Recent Advances in Color Object Recognition: A Review

Mahesh M. Solankar
Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad M.S. (India)

Pravin L. Yannawar, PhD
Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad M.S. (India)

ABSTRACT
The color object recognition is the unsolved problem in computer vision. The numbers of researchers are working to solve this problem. Various approaches to study the visual (color recognition) and geometric (shape recognition) properties of objects have been proposed. Objects are classified based on its features. In this paper various object properties are discussed. This paper reviewed the RGB, CMY and HSV color models and Texture information for visual recognition. As the surface color gets affected with the visible spectrum, solution to this illumination problem is also discussed. Objects Geometric properties of has a key role in physical representation of an object. The geometric properties like corners, edges, blobs, shapes, and region properties discussed. Finally, the four types of approaches like Appearance Base object recognition, Shape Based Object Recognition, Deformable part Based Object Recognition and Appearance plus Shape Based Object Recognition approaches are discussed.

Keywords: Object Recognition, Visual Properties, Geometric Properties, Deformable Parts.

1. INTRODUCTION
Computer's interaction with the world is limited because of their inability to see the things. Cameras and videos are able to capture and store the visual scene but they fails to understand the information they capture. Here the computer vision raises the solution to understand the image and video.

Object Recognition is a challenging task in computer vision. Objects are standalone things with well-defined boundary, shape and center. The goal of Object Recognition is to detect and categorize the object into one of the available object classes. Object classes are represented as a group of features, or parts, each part has a unique appearance and spatial position [1]. "Color Objects" are the objects having colored surface. Recognition of Color Objects is evolved as a 'Big Challenge' in computer vision.

Appearance of object color is dependent on the surface reflectance and changes according to visible spectrum [2]. Computer Vision understands videos on the basis of pattern recognition. A video is just a sequence of images. In these image sequences two types of patterns (i.e. Spatial Patterns and Temporal Patterns) are present which helps to understand the videos. Spatial Patterns: Spatial patterns are useful only for single image processing purposes. Temporal Patterns: In videos these spatial patterns are present with respect to time or temporal patterns [3].

Objects can be differentiated from each other on the basis of their visual (color and texture) and geometric (corners, edges, blobs and shapes) properties. These object properties play an important role in recognition of foreground objects which can be termed as an object of interest. In case of geometric features shape based features and their role in the classifying the object is one of the major challenge because of the wide range of invariance in illumination, scale, pose, shape, and view point in the scene [4]. In visual only features there are three issues critical for successful color classification. First is Color Constancy, Second is moving objects from video are not correctly segmented from background. Shadows are often part of objects. Lastly Complex objects are not predominantly one color [5]. Visual features are not relied on the spatial information of the objects whereas in case of geometric features spatial information of an object parts has a key role in describing the object physical structure [6]. Change in the viewpoint affects the scale and orientation of objects. The illumination properties and discriminative power can affect the visual appearance of object surface. Diagonal mapping and van Kries Model are used to model the illumination invariance [7]. Reflection can be removed by using chromatic properties of the reflection and also not relies on geometric model of the objects [8]. [9] Presented a method for modeling color adjacency relationship. Increased the intensity of preferred object boundary and weakening other edges using color adjacency model can more easily isolate the object boundary. It is observed that rather using only color features or only texture features, fusion of both of these features improves the classification accuracy to better extent [10].

Present approaches for object recognition are used visual features [2], [3], [5], [7], [8], [9], [10], [11] and geometric features [4], [6], [12], [13], [14], [15], [17], [18], [19], [22] or fusion of visual and geometric features [1], [2], [23], [24], [26]. Human lives can recognize the objects on the basis of objects properties. Every object is been represented by visual as well as geometric features.

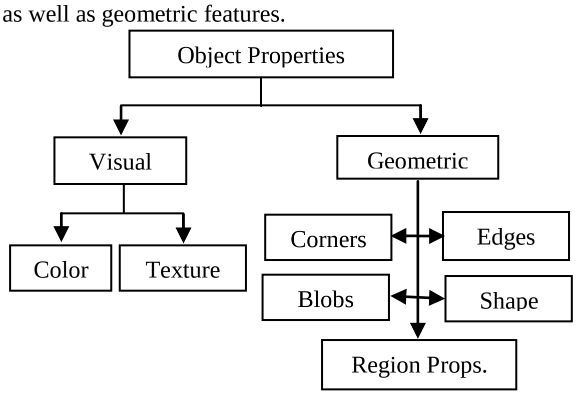

Fig. 1 Object Properties

2. OBJECT PROPERTIES

Every Object is having its unique properties, which makes it differentiate from other object classes. These properties are either Visual Properties or Geometric Properties.

2.1 Visual Properties

Visual properties are appearance based properties. These properties recognize objects on the basis of object appearance. These are not dependent on the spatial positions of Objects [6]. Visual Properties are of two types:

2.1.1 Color:

Color is an important cue for identification and extraction of object from scene. For System color is nothing more than the numerical values, but for human eyes color is key feature for object recognition.

2.1.2 Texture:

Texture is another important cue for object recognition. It represents the surface of an objects or area of interest. It provides information related to spatial arrangement of intensities or color.

2.2 Geometric Properties

Geometric properties are physical properties. These properties recognize objects on the basis of object shape, structure and geometric properties. These are dependent on the spatial positions of Objects [6]. Geometric Properties are:

2.2.1 Edges

Edges are helpful to detect and understand the object outline. Edges represent the boundary of different image regions.

2.2.2 Blobs

Blobs are connected regions in an Image. Blobs are having different properties compared to its background region.

2.2.3 Shape

Every class of an Object is having its unique shape shapes. Object shape may be Circle, Triangle, Rectangle, or Square.

2.2.4 Region-Properties

Region in an image is a connected component. Region Properties describe region to get geometric properties like Area, Convex Area, Perimeter, Bounding Box, Centroid, Extent, Orientation, Convex Hull, EquivDiamete, Extrema and Filled Area of a region.

3. LITERATURE REVIEW

Table 1 Literature review

Sr. No.	Year	Paper Name	Methods / Algorithms	Dataset	Result
1	2003	Object Class Recognition by Unsupervised Scale-Invariant Learning	SIFT, Maximum Likelihood Unsupervised learning algorithm, Gaussians Distribution, Poisson Distribution, Bayesian Decision, PCA, and Expectation-Maximization Algorithm.	Images of cars, Motorbikes, Airplanes and Faces.	Motorbikes = 92.5 Faces = 96.4 Airplanes = 90.2 Cars = 88.5 [1]
2	2005	Application and Evaluation of Color Constancy in visual surveillance	Grey-World Algorithm, Gamut Mapping Algorithm, Statistical Integration, HLS Color Space, Color Count, 3D Volume, 2D Area, Mean Saturation. Quantitative Evaluation is made by Angular errors, Euclidean errors, RMS.	Video clips of Car Parking at 1 frame per minute rate and Indoor clips.	Gamut-mapping algorithm achieves better color constancy than grey-world Algorithm. [2]
3	2005	Pattern Recognition in Video	Bayesian Decision Rule, Spatial Patterns, Temporal Patterns, Affine Appearance Model, Adaptive Velocity Model, Affine Transformations, Cylindrical Model, Markovian Model, Particle Filtering, Kalman Filter, Time-invariant first-order Morkov Gaussian Model, Dynamic Time Warping (DTW), Hidden Markov Model (HMM) .	-	Approaches based on pattern recognition to various problems like tracking, activity modeling, behavior analysis, abnormality detection are presented. [3]
4	2007	Shape based Object Classification for Automated Video Surveillance with Feature Selection	Neural Network (NN), Support Vector Machine (SVM), Support Vector Data Compression (SVDD).Shape Features : [Hu Invariant Moments, Height/Width ratio, Fill Ratio, Segment Area, perimeter, Compactness, Segment Convexity, Convex Deviation, and Projection Histogram], Online Feature Selection Method and Confusion Matrix.	Human and Vehicle Videos (1200 training samples and 800 testing samples)	Features (MBR Ratio, Fill Ratio, Hu Moment, Projection Histogram, Segment Area, Segment Convexity, and Convex Deviation) are selected. [4]
5	2008	Color Retrieval for Video Surveillance	Normalized cumulative color histogram, HSL color space, Bi-Conic Color Retrieval, Background Subtraction, Appearance based tracker, Sensitive Color Retrieval, interactive Retrieval.	1:8 hrs. video of 300Vehicles, 2: 12 Min. Video of 194 Vehicles	Highway Default Setting TP = 67% & FP = 33% Highway Tuned Setting : TP = 100% & FP = 33% Parking lot Entry Road : Default Setting : TP = 79 % FP = 21 % Tuned Setting: TP = 86 % FP = 14 % Sensitive Color Retrieval: TP = 78 % FP = 22 %. [5]

6	2008	Spatial Relationship for Object Recognition	Spatial relation, Visual words, SR-S Algorithm (Combiner Spatial Relation with statistical visual words), PCA, Maximum likelihood, Expectation Maximization Algo., Harris-affine scale and affine invariant key point's detector, k-means clustering.	Caltech 256 Dataset	Equal Error Rates Face: 3.7 Car: 6.8 Airplane: 1.2 Motorbike: 3.5 Watch: 4.9 [6]
7	2010	Evaluating Color Descriptors for Object and Scene Recognition	OpponentSIFT, Lambertian Reflectance, Histogram Based Descriptors, Color moments and color moment invariants, SIFT Descriptors, Harris-Laplace point detector, Bag-of-words model, SVM.	PASCAL VOC 2007 , Media-mill Challenge, ALOI	SIFT and Color SIFT Descriptors perform much better than Histogram Based Descriptors. [7]
8	2010	Reflection Removal in Color Videos	Chromatic Properties, Background Subtraction, Object Detection, Cut Line.	PETS 2006 Dataset	Algo. Reduces the error in estimation of actual height for objects with a reflection, while unreflected objects are left unchanged. [8]
9	2010	Color Adjacency Modeling for Improved Image and Video Segmentation [IEEE]	Graph-Cut Segmentation Algorithm, To model the Color adjacencies estimates likelihoods for data inputs using a kernel density estimation computed by Fast Gauss Transform. Method sequence is – (a) General Color Adjacency Modeling (b) Global vs. Local Information (c) Color Contamination	Pictures of Cat, Orca and Frog. Videos of Cat Football, Man and Ballet.	Color Adjacency improves the result of segmentation. [9]
10	2010	Fruit Recognition using Color and Texture Features	Fusion of Color Features and Texture Features, Minimum-Distance Classifies based on Statistical and Co-occurrences Features , Discrete Wavelet Transform, Background Subtraction.	2635 Images	Using Color: 45.49 Using Texture: 70.85 Color + Texture: 86.00. [10]
11	2010	Video Object Retrieval Based on Color Feature Modeling	Color Feature Model classifies color of an object into 13 diff. colors, Background Subtraction, Collision Avoidance, Shadow Suppression, Kalman Filter and Estimator, OpenCV, RGB to HSV.	20 videos of high scene and 20 videos of pedestrians.	The proposed Model of Color retrieval has accuracy rate more than 85 % on a average. [11]
12	2011	Automatic Object Extraction in Single Concept Videos	CRF (Conditional Random Field), Optical-Flow, HOG, GMM, Sparse Matrix Representation	Single Concept Video Frames	Better Segmentation compared to available segmentation methods. [12]
13	2011	Detection of Multiple Instances of Video Objects	Aggregation Mechanism, Greedy Algo. And Simulated Annealing Algo. , MGEP-7 DCD (Dominant Color Descriptor), Mean Shift Technique, Quadratic Form Distance Measure, Euclidian Distance.	TRECVID 2010	First Tier Score : 66 Bull Eye Score: 86. [13]
14	2011	Discriminative Models for Multi-Class Object Layout	Cutting plane algorithm, Non-Maxima Suppression, Mutual Exclusion, Greedy Forward Search,	PASCAL 2007 VOC [14]	-
15	2011	Extracting Foreground Masks towards Object Recognition	Geometric prior, Appearance prior, Graph-Cut based Energy Minimization.	VOC09 and VOC10 [15]	-
16	2012	Real Time Movement Detection for Object Recognition	Absolute Difference, Sum of Absolute Difference, Background Subtraction.	Videos containing Human and Non-Human Objects	Objects are classified as either Human or Non-Human Objects. [16]
17	2013	2D Basic Shape Detection Using Region Properties	Canny Edge Detection, Gaussian Filter, Image Region properties, Phase I: Gradient of an Image, Non Maximum Suppression, apply Hysteresis to eliminate streaking. Phase II: Centroid. Metric and circularity.	Image with various shapes of object filled by different colors.	The accuracy of this shape detection algorithm is 90.38 %. [17]
18	2013	2D Geometric Shape and Color	NTSC Standard Equation, area of an object, inclination of an object, Bounding Box,	Different shape objects of	The accuracy of this shape and color

		Recognition Using Digital Image Processing	Extent, impixel, principle of additive color mixing.	different colors.	recognition algorithm is 99.00%. [18]
19	2013	Building Part-Based Object Detectors via 3D Geometry	Deformable part based model, Strongly supervised part models, RGB Images.	NYU v2 Dataset [19]	-
20	2013	Efficient Object Detection and Segmentation for fine grained recognition	Laplacian Based Propagation, HOG Features, LLC Method, Linear SVM Classifier,	Oxford Flower Dataset, Oxford Cats and Dogs Dataset, Caltech-UCSD Birds Dataset.	Classification on Oxford Flower Dataset – 80.66, Classification on Oxford Cats and Dogs – 54.30, Classification on Caltech-UCSD Dataset – 30.17. [20]
21	2013	Existence Detection of Objects in Images for Robot Vision Using Saliency Histogram Features	Global Features, Probability Distribution Function (PDF), Saliency Map, PCA, Saliency Approaches: Frequency Tuned, Visual Attention (IT), Region Contrast (RC), Luminance Contrast (LC), Spectral Residual (SR), Precision, Recall and F-Measure.	1000 Objects and background Images captured in working robot environment [21]	-
22	2013	Learning Collections of part models for Object Recognition.	Deformable Part Model, Exemplar-based part detector, HOG Features, LDA, SVM, Sigmoid weak learner.	Pascal VOC 2010 [22]	-
23	2013	Segment, Classify and Search Objects Locally	Code maps, l2 Normalization, Non-Linear Kernel pooling, Mean Average Precision, SIFT Descriptors, Gaussian Mixture Model,	Pascal VOC 11, TRECVID 2012.	Fisher vectors require 18 seconds per image for evaluating all 20 classes, our code maps require only 6 seconds. [23]
24	2014	A Multi Scale Particle Filter Framework for Contour Detection	Contour Detection Algo. , Sequential Monte Carlo Approach, Approximated trajectory distribution, F-measure score.	Berkeley Segmentation Dataset 300, 500. [24	-
25	2014	The Role of Context for Object Detection and Semantic Segmentation in the Wild	Nearest Neighbor (NN), Markov Random Field, Support Vector Machine, SuperParsing Algo., O2P Algo. , Hierarchical Context Model.	PASCAL VOC 2010 [25]	-
26	2015	Learning Discriminative Collections of part Detectors for Object Recognition	HOG Features, Pooled Local Features, SVM, Region Based Features, SIFT Descriptor, Color and Texture, k-Means Clustering, Deformable Part Based Models.	PASCAL VOC 2010. [26]	-
27	2015	Zero-Aliasing Correlation Filters for Object Recognition.	Correlation Filters, DFT, FFT, HOG Features, Minimum Average Correlation Energy (MACE) Filter, Maximum Margin Correlation Filter (MMCF),	AT & T Database of Faces, ATR Algo. Development Image Dataset	CF Recognition Without Training for MACE (ATR Dataset): Classification : ZA : 51.9 Localization : ZA : 88.9 Recognition : ZA : 47.5 With Training : Classification : ZA : 57.7 Localization : ZA : 89.1 Recognition: ZA: 51.9. [27]
28	2015	Contextualizing Object Detection and Classification	Contextualized SVM, Linear Scaling Instantiation and Ambiguity-guided Mixture Model	PASCAL VOC 2007, 2010 and sun09	PASCAL VOC 2007 Context-SVM_LSI and Context-SVM_AMM achieves better result for classification and 18detection of 20 object classes. SUN09 Context -

					SVM_LSI–(Detection =8.39)(Classification=30. 12)And Context-SVM_ AMM (Detection=8.56) (Classification=31.43). [28]
29	2015	HFirst: A Temporal Approach to Object Recognition	Hierarchical Spiking Neural Network, Convolution Neural Network, Gabor Filter, Template Matching.	4 Class card pip, 36 class character	For 4 class card pip = 97.5%, For 36 class character = 84.9%. [29]
30	2015	Regionlets for Generic Object Detection	Deformable Part Model, Cascaded Boosting Classifier, Support Pixel Integral Image Technique, Deep CNN,	PASCAL VOC 2007, 2010, ILSVRC 2013.	VOC 2007 – 41.7 mAP VOC 2010 – 39.7 mAP ILSVRC 2013 – 16.3 mAP. [30]
31	2015	Recognising Planes in Single Image	Vector Machine Classifier, Relevance Vector Machine Classifier, Markov Random Field Segmentation Algo. , Histogram of Gradients and Color, Plane Recognition Algo. , Difference of Gaussians Detector, Support Vector Machine.	Single Images Containing Planar and Non-Planar Surfaces.	Gradient : 86.5% Color : 92.5% Gradient + Color: 93.9%. [31]
32	2015	Spatial Pyramid Pooling in Deep Convolution Networks for Visual Recognition	Convolution Neural Network, Spatial Pyramid Pooling.	ImageNet 2012 Dataset [32]	-
33	2015	Unsupervised Object Class Discovery via Saliency Guided Multiple Class Learning	Saliency Detection, Bottom-up Multiple Class Learning, K-means Clustering.	SIVAL	Proposed method can handle diverse and noisy Internet images for both clustering and detection task. [33]
34	2015	What We Can Learn From Primate's Visual System	Artificial Neural Network, Statistical Learning Technique	ImageNet Dataset	It' a study of biological visual system. [34]
35	2015	Bayesian Joint Modeling for Object Localization in weakly labelled Images	Weakly supervised learning, Joint Modelling of object classes, Bayesian Topic Model, Bayesian Priors, Probabilistic feature fusion, Non Maximum Suppression, k-means clustering, Local Binary Patterns.	Pascal VOC, ImageNet, YouTube Object Video Dataset	For Object Localization Our-Gaussian is better in simple dataset, Our-sampling is better in complicated situation. [35]

4. OBJECT RECOGNITION

The large amount of work is going on for Object Recognition. From the work reported on the same, it is been observed that the different types of approaches are tried to recognize the object classes. Fig. 2 shows the approaches used for object recognition.

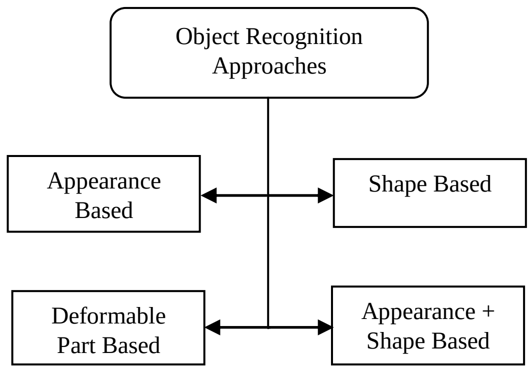

Fig. 2 Object Recognition Approaches

4.1 Appearance Based Approach

In Appearance Based approach of Object Recognition, only visual features are used for classification. These features are not dependent on the location of an objects. [2], [3], [5], [7], [8], [9], [10], [11], [21], [31], [32], and [33] used visual features. Color and Texture of an object are appearance based properties.

Color provides key information in object Recognition. The color models available in MATLAB or in OpenCV are used to recognize the object color. The color systems are RGB, CMY, and HSV.

RGB: Color system is composed of three (Red, Green, Blue) channels. Each channels values ranges from 0 to 255. These range of values indicates the intensity of particular channel. RGB color system can represents number of colors by its three basic channels at different scales. These channels can be separated to get particular color.

CMY: CMY Color system is complementary of RGB color system. It is composed of three (Cyan, Magenta and Yellow) basic channels.

HSV: HSV color system is composed of thee composed of three (Hue, Saturation, and Value) components. The luminance channel 'V' is used to extract texture related features and Chrominance channels 'H' and 'S' are used to extract color features. HSV color system is closer to the way human can recognize the colors.

The RGB color space in been converted to HSV. In HSV color space one channel contains the luminance information and remaining two channels contains chrominance information. HSV color space is mostly used because of its invariant properties under orientation and camera direction which is more suitable for object detection. [5] And [11] proposed methods to recognize color of objects. [5] Classify moving object into one of six colors: Black, White, Green, Yellow, Red, and Blue. HSL color model and Bi-Conic color retrieval algorithms are been used to classify the color objects. This method is able to predict the color on the basis of hue, saturation and luminance relation. [11] Designed a color feature model to classify the color of the video object into thirteen colors. This model converts the RGB pixels to Hue-Saturation-Value (HSV) color space, which is more close to human perception in the view of color. Background subtraction is used to segment the moving object. Collision avoidance and shadow suppression are used in order to extract the object more precisely. Kalman Filter and estimator are used to track segmented object. The result indicate that this proposed model of color retrieval is effective since the accuracy rate is more than 85 % on average. In [10] combines the color features and texture features to recognize fruits. The texture features are extracted from the luminance channel of HSV color space [10].

Appearance of object color changes according to the range of visible spectrum. Change in illumination leads to appear the object different than its original color. [2], [7], and [8] worked to solve the illumination problems. [2] Addressed Grey-World Algorithm and Gamut Mapping Algorithm to tackle the illumination problem. [7] Proposed color descriptors based on Histogram and SIFT to increase illumination invariance. Changes in the illumination can be modelled by diagonal mapping or van Kries Model. [8] Presents a novel method for reflection removal in the object detection system. The method depends on chromatic properties of the reflection and does not require a geometric model of the objects.

4.2 Shape Based Approach

Object Shape plays key role in computer vision to recognize the object class. [4], [6], [12], [14], [16], [17], and [18] have focused on the shape based features for object detection and recognition. In computer vision it is a challenge to deal with shape features of objects, because of the large amount invariance in scale, pose, rotation and camera position in the scene. These geometric invariance affects the classification result. To deal with these invariance properties of shapes, [4] discussed moment features formulated by Hu. These Hu Moment features are useful for object recognition because they are having invariance to the translation, rotation, and scale of an object. As the shape is a physical property, then the spatial position of an object also plays an important role in object detection and recognition. [6], and [14] discussed the issue of spatial positions of various objects which learns statistics that capture the spatial positions of various objects in real images.

In shape based object recognition, the first step is to locate the object to be recognized from the scene. Once the object is located, then the physically measurable properties regarding object shape are analyzed. To analyze object geometry, various methods and algorithms are used. [12] Proposed a motion-driven method to extract objects from videos. The shape model is constructed using motion cues, which helps to get foreground and background related information accordingly. Conditional Random Field (CRF) is used for video object segmentation. CRF is powerful technique to estimate structural information (i.e. class label) of set of variables. It is used to encode the known relationships between the observations and make reliable interpretations. It is often used for labeling of sequential data. In the area of computer vision, Conditional Random Field is often used for image segmentation and Object Recognition. Histogram of Oriented Gradients (HOG) features are used to describe the objects. HOG is used as a feature descriptor for object recognition. HOG computes occurrences of gradient orientation in image regions. The advantage of HOG is that objects visual appearance and shape based information within an image can be described by distribution of intensity gradients. It is also considered as invariant to the photometric and geometric transformations of objects.

There are cases, when multiple objects are presents in scene, then there may be chances of object occlusion. To deal with the problem of overlapping criteria [14] Proposed model that learns statistics both in case of which arrangements to suppress through Non-Maxima Suppression and to favor through spatial co-occurrence statistics. Non-Maxima Suppression is used to remove some detections results returned by a classifier depending on overlapping criteria or more complex heuristics. [16] Proposed an algorithm that automatically detects Human and Non-Human Object Detection. Correlation Method is used to classify detected object as a Human or Non-Human. The Foreground objects are extracted using Background. [17], and [18] worked for the 2D shapes of Objects. The 2D basic shapes like circle, triangle, rectangle or squares are detected by integrating the region properties and canny edge detection technique. The region properties represents the statistical information about an object to be recognized. The region properties gives information about area, perimeter, extent, centroid, BoundingBox of an object. On the basis of their extent value, these 2D basic shapes are classified. Extent is the ration of an object area to the area of its bounding box. The extend value is constant for each shape (I.e. extent for square and rectangle is 1, extent for circle is 0.78, extent for rectangle ranges from 0.25 to 0.50). The edges of these objects are extracted using the available edge detection techniques. The available Edge detections techniques are canny, sobel, prewitt, and LoG edge detection. Out of these techniques canny edge detection technique performs better in case of object recognition. The recognized objects are invariant to the rotation.

4.3 Deformable Part Based

Deformable part based object recognition approach classify the objects by understanding the object parts. Object parts are important regions of an objects, which are available in particular class of objects. [19], [20], [22], [26], and [30] has worked on deformable part based models to detect and recognize objects from the scene. In Deformable Part Model (DPM), Objects parts are depends upon their physical or geometric properties rather than visual properties. The important point about deformable part based object recognition is, object part is not same as the geometric element and geometric element acts as different object part depending upon its location in object. [19] Proposed three basic postulates regarding object part with respect to the object class. These postulates are: (1) object-part is an element of Object, (2) The relation of object-part with respect to the center of object, and (3) The spring model for the part, which denotes the way object part is consistent to the object.

[20] Proposed an algorithm which combines region-based detection of the object of interest and full object segmentation through propagation. The algorithm first detects low-level regions that could potentially belong to the object and then a whole object is segmented through propagation. [22] Present a framework to learn a diverse collection of discriminative parts that have high spatial consistency. To detect objects, [22] pool part detections within a small set of candidate object regions with loose spatial constraints and training a novel boosted-sigmoid classifier. HOG templates are used to model the appearance of each object part. The constellation model is one of the deformable part based object recognition algorithm. It represents an objects by a set of N parts on an object under mutual geometric conditions. This model considers the relationship between the entire object parts.

[26] defined a good part collection to have the following properties: (1) each part detector is discriminative. Relevant parts of the object should score higher than the large majority of background patches. (2) Each part detector localizes a specific part of the object or the entire object in a specific viewpoint. Detected parts should be able to predict the object pose. (3) The complete set of parts should build the object. In part based object recognition, spatial position of each part in an object has important in object class prediction.

4.4 Appearance + Shape Based Approach

According to literature survey, fusion of visual properties and shape based properties increases the object recognition rate as compared to the other approaches. [1], [13], [14], [15], and [35] used both visual as well as geometric features for recognition purpose.

[1] Proposed a technique to recognize object classes with unsupervised scale invariant learning. In this technique object regions and their scales within the scene are extracted using Entropy based feature detector. The objects are represented as a mutual position of the object parts. Every object is having appearance, relative scale and can be occluded or not. So appearance, scale, shape and occlusion are all modeled by probability density function. Features are found using the detector of Kadir and Brady. This method finds regions that are salient over both location and scale. In learning parameters of the scale-invariant object model are estimated. This is done using Expectation Maximization (EM) in a maximum-likelihood setting. The EM algorithm is to detect (locally) Maximum likelihood parameters of a statistical model when it is unable to solve equations directly. Generally these models have latent variable with unknown parameters and known data observations. In recognition, this model is used in a Bayesian Manner to classify images. Recognition result for Motorbikes 92.5 %, for Faces 96.4%, for Airplanes 90.2 and for cars (side) is 88.5%. It proves that, fusion of visual and shape based features increases the recognition result. [13] Have proposed two region-based object retrieval strategies which make it possible to detect throughout the video multiple instances of an object selected by the user. The first one is based on a greedy region construction, while the second involves a simulated annealing approach. Each key-frame is segmented using the standard algorithm: Mean Shift Technique. Promising retrieval results, with object detection rates of up to 66% (in FT(First Tier) Score) and 86% (in BE(Bull Eye) score) have been obtained on the challenging TRECVID 2010 instance search track video dataset which involves various characters in visually different scenes. [14] Learns statistics that capture the spatial arrangements of various object classes in real images. Both in terms of which arrangements to suppress through NMS (non-maxima suppression) and which arrangements to favor through spatial co-occurrence statistics. NMS is generally described in terms of intra-class inhibition, but it can be generalized to suppression of occluded detections between different classes. [15] Proposed a novel foreground / background segmentation algorithm that attempts to segment the interesting object from an image. This extracted foreground mask helps to identify the location of an object in an image. The model includes the geometric prior and appearance prior. The GIST descriptor is used to extract the geometric prior. In case appearance prior, standard Bag-Of-Words (BOW) method is used to find scenes having similar content. The graph-cut based energy minimization is used to enforce spatial coherence on the models output. [35] Address the problem of localization of objects as bounding box in images and videos with weak labels based on Bayesian joint topic modelling. Work reported on object recognition achieves increased recognition rate with the appearance plus shape based features as compared to visual only or shape only features.

5. CONCLUSION

The color object recognition is unsolved problem in computer vision. The color object are represented by visual as well as geometric properties. Various approaches are been tried to achieve accuracy in recognition.

In color recognition, HSV color space is supposed to be more reliable for accurate color recognition. Histogram of Oriented Gradients (HOG) features are widely used as region descriptors. From all the reviewed approaches, it is observed that, the fusion of appearance and shape features gives increased recognition result as compared to visual only or geometric only or deformable part based object recognition approach.

Future work will be to make the fusion of the visual and geometric object properties for automatic object identification.

6. REFERENCES

[1] R. Fergus et.al. , "Object Class Recognition by Unsupervised Scale-Invariant Learning," Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on Vol. 2. IEEE, 2003.

[2] John-Paul Renno et.al. , "Application and Evaluation of Color Constancy in Visual Surveillance," in Proceedings 2nd Joint IEEE International Workshop on VS-PETS, Beijing, pp. 301-308, October 15-16 2005.

[3] Rama Chellappa et.al. , "Pattern Recognition in Video," PReMI 2005, LNCS 3776, pp. 11-20, 2005. Springer-Verlag Berlin Heidelberg 2005.

[4] Rudra N. Hota et.al. , "Shape based Object Classification for Automated Video Surveillance with Feature Selection," 10th International Conference on Information Technology, pp. 97-99. 2007.

[5] Lisa M. Brown, "Color Retrieval for Video Surveillance." Advanced Video and Signal Based Surveillance, 2008. AVSS'08. IEEE Fifth International Conference on IEEE, pp. 283-290, 2008.

[6] Zhu, Lili and Hua Yuan, "Spatial Relationship for Object Recognition." Image and Signal Processing, 2008. CISP'08. Congress on. Vol. 2. IEEE, pp. 412-416, 2008.

[7] Van De Sande, Koen EA, Theo Gevers, and Cees GM Snoek. "Evaluating Color Descriptors for Object and Scene Recognition." Pattern analysis and Machine Intelligence, IEEE Transaction on 32.9(2010):1582-1596, 2010.

[8] D. Conte, et.al. "Reflection removal in color videos." In IEEE International Conference on Pattern Recognition, pp. 1788-1791, 2010.

[9] Price, Brian L., Bryan S. Morse, and Scott Cohen. "Color Adjacency modeling for improved Image and Video segmentation." Pattern Recognition (ICPR), 2010 20th International Conference on. IEEE, pp. 2390-2394, 2010.

[10] Arivazhagan S, et.al. "Fruit Recognition using Color and Texture Features." Journal of Emerging Trends in Computing and Information Sciences, pp. – 90-94, 2010.

[11] Ma. Xianheng, et al. "Video Object Retrieval Based on Color Feature Modeling." Machine Vision and Human-Computer Interface (MVHI), 2010 International Conference on. IEEE, pp. 101-104, 2010.

[12] Lien, Kuo-Chin, and Yu-Chiang Frank Wang. "Automatic Object Extraction in Single-Concept Videos." Multimedia and Expo (ICME), 2011 IEEE International Conference on. IEEE, pp. 1-6, 2011.

[13] Bursuc, Andrei, Titus Zaharia, and Francoise Preteux. "Detection of Multiple Instances of Video Objects." Signal-Image Technology and Internet-Based Systems (SITIS), 2011 Seventh International Conference on. IEEE. pp. 446-453, 2011.

[14] Desai Chaitanya, Deva Ramanan, and Charless C. Fowlkes. "Discriminative models for multi-class object layout." International Journal of Computer Vision 95.1, pp. 1-12, 2011.

[15] Rsenfeld, Amir, and Daphna Weinshall. "Extracting Foreground Mask towards Object Recognition." Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, pp. 1371-1378, 2011.

[16] Sreedevi M., et.al. "Real Time Movement Detection for Human Recognition." In Proceedings of the World Congress on Engineering and Computer Science 2012, Vol. I, WCECS 2012, October 24-26, San Francisco, USA, 2012.

[17] Shalinee Patel, et.al. "2D Basic Shape Detection Using Region Properties." International Journal of Engineering and Technology. Vol. 2. No. 5 (May-2013). ESRSA Publications, pp. 1147-1153, 2013.

[18] Sanket Rege, et.al. "2D geometric shape and color recognition using digital image processing." International journal of advanced research in electrical, electronics, and instrumentation engineering 2.6. (2013), pp. 2479-2487, 2013.

[19] Abhinav Shrivastava, et.al. "Building part-based object detectors via 3d geometry." Computer Vision (ICCV), 2013 IEEE International Conference on. IEEE, pp. 1745-1752, 2013.

[20] Anelia Angelova and Shenghuo Zhu. "Efficient object detection and segmentation for fine-grained recognition." Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. IEEE, pp. 811-818, 2013.

[21] Christian Scharfenberger, et.al. "Existence detection of objects in image for robot vision using saliency histogram features." Computer and Robot Vision (CRV), 2013 International Conference on. IEEE, pp.75-82, 2013.

[22] Ian Endres, et.al. "Learning collection of part models for object recognition." Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. IEEE, pp. 939-946, 2013.

[23] Zhenyang Li, et.al. "Codemaps-Segment, Classify and Search Objects Locally." Computer Vision (ICCV), 2013 IEEE International Conference on. IEEE, pp. 2136-2143, 2013.

[24] Nicolas Widynski and Max Mignotte. "A Multiscale Particle Filter Framework for Contour Detection." Pattern Analysis and Machine Intelligence, IEEE Transaction on 36.1, pp. 1922-1935, 2014.

[25] Roozbet Mottaghi, et.al. "The Role of Context for Object Detection and Semantic segmentation in the wild." Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on. IEEE, pp. 891-898, 2014.

[26] Kevin J. Shih, et.al. "Learning Discriminative Collections of Parts Detectors for Object Recognition." In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 8, pp. 1571-1584, 2015.

[27] Joseph A. Fernandez, et.al. "Zero-Aliasing Correlation Filters for Object Recognition." In IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 37. No. 8. pp. 1702-1715, 2015.

[28] Qiang Chen, et.al. "Contextualizing Object Detection and Classification." Pattern Analysis and Machine Intelligence, IEEE Transaction on 37.1, pp.13-27, 2015.

[29] Garrick Orchard, et.al. "HFirst: A Temporal Approach to Object Recognition." In IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 37. No. 10, pp. 2028-2040, 2015.

[30] Xiaoyu Wang, et.al. "Regionlets for generic Object Detection." In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 37. No. 10, pp. 17-24, 2015.

[31] Osian Haines ad Andrew Calway. "Recognising Planes in a Single Image." In IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 37. No. 9. pp. 1849-1861, 2015.

[32] Kaiming He, et.al. "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition." In IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 37. No. 9. pp. 1904-1916, 2015.

[33] Jun-Yan Zhu, et.al. "Unsupervised Object Class Discovery via Saliency-Guided Multiple Class Learning." In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 37. N. 4. pp. 862-875, 2015.

[34] Norbert Kruger, et.al. "What we can learn from the primate's visual system." KI0Kunstliche Intelligenz 29.1, pp. 9-18, 2015.

[35] Zhiyuan Shi, et.al. "Bayesian Joint Modelling for Object Localization in Weakly Labelled Images." In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 37. No. 10. pp. 1959-1972, 2015.