Machine Learning Approach for Object Recognition

V. N. Pawar and S. N. Talbar

***Abstract*—Object recognition is basically invariant to the dramatic changes caused in objects' appearance such as location, size, viewpoint, illumination, occlusion and more by the variability in viewing conditions. In this paper, we employ an efficient approach for object recognition using invariant features and machine learning technique. The invariant features namely color, shape and texture invariant features of the objects are extracted separately with the aid of suitable feature extraction techniques. In the proposed approach, we integrate the color, shape and texture invariant features of the objects at the feature level, so as to improve the recognition performance. The fused feature set serves as the pattern for the forthcoming processes involved in the proposed approach. We employed the pattern recognition algorithms, like Discriminative Canonical Correlation (DCC) and attain distinct or identical results concerned with false positives. Our proposed approach is evaluated on the ALOI collection, a large collection of object images consists of 1000 objects recorded under various imaging circumstances. The experiments clearly demonstrate that our proposed approach significantly outperforms the state-of-the- art methods for combining color, shape and texture features. The proposed method is shown to be effective under a wide variety of imaging conditions.**

***Index Terms*—Computer vision, objects recognition, feature extraction, color, shape, texture, discriminative canonical correlation (DCC).**

I. Introduction

Generally in a cluttered real-world scene, identifying a correspondence between a 3-D object and some part of a 2-D image taken from an arbitrary viewpoint serves as the major task in object recognition. Practically, in many cases object recognition systems are appearance- or model- based. To identify the model (image) corresponding the portion of the target image [1], the recognition module consists of a matching process in which the stored models (model-based) or images (appearance-based), encapsulated in the representation scheme is matched against the target image. Recently, Generic object detection and recognition have gained quite various focus of attraction in computer vision [2]-[6]. Generic object recognition strategies enterprise to recognize objects based upon coarse, protypical representations by considering possible variability's of the object appearance.

In this paper, we present a new strategy for object recognition with invariant features and machine learning techniques. Here we employ an approach for object recognition with three different phases namely feature extraction, feature level fusion, Recognition. In the proposed approach, we integrate the color, texture and shape invariant features of the objects at the feature level, so as to improve the recognition performance. At first the invariant features namely color, shape and texture invariants of the objects are extracted independently using proper feature extraction techniques. For color invariant features, with help of photometric invariants of the object, we have derived a set color descriptors. Consequently, for shape invariant features we employed variational level set formulation and convex hull method. Finally for texture invariant, features are derived with the deliverance of haralick texture features. By providing a Color only, shape only and texture only representation, we focus attention on how they can be combined together as they are, rather than defining a new representation. Here we employ a simple concatenation procedure to integrate the extracted features together. The fused feature set serves as the pattern for the forthcoming processes involved in the proposed approach. The object recognition process is employed by the pattern recognition algorithms, like Discriminative Canonical Correlation (DCC). We are likely to attain distinct or identical results concerned with false positives.

II. Feature Extraction

Extracting features of interest from large and possibly heterogeneous data sets is a crucial task because it has to face many communities of end-users. Feature extraction is a must in object recognition, so as to locate those pixels in an image/object that have some distinctive characteristics. Typically those characteristic poses some inhomogeneity especially in local image properties such as Color, Shape (Geometric) and Textures.

A. Color Invariant Features

The changes in the color of the light source (e.g. Spectral Power Distribution (SPD) and object-illumination geometry (e.g. angle of incidence and reflectance) automatically reflects in the color of an object. Consequently, with the time-of-day, cloud cover and other atmospheric conditions, the color of the illuminant (i.e. daylight) varies in outdoor images. As a result, the color of an object may change drastically due to these variations in imaging conditions. In Gevers and Smeulders method [7], in order to recognize an object, different color models are proposed which show some degree of invariance. To achieve a broad usage of the color descriptor, it should be robust to: 1. photometric changes frequently identified in the real world, 2. Difference in image quality, from high quality images to snap-shot photo quality and compressed internet images [8]. In this paper, we develop a set of color descriptors, which are more consistent under photometric and geometrical changes, and reliable under decreasing image quality. By applying nonlinear diffusion,

Manuscript received July 14, 2012; revised August 17, 2012.

V. N. Pawar is with A. C. Patil College of Engineering, Navi Mumbai, Maharashtra, India 410206 (e-mail: vnpawar2000@ gmail.com).

S. N. Talbar is with S. G. G. S. Institute of Engineering and Technology, Nanded, Maharashtra, India 431606 (e-mail: sntalbar@yahoo.com).

initially the color image is diffused. Nonlinear diffusion filtering is a module, based on nonlinear evolution partial differential equations, which inquires to increase the quality of images, by removing noise while retaining details and constantly reinforcing edges [9].

On the basis of the photometric robustness, geometric robustness, photometric stability and generality, we procure a set of photometric invariant color histograms which can often be used as color descriptors after the object/image is being diffused using a nonlinear diffusion. A method is being used as a global image feature by combining illuminant and geometric invariant [10] and is called as comprehensive color image normalization (CCIN), and in this we have applied it as a local image feature. An iterative usage is being utilized in this method and thereby it is invariant for both lighting geometry and illuminant color.

B. Shape Invariant Features

In this section, shape invariant descriptors are discussed in order to express geometric properties of a set of coordinates of an image object free from a specific coordinate transformation. In particular, an image (or shape) invariants depend based on the shape of an object. In our work, we employ variational level set formulation of active contours and convex hull method for geometric invariant extraction.

Initially we utilized a variational formulation to completely eliminate the need of the costly re-initialization procedure, for geometric active contours [11] which force the level set function to be close to a signed distance function. The variational formulation comprises of two energy terms namely an internal energy and an external energy. An internal energy penalizes the deviation of the level set function from a signed distance function; on the other hand, the later one drives the motion of the zero level set toward the desired image features, such as object boundaries. Beyond the traditional level set formulations, the used variational level set formulation [11] has three significant advantages. They are as follows.

1) To numerically solve the evolution PDE, a significantly larger time step can be utilized and that automatically speeds up the curve evolution.
2) Comparatively, with signed distance function, the level set function should be initialized as functions that are computationally more efficient to generate.
3) Rather using complex upwind scheme as in traditional level set formulations, the used level set evolution can be implemented using simple finite difference scheme.

Variational Level Set Formulation of Active Contours without Re-initialization:

In image segmentation, active contours are represented as dynamic curves that move towards the object boundaries. In order to achieve this objective, we explicitly define an external energy that can move the zero level curves toward the object boundaries [11]. Let I be an image, and g be the edge indicator function defined by

$$g = \frac{1}{1 + |\nabla G_\sigma * I|^{2*}} \tag{1}$$

where G_σ is the Gaussian kernel with standard deviation σ. We describe an external, for a function $\phi(x, y)$ as below.

$$\in_g, \lambda, \nu(\phi) = \lambda L_g(\phi) + \nu A_g(\phi) \tag{2}$$

where $\lambda > 0$ and ν are constants, and the terms $L_g(\phi)$ and $A_g(\phi)$ are defined by

$$L_g(\phi) = \int_\Omega g^\delta(\phi)|\nabla\phi|\, dxdy \tag{3}$$

and

$$L_g(\phi) = \int_\Omega gH(-\phi)\, dxdy \tag{4}$$

likewise, where δ is the univariate Dirac function and H is the Heaviside function. Now, we define the subsequent total energy functional

$$\varepsilon(\phi) - \mu P(\phi) + \varepsilon_{g,\lambda,\nu}(\phi) \tag{5}$$

The external energy $\varepsilon_{g,\lambda,\nu}$ derives the zero level set toward the object boundaries, although the internal $\mu P(\phi)$ penalizes the deviation of ϕ from a signed distance function throughout its evolution.

Once we have identified all contour points of the object, then we search for convex hull which acquiesce the shape invariant features.

Convex Hull

The convex hull is defined for any kind of objects made up of points in a vector space that may perhaps have any number of dimensions, including infinite-dimensional vector spaces [12]. The convex hull of *X* is constructively described as the set of convex combinations of finite subsets of points from *X*: that is, the set of points of the form $\sum_{j=1}^{n} t_j x_j$, where n is an arbitrary natural number, the numbers t_j are non-negative and sum to 1, and the point's x_j are in *X*. It is uncomplicated to verify that this set assure either of the two definitions above. So the convex hull $H_{convex}(X)$ of set *X* is:

$$H_{convex}(X) = \left\{ \sum_{i=1}^{k} \alpha_i x_i \mid x_i \in X, \alpha_i \in \Re, \alpha_i \geq 0, \sum_{i=1}^{k} \alpha_i = 1, k = 1,2,.... \right\} \tag{6}$$

In fact, if *X* is a subset of an N-dimensional vector space, convex combinations of at most $N+1$ points are adequate in the definition above. This is similar to saying that the convex hull of *X* is the union of all simplexes with at most $N+1$ vertices from *X*.

C. Texture Invariant Features

Texture is a significant feature because the images can be considered as the composition of different texture regions. Dislike shape-based objects, Texture-based objects are one for which, there is no obvious visible inter-object part-wise correspondence. These objects are better detailed by their texture rather than the geometric structure of reliably detectable parts. There are numerous techniques for texture feature extraction. In this paper, we employed Haralick Texture feature calculation. These features capture information regarding the patterns that emerge in patterns of texture. Haralick texture feature calculation can be broken down into two parts or modules; (1) the construction of the co-occurrence matrices and (2) the calculation of the 13 texture features based on the co-occurrence matrices.

Co-occurrence Matrix

In an image, a co-occurrence matrix *P* is used to describe the patterns of neighboring pixels at a given distance d. In the computation of the texture features, 4 such matrices are required to depict different orientations. In particular, single co-occurrence matrix explains pixels that are adjacent to one another horizontally P^0. Moreover, there is a co-occurrence matrix for the vertical direction and diagonally in both directions. The co-occurrence matrices are symmetric matrices amid the dimensionality of $N_g \times N_g$ where N_g is the number of possible gray levels for a particular image. For every layer of d, *a* few statistical properties can be defined [13]. The 13 Haralick texture features are detailed in [14].

III. Feature Level Fusion

Feature-based fusion at feature level needs an extraction of objects recognized in the various data sources. It necessitates the extraction of salient features that depend on their environment such as Color, shape and textures. Fusion at this level involves the integration of feature sets, subsequent to multiple information sources. The extracted feature sets can be fused to create a new feature set to represent individual. The feature set has richer information regarding the raw object data that expects integration at this level, to provide better recognition results.

In this work, feature level fusion is proficient by a simple concatenation of the feature sets acquired from multiple information sources. Let $P = \{p_1, p_2, p_3, \ldots\ldots\ldots p_l\}$, $Q = \{q_1, q_2, q_3, \ldots\ldots\ldots q_m\}$ and $R = \{r_1, r_2, r_3, \ldots\ldots\ldots r_n\}$ indicate feature vectors $(P \in X^l, Q \in X^m and\ R \in X^n)$ representing the information extracted via three different sources. The aim is to combine these three feature sets in order to capitulate a new feature vector, S, to facilitate a better representation of the individual. The vector *S* is generated by first augmenting vectors *P*, *Q* and *R*, and then performing feature selection on the resultant feature vector. The concatenated feature set exhibits better discrimination capability than the individual feature vectors obtained from Color, Shape and Texture of the Objects individually.

IV. Recognition Technique

The recognition technique employed in our proposed object recognition system is detailed below.

A. Discriminative Canonical Correlation

In 1936, Hotelling [15] was the first to introduce Canonical Correlation Analysis (CCA). This technique detects the degree of intersections between two linear subspaces. From a geometric point of view, the correlation of the transformed data is maximized, by rotating the two coordinate frames simultaneously using CCA. Finding the rotation matrices is essentially equal to seek linear transformations of variables [16].

CCA seeks two projection vectors which yield a maximum correlation in a new coordinate system [17] for a given pair of images $x \in R^n$ and $y \in R^n$. Initially, the canonical correlation between two projected vectors can be expressed as

$$\rho = \max_{w_x, w_y} \frac{Cov(w_x^T x, w_y^T y)}{\sqrt{\mathrm{var}(w_x^T x)\,\mathrm{var}(w_y^T y)}} = \max_{w_x, w_y} \frac{w_x^T C_{xy} w_y}{\sqrt{(w_x^T C_{xx} w_x)(w_y^T C_{yy} w_y)}} \tag{7}$$

where C_{xy} is xy^T, C_{xx} is xx^T, and C_{yy} is yy^T. Equation (7) can be rewritten as

$$\rho = \max_{w_x, w_y} w_x^T C_{xy} w_y \tag{8}$$

Subject to $w_x^T C_{xx} w_x = 1, \quad w_y^T C_{yy} w_y = 1$

Equation (8) can then be reformulated as a Lagrangian formulation exposed as

$$L(w_x, w_y, \lambda_x, \lambda_y) = w_x^T C_{xy} w_y + \lambda_x (1 - w_x^T C_{xx} w_x) + \lambda_y (1 - w_y^T C_{yy} w_y) \tag{9}$$

where λ_x and λ_y are the Lagrangian multipliers. By differentiating L with respect to w_x and w_y, the solution of Equation (9) can be a more generalized Eigen-system described as follows:

$$\begin{bmatrix} 0 & C_{xy} \\ C_{yx} & 0 \end{bmatrix} \begin{bmatrix} w_x \\ w_y \end{bmatrix} = \Lambda \begin{bmatrix} C_{xx} & 0 \\ 0 & C_{yy} \end{bmatrix} \begin{bmatrix} w_x \\ w_y \end{bmatrix} \tag{10}$$

where Λ is the canonical correlation.

Generalizing for a set of images, let $X = [x^{(1)} \cdots x^{(p)}]$ and $Y = [y^{(1)} \cdots y^{(q)}]$ be two image sets where the number of samples for a data set *X* and *Y* are *P* and *q*, respectively. The combined data samples can be symbolized as $Z = [X \mid Y]^T$. The covariance matrix, $Cov(Z) = ZZ^T$, can then be written as a block matrix:

$$Cov(Z) = \begin{bmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{bmatrix} \tag{11}$$

where $C_{xy} = C_{yx}^T$. Note that C_{xy} represents the covariance between sets, whereas C_{xx} and C_{yy} represent the covariance within sets. Applying the Eigen-solution from Equation (10), the canonical correlations between data set *X* and data set *Y* are embedded in the diagonal elements of Λ [17].

V. The Proposed Approach for Object Recognition

Pattern recognition refers to the act of taking in raw data and taking an action based on the category of the data. Recognizing objects through perception is an important part of our lives: we recognize people while talking with them; we recognize a cup placed on the breakfast table, our car parked in a lot, and so on. Although this task is performed with great accuracy and apparent little effort by humans, it is still unclear to note that how this performance is achieved. Object recognition is one of the most fundamental issues in image processing from the human visual perspective and it has been a most important objective. The two major category of methodologies for object recognition are: 1) feature based, which make use of spatial arrangements of extracted features such as edge elements or junctions and 2) brightness-based (or appearance-based), which directly uses the pixel brightness. Generally, instead of focusing on the shape of the occluding contour or other extracted features, brightness-based methods make direct use of the gray values within the visible portion of the object, on the other hand feature based methods make use of the shape and texture properties of the objects. Current approaches in computer vision make use of either color, texture or shape properties of the objects for recognition. An extensive survey over the literature shows that the recognition performance of the methods could be enhanced by incorporating some invariant features namely color, texture and shape groupings in the same approach.

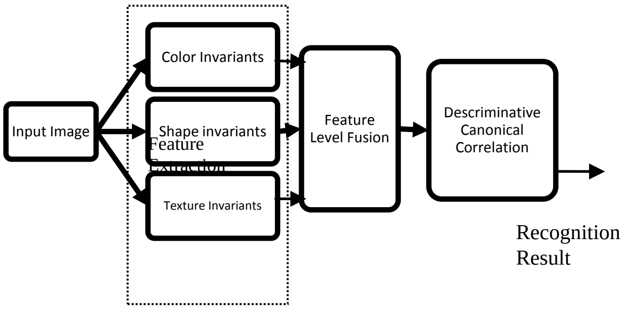

Fig. 1. Block diagram of the proposed approach for object recognition

Object recognition system used in the paper can be summarized as shown in Fig.1. The proposed research is to intrigue an efficient approach for object recognition using invariant features and machine learning technique. The proposed object recognition system follows three phases namely; (i) Feature Extraction (ii) Feature level fusion (iii) Object recognition using machine learning technique. In order to improve the recognition performance, at the feature level, the proposed system has integrated the following properties of the object namely color, texture and shape. An image of the object to be recognized is fed as input to the proposed approach. At first, the invariant features namely Color, Shape and Texture invariant features of the image are extracted independently using appropriate feature extraction techniques. For color invariant features, initially the object is diffused with the aid of nonlinear diffusion and with assistance of photometric invariants of the object, we have derived a set color descriptors. Subsequently, for shape feature we employed variational level set formulation and convex hull method. The variational level set method can be easily implemented by using simple finite difference scheme and is computationally more efficient than the traditional level set methods. In this method, significantly larger time step can be used to speed up the curve evolution, while maintaining stable evolution of the level set function. Moreover, the level set function is no longer required to be initialized as a signed distance function. Finally for texture invariants are derived with the deliverance of haralick texture features. These features capture information about the patterns that emerge in patterns of texture. The 13 features include are angular second moment, contrast, correlation and the remaining follows a variety of entropy measures, and these are calculated followed by the construction of a co-occurrence matrix that is by tradition computationally expensive.

Once the features of the object are derived independently, we intended to perform feature level fusion. Here feature level fusion is accomplished with the aid of simple concatenation of possibly three extracted features namely color, shape and texture which are expected to provide better recognition results. These combined feature set, serves as a pattern for the forthcoming processes involved in our proposed approach. After that we have the process of pattern recognition Discriminative Canonical Correlation (DCC), with this we are likely to gain distinct or identical results concerned with recognition rate and false positives.

VI. Experimental Results and Discussion

In this section, several experiments are carried out to demonstrate the effectiveness of our proposed Object recognition based on approaches. The experiments are conducted in different levels such as feature extraction, Feature level Fusion, Recognition and decision under varying conditions. Performance is evaluated with an object recognition task on the Amsterdam Library of Object Images (ALOI) dataset [18], [19]. This public database has been widely used in object recognition literature. The ALOI collection consists of 1000 objects recorded under various imaging circumstances. For each object the viewing angle, illumination angle, and illumination color are varied. The ALOI database consisting of 72000 images of 1000 different objects with 72 diverse views each. The sizes of the images are 256x256 with 8 bits per color. The images show a considerable amount of shadows, shading, and highlights. The combination of a large image dataset with a considerable variety of appearance, offers a formidable challenge for object recognition. Object recognition is the problem of matching one appearance of an object against a standardized version. One object may give rise to millions of different images, as camera conditions may be assorted endlessly. In our recognition experiment, one prototypical version of each object in the ALOI dataset is indexed and for querying, the diversity of the recorded object variations in the collection is being used. An object is accurately recognized when for all different variations the correct indexed object is returned. In this case, one may assume that the object can be recognized under a wide variety of real-life imaging circumstances. The objects of the data set are presented in Fig. 2.

Fig. 2. Sample objects from the ALOI database

A. Color Invariant Features:

In Photometric Invariants Extraction, we derive a set of photometric invariant color histograms based on photometric robustness, geometric robustness, photometric stability and generality that are used as a color descriptor. Initially the image is diffused with the aid of anisotropic diffusion and after that convert the image into different number of patches. Subsequently, for each patch color descriptors are being computed. To accomplish that initially color normalization is applied. Next the color descriptors, being the histograms of rgb, hue, Sat are computed, Where as r g b and hue are computed as follows.

$$r = \frac{R}{R+G+B}\ ,\quad g = \frac{G}{R+G+B}\ ,\quad b = \frac{B}{R+G+B} \tag{12}$$

$$hue = \arctan\left(\frac{\sqrt{3}(b^R - b^G)}{(b^R + b^G - 2b^G)}\right) \tag{13}$$

The photometric invariant features such as Saturation, Hue and r g b values computed for some sample patches of sample class object from ALOI database is depicted in table 1.

TABLE I: PHOTOMETRIC INVARIANTS (10 SAMPLE PATCHES FROM EACH OBJECT) FROM ALOI DATABASE

Saturation	Hue	rgb
0.101289	4.405676	0.063805,-0.047281,0.033477
0.101337	4.408993	0.020260,0.000853,0.028887
0.101491	4.419952	0.031550,0.063443,-0.044993
0.101289	4.526919	0.037678,-0.036760,0.049082
0.101495	4.511829	0.013007,0.014787,0.022205
0.101566	4.462959	0.013394,0.011369,0.025237
0.101700	4.448676	-0.018203,0.033160,0.035043
0.101681	4.469517	0.178874,0.015836,-0.144709
0.101660	4.468705	-0.007702,-0.009512,0.067214
0.101299	4.456538	-0.003627,0.010705,0.042923

B. Shape Invariant Features:

The proposed variational level set method has been applied to a variety of images from ALOI database in different modalities. A new variational formulation for geometric active contours that forces the level set function to be close to a signed distance function, and therefore completely eliminates the need of the costly re-initialization procedure. We used a region-based initialization of level set function, which is not only computationally more efficient than computing signed distance function, but also allows for more flexible applications. Our variational formulation consists of an internal energy term that penalizes the deviation of the level set function from a signed distance function, and an external energy term that drives the motion of the zero level set toward the desired image features, such as object boundaries. The resulting evolution of the level set function is the gradient flow that minimizes the overall energy functional. The intermediate results for geometric active contours are depicted in Fig. 3.

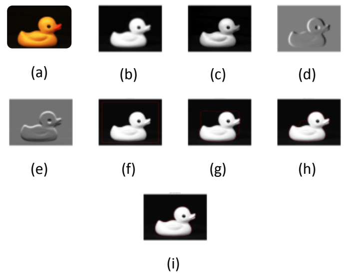

Fig. 3. Intermediate Results of Geometric (Shape) Invariant Feature Extraction (a) Input Object (b) Smoothed Object (c) Gray object (d) X gradient (e) Y gradient (f) Initial Active Contour (g) 100 Iteration (h) 200 Iteration and (i) Final Active contour (350 Iteration).

C. Texture Invariant Features:

In Texture Invariant Extraction, the Haralick Texture Features comprise of totally 14 features namely Angular Second Moment, Contrast, Correlation, Sum of Squares: Variance, Inverse Difference Moment, Sum Average, Sum Variance, Sum Entropy, Entropy, Difference Variance, Difference Entropy, Info measure of correlation1, Info measure of correlation2, and Maximum Correlation Coefficient. In our work we optimize the first 13 Haralick Texture Features and we do not compute 14th Feature number (Maximum Correlation Coefficient). Most of the features have a visual meaning, e.g. (1) the angular second moment is a measure of the smoothness and (2) a measure of contrast in the image and some of the features depend on other features as well as on intermediate results. Here we employ the 13 Haralick texture feature values for some sample objects from the ALOI database. Table 2 shows the Haralick Texture features for 2 sample images from the ALOI database.

TABLE II: THE HARALICK TEXTURE FEATURES FOR SAMPLE IMAGES FROM THE ALOI DATABASE

Images		
Angular Second Moment	1.142646	0.878005
Contrast	0.345515	0.817617
Correlation	116.5517	166.9108
Sum of Squares: Variance	7.540083	7.989534
Inverse Difference Moment	1.890158	1.716345
Sum Average	2.584632	2.711508
Sum Variance	5.510017	5.957783
Sum Entropy	0.878656	1.114017
Entropy	1.971508	2.511118
Difference Variance	0.001028	0.000991
Difference Entropy	0.439542	0.608101
Info measure of correlation1	0.049013	0.074823
Info measure of correlation2	0	0

D. Recognition Results:

The used ALOI database comprises of 1000 class image objects; for each objects 72 views are gathered, with a separation of 5°. In our work, for analysis we have chosen 50 sample class objects randomly from the ALOI database. The classification accuracy was computed using the cross-validation holdout scheme: the dataset is divided in two mutually disjoint sets, one used for the training phase and another for testing. The fact that the system is tested on a set, different from the one used for training, is a crucial factor for assessing the capability of the system in generalizing to different views. In particular, odd views are used to build the classifier and even views are used for testing. The state-of-art performance of the dataset by using the recognition techniques is analyzed with the aid of classification accuracy and error rate. The experiments results of our proposed approach for object recognition system based on the recognition technique was listed in the Table.3. Experiment results showed that significant improvement was achieved by the proposed approach.

$$ACCURACY\ (\%) = \frac{number\ \ of\ \ classified\ \ test\ \ objects}{total\ \ number\ \ of\ \ tested\ \ samples} \times 100 \tag{14}$$

$$ERROR\ \ RATE\ \ (\%) = \frac{number\ \ of\ \ misclassified\ \ test\ \ objects}{total\ \ number\ \ of\ \ tested\ \ samples} \times 100 \tag{15}$$

TABLE III: THE EXPERIMENTS RESULTS OF OUR PROPOSED APPROACH FOR OBJECT RECOGNITION SYSTEM BASED ON THE RECOGNITION TECHNIQUE

Images Class	Accuracy %	Error Rate %
Apple teabags	94.444444	5.555556
Christmas bear	88.888889	11.111111
Shell	88.888889	11.111111
Green box	94.444444	5.555556
Shoe	88.888889	11.111111
Matchbox	91.666667	8.333333
Baby cream tube	94.444444	5.555556
Yellow bear	97.222222	2.777778
Blue bear	91.666667	8.333333
Optic lamp box	94.444444	5.555556
Colorful toy square	91.666667	8.333333
Chocolate box	94.444444	5.555556
Brown bear	94.444444	5.555556
Chess horse	88.888889	11.111111
Vitamin pills	88.888889	11.111111
Alarm clock	88.888889	11.111111
Halter toy	97.222222	2.777778
Blue cell phone	97.222222	2.777778
Yellow duck	97.222222	2.777778
Red/white wire	88.888889	11.111111
Bambix	88.888889	11.111111
Emma eend book	91.666667	8.333333
Vaseline	91.666667	8.333333
Boat	94.444444	5.555556
Smiling duck	94.444444	5.555556
Wooden glue	91.666667	8.333333
Kleenex tissues	88.888889	11.111111
Picture frame	88.888889	11.111111
Droste box	97.222222	2.777778
Nail brush	91.666667	8.333333
Wooden tea box	91.666667	8.333333
Pinokio	91.666667	8.333333
Yellow can	91.666667	8.333333
Yellow/blue play mobile	94.444444	5.555556
Red car	88.888889	11.111111
Fire truck	97.222222	2.777778
Police truck	91.666667	8.333333
Alarm clock	94.444444	5.555556
Garlic 1	94.444444	5.555556
Tiger	97.222222	2.777778
White thing with handles	91.666667	8.333333
Sport life	91.666667	8.333333
Blue/white windmill	91.666667	8.333333
Pirate box	91.666667	8.333333
Sitting doll	94.444444	5.555556
Square toy	94.444444	5.555556
Cancer book	94.444444	5.555556
Yellow Kleenex tissues	94.444444	5.555556
Doll in cup	97.222222	2.777778
Yellow bag	97.222222	2.777778

VII. CONCLUSION

We have proposed an efficient approach for object recognition using invariant features and machine learning techniques. In this paper, new sets of feature descriptors have been proposed which is invariant to the viewpoint, geometry of the object and illumination conditions. In the proposed approach, we have integrated the color, texture and shape features of the objects at the feature level, so as to improve upon the recognition performance. The fused features set have been served as the pattern for the forthcoming process involved in the proposed approach. We have attained distinct results concerned with recognition rate and false positives. The experiments clearly demonstrated that our proposed approach significantly outperforms the state-of-the- art methods for combining color, shape and texture features. The proposed method is shown to be effective under a wide variety of imaging conditions.

REFERENCES

[1] A. Diplaros, T. Gevers, and I. Patras, "Combining Color and Shape Information for Illumination-Viewpoint Invariant Object Recognition," *IEEE Transactions on Image Processing*, vol. 15, no. 1, pp. 1-11, 2006.

[2] S. Agarwal and D. Roth, "Learning a sparse representation for object detection," in *Proceedings of the 7th European Conference on Computer Vision*, Copenhagen, Denmark, vol. 4, pp. 113–130, 2002.

[3] E. Borenstein and S. Ullman, "Class specific top down-segmentation," in *Proceedings of the 7th European Conference on Computer Vision*, Copenhagen, Denmark, vol. 2, pp. 109–124, 2002.

[4] R. Fergus, P. Perona, and A. Zisserman, "Object class recognition by unsupervised scale-invariant learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2003.

[5] A. Opelt, M. Fussenegger, A. Pinz, and P. Auer, "Weak hypotheses and boosting for generic object detection and recognition," in *Proceedings of the 2004 European Conference on Computer vision*, pp. 71–84, 2004.

[6] X. J. Li and I. Guskov, "3D object recognition from range images using pyramid matching," in *Proceedings of IEEE 11th International Conference on Computer Vision* (ICCV 2007), Rio de Janeiro, Brazil, pp. 1-6, October 14-20, 2007.

[7] Th. Gevers and A. W. M. Smeulders, "Color based object recognition," *Pattern Recognit*., vol. 32, pp. 453–464, 1999.

[8] J. van de Weijer and C. Schmid, "Coloring Local Feature Extraction," *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, vol. 3952, pp. 334-348, 2006.

[9] P. Perona and J. Malik, "Scale space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 12, no.7 pp. 629–639, 1990.

[10] G. D. Finlayson, B. Schiele, and J. L. Crowley, "Comprehensive colour image normalization," in *Proceedings of the 5th European Conference on Computer Vision* (In: ECCV '98), Springer-Verlag, vol. I, pp. 475-490, 1998.

[11] C. M. Li, C. Y. Xu, Chang Feng Gui, and Martin D. Fox, "Level Set Evolution Without Re-initialization: A New Variational Formulation," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (CVPR'05), vol. 1, pp. 430 - 436, 2005.

[12] F. P. Preparata and S. J. Hong, "Convex Hulls of Finite Sets of Points in Two and Three Dimensions," *Commun. ACM*, vol. 20, no. 2, pp. 87–93, 1977.

[13] E. Miyamoto and T. Merryman, *Fast Calculation of Haralick Texture Features*, 2008.

[14] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural Features for Image Classification," in *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621,1973.

[15] H. Hotelling, "Relations between two sets of variants," Biometrika, vol. 28, pp. 321–377, 1936

[16] D. Hardoon, S. Szedmak, and J. Shawe-Taylor, "Canonical correlation analysis; an overview with application to learning methods," *Neural Computation*, vol. 16, no. 12, pp. 2639–2664, 2003.

[17] Y. M. Lui, J. R. Beveridge, B. A. Draper, and M. Kirby, "Image-Set Matching using a Geodesic Distance and Cohort Normalization," in *IEEE International Conference on Automatic Face and Gesture Recognition, Amsterdam*, The Netherlands, 2008.

[18] J. M. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders, "The Amsterdam library of object images," Int. J. Comput. Vision, vol. 61, no. 1, pp. 103-112, January, 2005.
[19] Amsterdam Library of Object Images (ALOI) Datasets. [Online]. Available: http://staff.science.uva.nl/~aloi/

V. N. Pawar is a Ph. D. student at SGGS Institute of Engineering & Technology, Vishnupuri, Nanded, India doing research in the field of object recognition systems. He has received B.E. Electronics and M.E. (Control Systems) from Shivaji University, Kolhapur, India in 1990, and 1997 respectively. He is presently working at Electronics Department, A.C. Patil College of Engineering, Navi Mumbai. He is life fellow member of IETE and member of ISTE. His research interests are in Embedded Systems, Image Processing and Robotics.

Sanjay N. Talbar has received his B.E and M.E degrees from SGGS Institute of Engineering & Technology, Nanded, India in 1985 and 1990 respectively. He obtained his PhD (Highest degree) from SRTM University, India in 2000. He has received the Young Scientist Award by URSI, Italy in 2003. He has Collaborative research programme at Cardiff University Wales, UK. Currently he is working as Professor in Electronics & Telecommunication Department of SGGS Institute of Technology Nanded, India. He is a member of many prestigious committees in academic field of India. His research interests are Signal & Image processing and Embedded Systems.