of AE pixels along the directions 0° , 45° , 90° , 135° . The counts are converted to probability values by dividing by the sum of all counts. The four edge direction features F_1 – F_4 are computed in (1)–(4).

(b) *The edge direction continuity features:* For each 0° edge-AE (E-AE) instance, let the angle between the AE-NAE be denoted by θ_{NAE} . For continuity in the edge orientation, it is desired that this angle be 0° or at least not deviate significantly from it. The magnitude of this deviation denoted by $|\theta_{\text{NAE}} - 0^\circ|$ describes the deviation in edge continuity. The negative exponential function transforms this deviation non-linearly, and a high value of 1 is achieved if $\theta_{\text{NAE}} \cong 0^\circ$. The same process is repeated for the E-AE orientations of 45° , 90° and 135° . In this manner, we obtain the four edge direction continuity features F_5 – F_8 for all the edge pixels in the image. We summarise these features by averaging over the count of all possible NAEs for each edge pixel

$$F_5 = \frac{\sum_{\text{NAE}, \forall E - \text{AE} = 0^\circ} e^{-\frac{|\theta_{\text{NAE}} - 0^\circ|}{135^\circ}}}{\text{COUNT}_{\text{NAE}} \forall E - \text{AE} = 0^\circ} \quad (5)$$

$$F_6 = \frac{\sum_{\text{NAE}, \forall E - \text{AE} = 45^\circ} e^{-\frac{|\theta_{\text{NAE}} - 45^\circ|}{135^\circ}}}{\text{COUNT}_{\text{NAE}} \forall E - \text{AE} = 45^\circ} \quad (6)$$

$$F_7 = \frac{\sum_{\text{NAE}, \forall E - \text{AE} = 90^\circ} e^{-\frac{|\theta_{\text{NAE}} - 90^\circ|}{135^\circ}}}{\text{COUNT}_{\text{NAE}} \forall E - \text{AE} = 90^\circ} \quad (7)$$

$$F_8 = \frac{\sum_{\text{NAE}, \forall E - \text{AE} = 135^\circ} e^{-\frac{|\theta_{\text{NAE}} - 135^\circ|}{135^\circ}}}{\text{COUNT}_{\text{NAE}} \forall E - \text{AE} = 135^\circ} \quad (8)$$

3.3 Suitability of the exponential function to measure edge continuity

Equations (5)–(8) can be defined as the weighted sums of exponentials that are non-additive by nature. The maximum value of each of these features is 1. The non-additive nature of the exponentials has been explored in [22] for the non-extensive entropy with Gaussian gain that identifies regular textures from sparse image data. Its exponential function has a quadratic exponent, and the low and sparse probabilities that constitute the texture data are confined in the bell of the Gaussian curve. In another work, the weighted average of this exponential non-extensive entropy computed from multiple frames was plotted over time to indicate anomalies (from the location of peaks) in a state of usual order such as the general motion of a crowd in a video [23]. Taking a cue from these indicative works, the weighted sum of exponentials defined in our work (5)–(8) represents strong edge continuity (approaching 1) when almost all or several of the individual exponentials in the summation approach 1, i.e. when the orientation between the edges E-AE-NAE is preserved (0°) for most of the edges E in the image. Since the response is heavily skewed to 1 for the 0° response, and more or less low for higher deviation angles, the decreasing

Angle between edge E and adjacent edge AE	Edge (red)-Adjacent edge (blue) $E-AE$ orientation	Edge Continuity over the next adjacent pixel NAE (black)	Edge Deviation by $\pm 45^\circ$ for the next adjacent pixel NAE (black)	Edge Deviation by $\pm 90^\circ$ for the next adjacent pixel NAE (black)	Edge Deviation by $\pm 135^\circ$ for the next adjacent pixel NAE (black)
0°					
45°					
90°					
135°					
Deviation: $\Delta\theta$		0°	$\pm 45^\circ$	$\pm 90^\circ$	$\pm 135^\circ$
Normalized deviation $\Delta\theta_N = \frac{ \Delta\theta }{135^\circ}$		0	$\frac{1}{3}$	$\frac{2}{3}$	1
Edge Continuity $e^{-\Delta\theta_N}$		$e^0 = 1$	$e^{-\frac{1}{3}} = 0.7165$	$e^{-\frac{2}{3}} = 0.5134$	$e^{-1} = 0.3679$

Fig. 1 Computation of edge continuity along each of the four directions 0° , 45° , 90° , 135° based on the orientation of the NAE pixel

exponential function with the linear exponent is justified for the non-linear mapping of edge-edge deviations.

3.4 Some examples of common object shapes

The eight-dimensional feature set F1–F8 given by (1)–(8) is now extracted for a variety of synthetic shapes and patterns, for testing its credibility as a distinguisher of object contours. Fig. 3 demonstrates the evaluation of the feature set for a variety of shapes – *circle*, *oval*, *rectangle*, *polygon*, *irregular shape*. We

observe that in the case of the *circle* and the *oval* in Figs. 3a and b, the direction-breakup of the edge-edge relationship F1–F4 shows a different and unique pattern with the *circle* being equi-directional; and the *oval* showing high readings in the horizontal direction (F1), weak and equal orientations along the two diagonals (F2, F4), and negligible reading along the vertical direction (F3). A similar pattern is observed in the edge direction continuity features (F5–F8) with the horizontal direction recording the highest edge continuity for the *oval*. The *rectangle* in Fig. 3c shows higher values along horizontal (F1) with almost zero readings along the diagonals (F2, F4). The continuity factors along

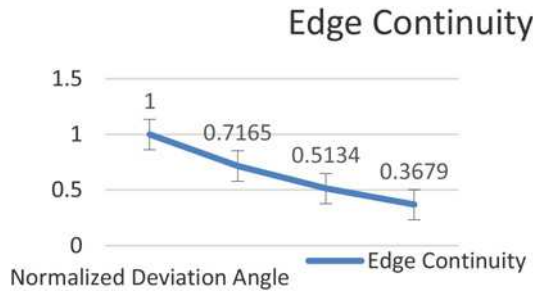


Fig. 2 Decreasing exponential function graph of edge continuity (refer Fig. 1) versus the normalised deviation $\Delta\theta_N$

the horizontal and vertical (F5, F7) are extremely high, >0.9 (maximum value for continuity factor is 1), higher than even the oval. The polygon in Fig. 3d and irregular shape in Fig. 3e, on the other hand, do not show very high edge direction readings along any one orientation.

4 Feature extraction steps and classification methodology

As for any pattern recognition process, the basic experimentation steps involved are that of feature extraction and classification.

4.1 Feature extraction steps

The feature extraction step consists of

1. Edge detection to highlight the discontinuities in the scene
2. Computation of 1×8 edge continuity feature vector

The Sobel edge detector: The Sobel is a first derivative based edge detection tool that is robust to noise and was proved reliable for contour formation in [21]. The Sobel operator is defined by the matrices in (9) and (10) for detecting horizontal and vertical edges, respectively, by convolving with overlapping 3×3 image windows

$$S_H = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad (9)$$

$$S_V = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad (10)$$

The input image is convolved with both matrices in (9) and (10) one after the other for detecting the horizontal and vertical edges. The edge continuity feature set (described in Section 3) is extracted

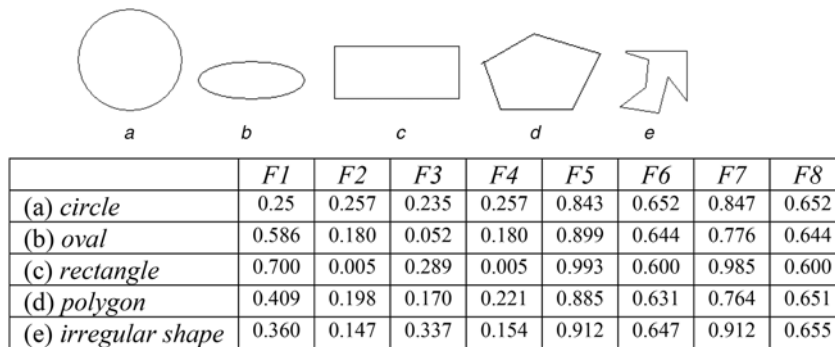


Fig. 3 Some instances of computed edge continuity features:

(a)–(e) Some common object shapes and their eight-dimensional feature set (F1–F8) computed by (1)–(8)

from the detected edges that are indicated by white pixels in a black background. Distance measures are next investigated for the classification part.

4.2 Colour-independent template matching

We propose grey-to-RGB template matching to improve the distance-based classification performance for object recognition. The grey-to-RGB comparison has been explored in [24] where the features extracted from the grey training image are compared with those of the RGB colour channels of the test image. A fuzzy classification process achieves this comparison in [24]. Colour invariance is achieved by this scheme since the object colours would stand out with respect to the background at least in one colour plane and would provide a better match with the grey training template. Colour invariance was achieved in [12] by averaging the R, G, B entropy features due to redundancy in entropy values across colour planes. In our work, we investigate various distance measures for the comparison of the grey training features with each of the test R, G, B colour channels' features. No averaging of R, G, B features is done since the colour features derived from each colour plane are unique and carry distinctive information of colour edges that may individually identify with the grey training template. Let $\text{Dist}(\text{Tr}, \text{Tst})$ be the distance function between two feature sets Tr and Tst derived from the grey training template and the colour test image, respectively. The formula for distance-based matching between the grey training template Tr and each of the R, G, B colour channels of the i th test image Tst_i is expressed as

$$\text{Dist}(\text{Tr}, \text{Tst}_i) = \min\{\text{Dist}(\text{Tr}, \text{Tst}_i^R), \text{Dist}(\text{Tr}, \text{Tst}_i^G), \text{Dist}(\text{Tr}, \text{Tst}_i^B)\} \quad (11)$$

Averaging the distance function over all or a subset of the training samples yields higher throughput as observed in [25]. In our experiments, the top K matches are retrieved for the test image from every training class, and the distance values are averaged over the top K matches for every class. The test is assigned the class of the training category it is closest to. This process is summarised in the following equation:

$$\text{Class}(\text{Tst}_i) = \arg \min_{\forall \text{Categories}} \left\{ \frac{\sum_{k=1}^K \text{Dist}(\text{Tr}_k, \text{Tst}_i)}{K} \right\} \quad (12)$$

A demonstration of this template matching scheme for our 'edge continuity' features is shown in Fig. 4 for the city-block distance function redefined for the eight features as

$$\text{Dist}(\text{Tr}, \text{Tst}) = \sum_j |\text{Tr}(j) - \text{Tst}(j)|, \quad j = 1, 2, \dots, 8 \quad (13)$$

The object shown in Fig. 4 is the *Harp* that has a strong vertical orientation as well as a high edge continuity along the vertical direction, as corroborated from the eight features F1–F8 shown computed in Fig. 4 for the grey training template. The R, G, B

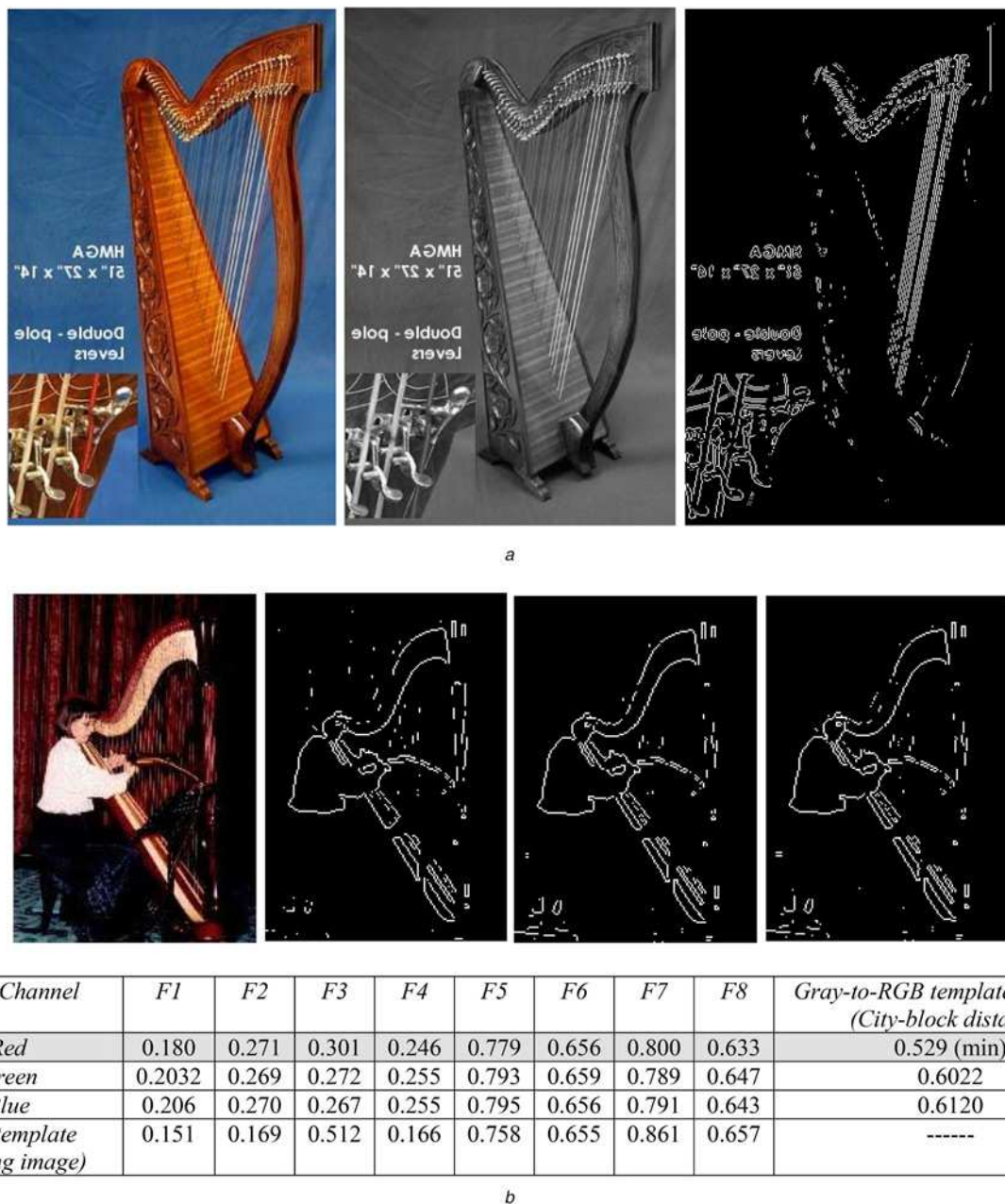


Fig. 4 Grey-to-RGB template matching shown for *a*

a Training template,

b Test image Red Green Blue images in the order from left to right (the reading corresponding to the best match with the grey template is shown highlighted in grey)

colour channels of a test instance of the *Harp* is shown in Fig. 4b, and the vertical edge orientation appears strong here as well. It is verified both visually and from the computed features F1–F8 that the Red (R) colour channel in Fig. 4b gives the best match with the grey training instance in Fig. 4a. The city-block distance shown computed for the three colour channels as per (13) is found minimum for the R channel as observed.

4.3 Object recognition pipeline

The detailed flowchart is presented in Fig. 5. The process is different for the training images and the test images. The grey images are retained for the training phase and the colour images are processed for the test images. The RGB colour space is used. The feature extraction stage is common for the grey template as well as each of the R, G, B colour channels. In the classification stage, the city-block distance is used to compare the grey training template with each of the R, G, B colour channels to find the best match.

The average distance of the test image from the top K closest matches in each category is used to determine the class of the test instance. In our experiments, $K=17$. We also conduct a sensitivity analysis for K in the later part of our experiments.

5 Experimental results and discussions

The experiments are performed in MATLAB 7.13 software on a Pentium computer clocked at 2.6 GHz. Our program just takes a fraction of seconds to execute for an average image in the dataset. Our code implementation of the proposed ‘edge continuity’ features is provided online at [26] for the reference of readers. We shortlisted a small subset of challenging objects, of varying image sizes, that have overlapping object shapes, to test for the worst-case scenario. A Caltech-256 dataset [27] subset containing the categories of *Bagpack*, *Bike*, *Binocular*, *Butterfly*, *Harp*, *Ketch* and *Starfish* was used for the experimentation, some

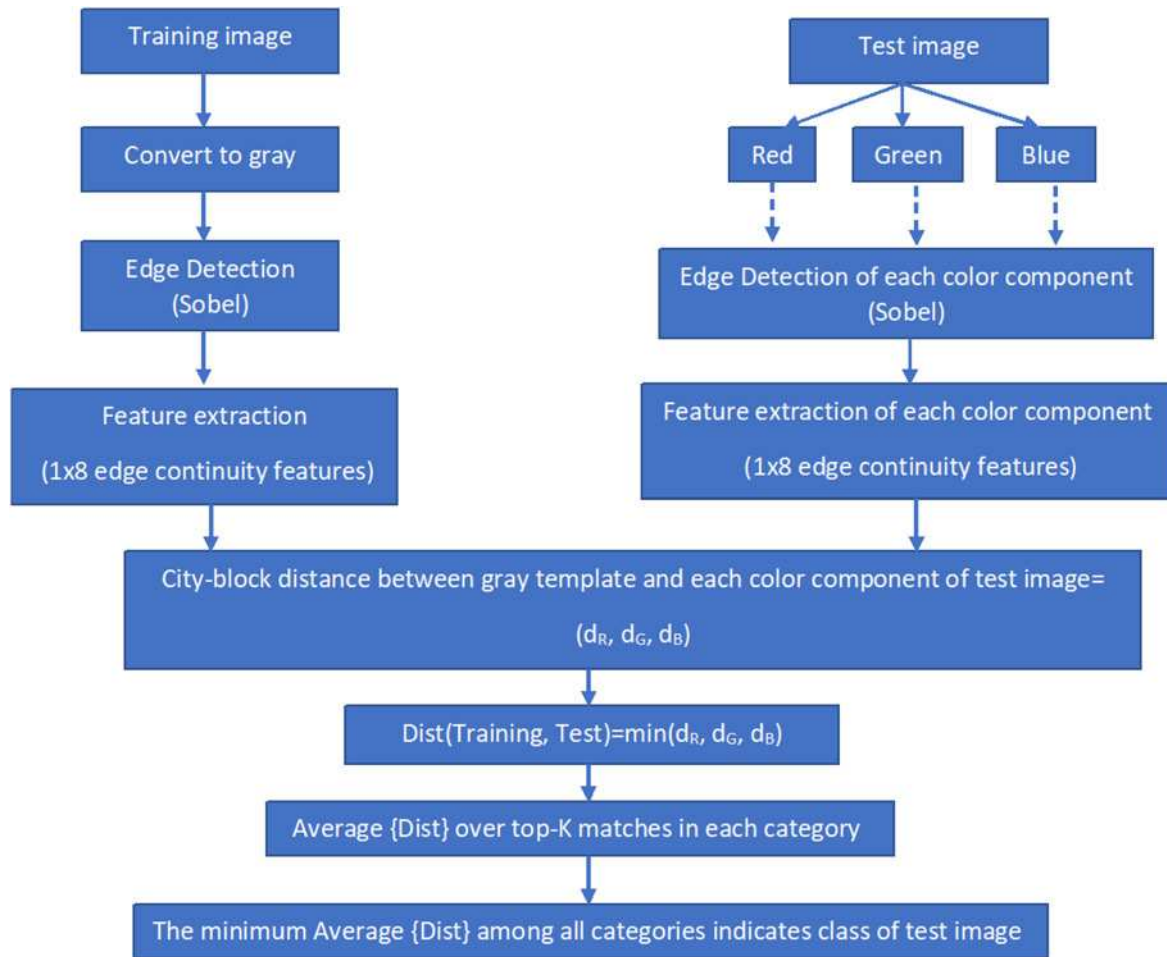


Fig. 5 Flowchart for the object recognition pipeline given a grey training template and a colour test image

samples of which are shown in Fig. 6. The categories were chosen such that there is an overlap of shapes and colours between categories such as among the *Bike*, *Butterfly*, *Starfish* and the *Harp* and also amongst *Bagpack*, *Ketch*, *Binocular* and the *Butterfly*. Sixty-eight images in the chronological order from each category were used, with alternate images chosen for training and testing that amounts to 476 images in all. In our first stage of experiments, we test our proposed schema (*sobel edge detection + feature extraction + grey-to-RGB template matching using city-block distance*) on the eight edge continuity features, as well as the image descriptors HOG [19] and LBP [28]. We average the distance as in (12) over 17 training samples that is 50% of the total number of training samples (=34). Thus, $K=17$ in (12). The confusion matrices are shown in the form of graphs in Fig. 7 for the proposed method.

The results in Table 1 and Fig. 7 show the highest accuracies for the *Ketch* class followed by the *Bike* category. The worst results, on a relative scale (failure instances), are observed for the *Binocular* category some instances of which get misclassified to the *Bagpack*, *Bike*, *Butterfly* and *Ketch* categories. Another failure instance is that of the *Butterfly* that sometimes gets misclassified as *Bike* image (refer Figs. 6 and 7 *Butterfly* graph). We next test affine-invariance and noise sensitivity on our feature set. The following transformations on the *Bagpack* image sample in Fig. 8 are shown: scaling, illumination change by contrast intensification, addition of salt and pepper noise, addition of Gaussian noise and rotation. The eight features $F1-F8$ computed for each of these transformations are shown in Table 2. We observe that scaling, illumination change and Gaussian noise do not affect the values of these features. Gaussian noise is handled well by the Sobel edge detection. However, the salt and pepper noise induces edge noise that is a set of short, abrupt and discontinuous intensities. Though

they cause a disturbance in the probability distribution of the direction histogram that records the orientation of the *E-AE* edge pair (features $F1-F4$), the edge direction continuity features ($F5-F8$) are undisturbed due to the noise being discontinuous. However, in the case of rotation, the features are different due to reordering of the orientations. A simple cyclic ordering of the eight features among $F1-F4$ and among $F5-F8$ yields the rotation-invariant version of our features as

$$\begin{aligned}
 &\{F1, F2, F3, F4, F5, F6, F7, F8\}_{RI} \\
 &= \{0^\circ, 45^\circ, 90^\circ, 135^\circ, 0^\circ, 45^\circ, 90^\circ, 135^\circ\} \text{ OR} \\
 &\{45^\circ, 90^\circ, 135^\circ, 0^\circ, 45^\circ, 90^\circ, 135^\circ, 0^\circ\} \text{ OR} \\
 &\{90^\circ, 135^\circ, 0^\circ, 45^\circ, 90^\circ, 135^\circ, 0^\circ, 45^\circ\} \text{ OR} \\
 &\{135^\circ, 0^\circ, 45^\circ, 90^\circ, 135^\circ, 0^\circ, 45^\circ, 90^\circ\}
 \end{aligned} \tag{14}$$

The computational time for the rotation-invariant version is increased to about four times which is still less than 1 s. It remains to be seen how the results are affected by this rotation invariance, since objects rotated from their usual form start resembling other objects and their rotated versions. The comparison of the rotated version of the feature set with its original form is shown in Fig. 9. The results indicate that the non-rotated (original) versions of images yield higher results than their rotated counterparts.

We compared our approach to some popular image descriptors used for object recognition namely SURF [29], HOG colour features with SVM [19], difference theoretic texture features (DTTF) with KNN classifier [28]. The results summarised in Table 3 indicate that experimentation on the challenging dataset gives the highest results for our method as compared to all other image descriptors. In order



Fig. 6 Caltech-256 categories of Backpack, Bike, Binocular, Butterfly, Harp, Ketch, Starfish (in the order from left to right and top to bottom) used for the experimentation

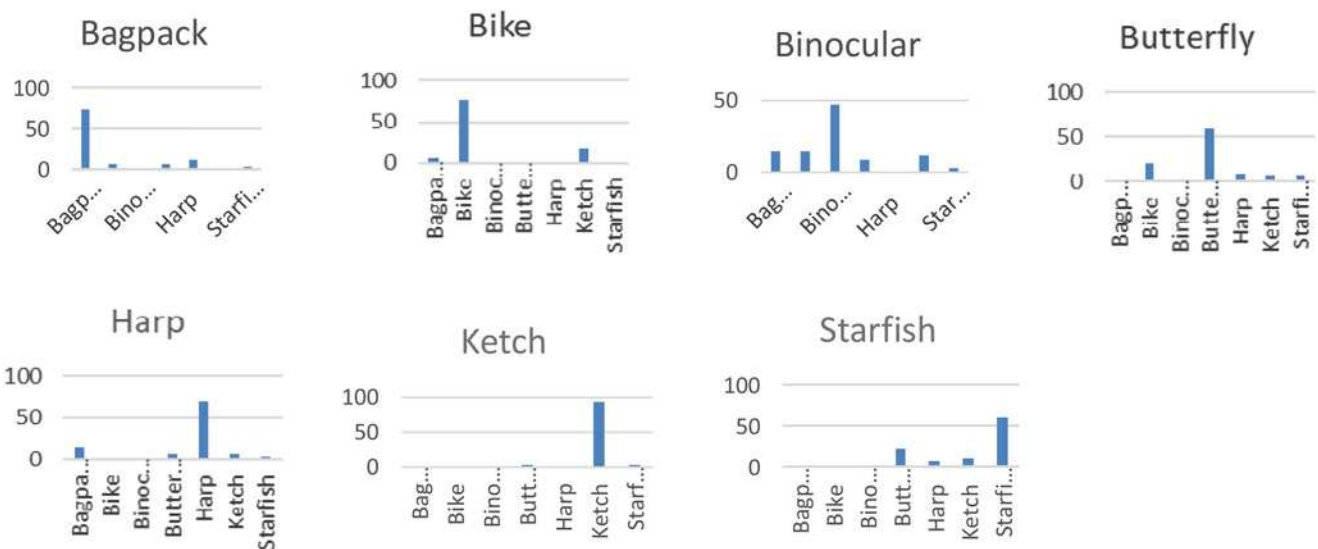


Fig. 7 Confusion matrices in the form of graphs that depict the distribution of classification accuracies over all categories for the proposed method

Table 1 Results of object recognition using the proposed scheme (Sobel edge detection + feature extraction + grey-to-RGB template matching using city-block distance) for various image features in terms of percentage accuracies (%) for the Caltech-256 dataset

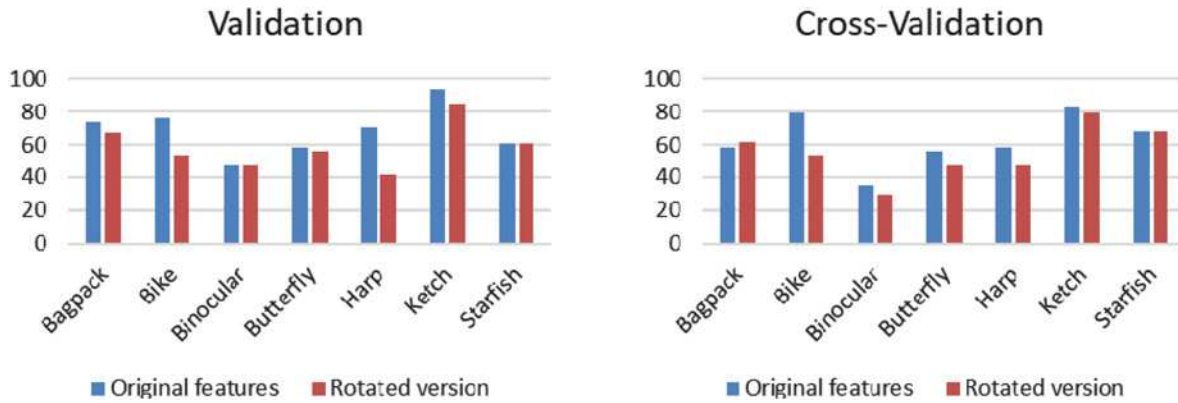
Class	Proposed edge continuity features (Dim:1 × 8)		HOG [19] (Dim: 1 × 20736)		LBP [28] (Dim:1 × 59)	
	V	CV	V	CV	V	CV
Backpack	76.47	64.71	2.94	0	32.35	41.18
Bike	73.53	79.41	79.41	94.12	35.29	23.53
Binocular	50	35.29	0	0	23.53	29.41
Butterfly	52.94	58.82	79.41	94.12	47.06	41.18
Harp	64.71	58.82	2.94	2.94	44.12	44.12
Ketch	94.12	82.35	8.82	2.94	88.24	82.35
Starfish	82.14	89.29	32.14	46.43	28.57	32.14
Average	70.56	65.93	29.38	34.36	42.73	42.03
time complexity (1200 × 1600 high-resolution image)	4.03 s		3.31 s		1.40 s	



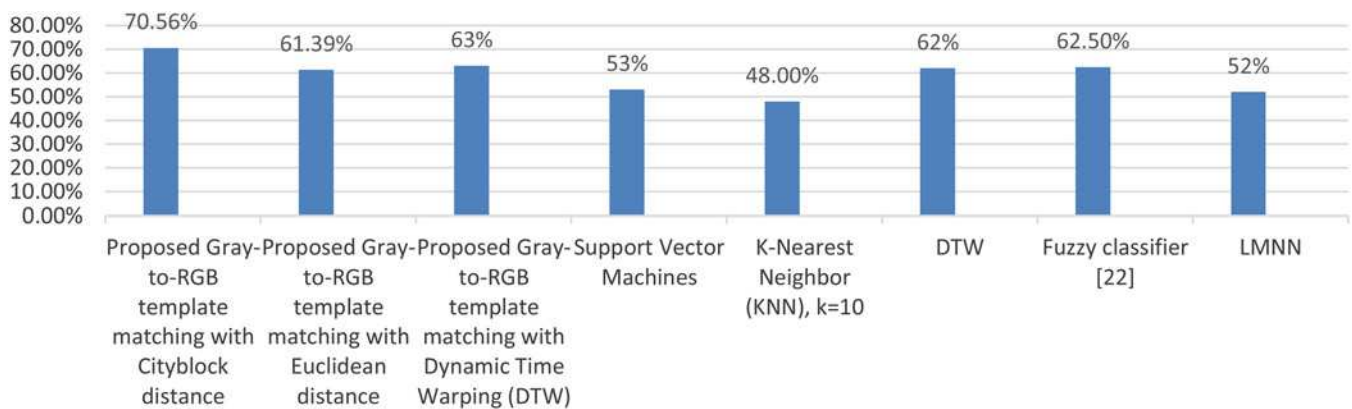
Fig. 8 Test image from Caltech-256 for testing affine invariance and robustness to noise on the proposed feature set

Table 2 Test for affine invariance and noise on the proposed feature set for the test image in Fig. 8

Transformation type	F1	F2	F3	F4	F5	F6	F7	F8
normal image	0.340	0.119	0.423	0.115	0.890	0.686	0.894	0.699
rotation	0.423	0.115	0.340	0.120	0.894	0.699	0.889	0.687
scaling	0.261	0.136	0.477	0.124	0.885	0.685	0.901	0.730
contrast-intensified image	0.218	0.124	0.532	0.124	0.872	0.707	0.919	0.723
noise (Gaussian)	0.256	0.135	0.471	0.137	0.868	0.695	0.893	0.699
noise (salt and pepper)	0.144	0.327	0.198	0.328	0.752	0.566	0.801	0.564

**Fig. 9** Performance comparison of the proposed feature set with its rotation invariant version shown for (a) Validation, (b) Cross-validation**Table 3** Results of object recognition using the proposed scheme (Sobel edge detection + feature extraction + grey-to-RGB template matching using the city-block distance classifier) versus various other methods in terms of percentage accuracies (%) for the Caltech-256 dataset

Class	Proposed edge continuity features (Dim:1×8)		HOG + SVM [19] (Dim: 1×20736)		DTTF + KNN [28] (Dim:1×11)		SURF [29] (Dim:1×59)	
	V	CV	V	CV	V	CV	V	CV
Backpack	76.47	64.71	64.71	52.94	41.18	23.53	55.88	47.05
Bike	73.53	79.41	79.41	91.18	52.94	76.47	55.88	82.35
Binocular	50	35.29	47.06	44.12	35.29	38.24	41.27	44.11
Butterfly	52.94	58.82	14.71	32.35	50	47.06	20.59	11.76
Harp	64.71	58.82	38.24	47.06	17.65	23.53	14.7	17.64
Ketch	94.12	82.35	91.18	88.24	79.41	58.82	17.64	14.71
Starfish	82.14	89.29	76.67	55.88	36.67	40	20	23.33
Average	70.56	65.93	58.85	58.82	44.73	43.95	32.38	34.42
time complexity (1200×1600 high-resolution image)	4.03 s		3.31 s		0.386 s		4.75 s	

**Fig. 10** Comparison of different classifiers for the proposed edge continuity features

to evaluate the performance of the grey-to-RGB template matching, we have included the comparison of our grey-to-RGB template matching classifier performance with that of SVM, KNN, LMNN,

fuzzy classifier for object recognition [24] and dynamic time warping (DTW) classifiers in Fig. 10. The higher accuracies clearly indicate the suitability of the grey-to-RGB template matching, using

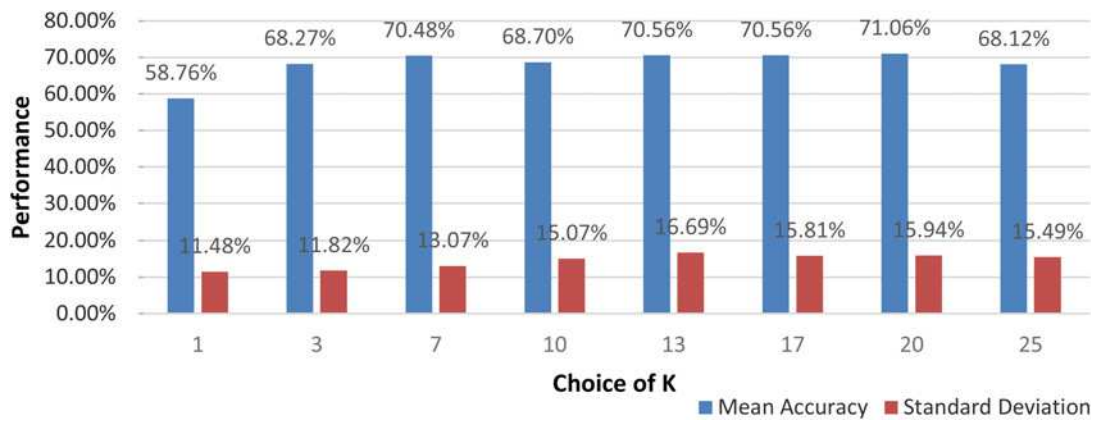


Fig. 11 Performance versus choice of minimum number of training samples (K) [refer (12)] for the proposed method

city-block distance, for object recognition from colour images using our proposed edge continuity features. Furthermore, a sensitivity analysis was conducted on the optimum choice of the minimum number of training samples per category (K in (12)), over which the cost function/distance should be averaged for achieving the best results in our method. The graphical comparison is shown in Fig. 11 for $K=1, 3, 7, 10, 13, 17, 20, 25$. Our choice of $K=17$ training samples (50% of the total) is found justified due to the high accuracy of 70.56% (~71%) obtained at this point. The graphical comparison in Fig. 11 shows the highest accuracy of 71.06% achieved for $K=20$ training samples per category, though for a higher standard deviation than at $K=17$. The time complexities of various methods, for a high-resolution image (*Harp* image), are shown in Tables 1 and 3, and these values indicate that our low-dimensional feature set has low computational complexity as well, comparable to the state-of-the-art.

6 Conclusions

A new set of edge continuity features are proposed for object recognition that capture the contour information from the edge pixels. Eight features are formulated based on the edge-edge orientation over three consecutive pixels, subsequently averaged over all the edge pixels. A grey-to-RGB template matching using city-block distance is proposed for our scheme. The new feature set is applied for object recognition from colour images, with successful results. Comparison with the state-of-the art on a challenging dataset reiterates the efficacy of our object recognition pipeline.

7 References

- [1] Zhang, S., Wang, C., Chan, S.-C., *et al.*: 'New object detection, tracking, and recognition approaches for video surveillance over camera network', *IEEE Sens. J.*, 2015, **15**, (5), pp. 2679–2691
- [2] Satpathy, A., Jiang, X., Eng, H.-L.: 'LBP-based edge-texture features for object recognition', *IEEE Trans. Image Process.*, 2014, **23**, (5), pp. 1953–1964
- [3] Zitnick, C.L., Dollár, P.: 'Edge boxes: locating object proposals from edges'. European Conf. on Computer Vision, Zurich, Switzerland, 2014, pp. 391–405
- [4] Gevers, T., Smeulders, A.W.M.: 'Color-based object recognition', *Pattern Recognit.*, 1999, **32**, (3), pp. 453–464
- [5] Van De Sande, K., Gevers, T., Snoek, C.: 'Evaluating color descriptors for object and scene recognition', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2010, **32**, (9), pp. 1582–1596
- [6] Chen, H., Bhanu, B.: '3D free-form object recognition in range images using local surface patches', *Pattern Recognit. Lett.*, 2007, **28**, (10), pp. 1252–1262
- [7] Ren, Z., Gao, S., Chia, L.-T., *et al.*: 'Region-based saliency detection and its application in object recognition', *IEEE Trans. Circuits Syst. Video Technol.*, 2014, **24**, (5), pp. 769–779
- [8] Gevers, T., Stokman, H.: 'Robust histogram construction from color invariants for object recognition', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2004, **26**, (1), pp. 113–118
- [9] Chang, P., Krumm, J.: 'Object recognition with color cooccurrence histograms'. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Fort Collins, Colorado, 1999, vol. 2, pp. 498–504
- [10] Ke, W., Zhang, T., Chen, J., *et al.*: 'Texture complexity based redundant regions ranking for object proposal'. Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops, Las Vegas, USA, 2016, pp. 10–18
- [11] Basu, M.: 'Gaussian-based edge-detection methods-a survey', *IEEE Trans. Syst. Man Cybernet. C (Appl. Rev.)*, 2002, **32**, (3), pp. 252–260
- [12] Susan, S., Hanmandlu, M.: 'Color texture recognition by color information fusion using the non-extensive entropy', *Multidimens. Syst. Signal Process.*, 2018, **29**, (4), pp. 1269–1284
- [13] Valgren, C., Lilienthal, A.J.: 'SIFT, SURF & seasons: appearance-based long-term localization in outdoor environments', *Robot. Auton. Syst.*, 2010, **58**, (2), pp. 149–156
- [14] Cheng, G., Zhou, P., Yao, X., *et al.*: 'Object detection in VHR optical remote sensing images via learning rotation-invariant HOG feature'. 2016 4th Int. Workshop on Earth Observation and Remote Sensing Applications (EORSA), Fort Collins, Colorado, 2016, pp. 433–436
- [15] Susan, S., Mittal, M., Bansal, S., *et al.*: 'Dynamic texture recognition from multi-offset temporal intensity co-occurrence matrices with local pattern matching', in Verma, N.K., Ghosh, A.K. (Eds.): 'Computational intelligence: theories, applications and future directions-volume II' (Springer, Singapore, 2019), pp. 545–555
- [16] Guo, Y., Bennamoun, M., Soheli, F., *et al.*: '3d object recognition in cluttered scenes with local surface features: A survey', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2014, **36**, (11), pp. 2270–2287
- [17] Lin, L., Wang, X., Yang, W., *et al.*: 'Discriminatively trained and-or graph models for object shape detection', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2015, **37**, (5), pp. 959–972
- [18] Wang, X., Feng, B., Bai, X., *et al.*: 'Bag of contour fragments for robust shape classification', *Pattern Recognit.*, 2014, **47**, (6), pp. 2116–2125
- [19] Dalal, N., Triggs, B.: 'Histograms of oriented gradients for human detection'. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 2005. CVPR 2005, San Diego, CA, USA, June 2005, vol. 1, pp. 886–893
- [20] Olson, C.F., Huttenlocher, D.P.: 'Automatic target recognition by matching oriented edge pixels', *IEEE Trans. Image Process.*, 1997, **6**, (1), pp. 103–113
- [21] Mahmoudi, F., Shanbehzadeh, J., Eftekhari-Moghadam, A.-M., *et al.*: 'Image retrieval based on shape similarity by edge orientation autocorrelation', *Pattern Recognit.*, 2003, **36**, (8), pp. 1725–1736
- [22] Susan, S., Hanmandlu, M.: 'A non-extensive entropy feature and its application to texture classification', *Neurocomputing*, 2013, **120**, pp. 214–225
- [23] Susan, S., Hanmandlu, M.: 'Unsupervised detection of nonlinearity in motion using weighted average of non-extensive entropies', *Signal. Image. Video. Process.*, 2015, **9**, (3), pp. 511–525
- [24] Susan, S., Chandna, S.: 'Object recognition from color images by fuzzy classification of gabor wavelet features'. 2013 5th Int. Conf. on Computational Intelligence and Communication Networks (CICN), Mathura, India, 2013, pp. 301–305
- [25] Susan, S., Sharma, S.: 'A fuzzy nearest neighbor classifier for speaker identification'. 2012 Fourth Int. Conf. on Computational Intelligence and Communication Networks (CICN), Mathura, India, 2012, pp. 842–845
- [26] <https://in.mathworks.com/matlabcentral/fileexchange/71160-edge-continuity-features>, Accessed online on 9th April 2019
- [27] Griffin, G., Holub, A., Perona, P.: 'Caltech-256 object category dataset', 2007
- [28] Susan, S., Hanmandlu, M.: 'Difference theoretic feature set for scale-, illumination-and rotation-invariant texture classification', *IET Image Process.*, 2013, **7**, (8), pp. 725–732
- [29] Bay, H., Tuytelaars, T., Van Gool, L.: 'Surf: speeded up robust features'. European Conf. on computer vision, Graz, Austria, 2006, pp. 404–417