Object Detection In Cluttered Background using Color Clusters

Chetan S. Gode, Atish S. Khobragade

Abstract: ***Object detection in presence of complex background and illumination variation is important image analysis problem with many applications. Most of the object detection algorithms use local image descriptors which are computed from interest points based on luminance information and neglect precious color information of an object. If appearances of the object to be detected contain multiple colors in non-homogeneous distributions then it makes it difficult to detect these objects using shape features. In this context, we propose a robust algorithm designed to detect a class of objects using a descriptor which is computed from color information of an object. Clusters are formed in Hue and Saturation (HS) color space of an object using k-means clustering and cluster analysis based on number of pixels belong to each cluster, object detection is performed. Use of clustering algorithm in color space of an object to form descriptor reduces the large dimensionality of the histogram bins in the computation. The performance of the algorithm is demonstrated by experimentation carried out on standard dataset GroZi-120. Experimental results shows that the proposed algorithm is insensitive to scaling, object rotation, illumination variations and capable of handling cluttered background effectively. Finally results shows that proposed algorithm outperforms closely related algorithm by a decisive margin.***

Keywords : Object detection, k-mean clustering, target image, image descriptor, and background subtraction.

.

I. INTRODUCTION

Object detection is a one of the important steps of various computer vision applications such as video surveillance ,image retrieval, etc. which needs segmentation of regions of interest in images which are further analyzed by more computationally demanding techniques on the basis of which it can be correctly interpreted that whether object of interest is present in image or not. Computer vision systems demand precise and efficient object detection techniques in support to other high end imaging systems. Object detection algorithm should detect the object in arbitrary scenes efficiently in case of changes in illumination, object scaling, rotation and changes in viewpoint. Most of the object detection technique extract shape information from the segmented image and use it for object detection. Shape based information is provided by edges and corners of segmented object. Shape based information is used by many authors in different way such as Chang and Lee [1] used segments of circular arcs to describe the shape of object while Gonzalez-Sosa et al. [2] used shape-based feature approaches, such as shape contexts and contour coordinates to recognize person from his millimeter waves images. To extract proper shape information, it is very important that object should be segmented properly. Segmentation of object having multiple colors and having non homogeneous color distribution is difficult which makes it difficult to detect object using shape features and therefore there is need of object detection algorithm which uses color information of object to detect the object. Color information is most important feature of any object which can be used to detect the object efficiently.

II. RELATED WORK

This section provides brief overview of various related approaches in which color information is effectively used for detection object in an image.

Panda and Meher [3] proposed use of Fuzzy color difference histogram (FCDH) along with background subtraction method to detect the moving object. Color difference histogram is achieved by measuring color difference between pixel and its neighborhood pixel. Use of FCDH along with background subtraction is efficient in tracking object in case of non-stationary background and illumination variation. Color and its spatial distribution are the main source of information used to detect salient objects in an image, but use of these parameters do not always provide satisfying performance therefore Li et al. [4] proposed use of mean shift filtering along with color information to detect silent object effectively. As popular methods of color image edge detection usually neglect the use of hue, some edges which are caused by hue changes are missed. To solve this problem of edge detection, fusion of principal component analysis and hue component is proposed by Lei et al [5]. Tzanidou et al. [6] proposed use of color information to detect torso clothes of the person carrying baggage. This method takes the advantage of the fact that color of clothes and baggage is different. This approach used high contrast attributes a* and b* component of La*b* color space to segment the torso from baggage.

Detection of interest point on the basis of which local image descriptors are computed plays most important role in object categorization and image retrieval. The most of the interest point extraction algorithms are based on intensity information. In general color information of an object is ignored while extracting interest points. Stottinger et al.[7] proposed use of color based interest point for sparse image representation. Use of color interest points reduces the number of features and therefore it takes lesser time to process larger data sets. Habili et al.[8]

Revised Manuscript Received on October 15, 2019

Chetan S. Gode, Department of Electronics and Telecommunication Engineering ,Yeshwantrao Chavan College of Engineering , Nagpur, India. Email: chetan_gode@rediffmail,com

Atish S. Khobragade, Department of Electronics Engineering, Yeshwantrao Chavan College of Engineering, Nagpur, India. Email: atish_khobragade@rediffmail.com

used color information efficiently to segment the face and hands in sign language video sequences. In this approach a bivariate normal distribution of skin color in CbCr plane is obtained and image pixels are classified as skin and non skin on the basis of Mahalanobis distance between the mean vector of the skin class and their feature vectors. If Mahalanobis distance is less than predetermined segmentation threshold, pixel is classified as skin. Finally this segmentation results along with change detection results are analyzed to segment the face and hands. Spatial-color joint probability functions along with color edge co-occurrence histogram (CECH) are effectively used by Luo and Crandall [9] to detect the object. Spatial-color joint probability functions provide spatial relationships between the colors of an object. This approach is flexible than pixel-by-pixel template matching and do not require large amount of training data sets. Chen and Juang [10] proposed use of color-based entropy features and a fuzzy classifier (FC) to detect the object in complex scenes. This approach is more useful when object is difficult to detect using shape features as it is contains multiple colors in non-homogeneous distributions.

III. PROPOSED APPROACH

In order to detect object using color information we proposed use of color clusters formed in HS plane of object, we assume that algorithm is provided with an image of object being detected and input search image, which may contain one or zero target object. It is assumed that object to be detected of any reasonable size may appear anywhere in the image.
Our proposed algorithm comprises the following steps.

- Descriptor computation using query object
- Object detection in target image

Descriptor computation using query object

In this step query object descriptor is computed using color information of object. Object descriptor is further used for object detection. Following are the steps involved in computing descriptor.

A. Formation of black and white cluster

In this step query object i.e. image of object to be detected is converted from RGB to HSL color space using equations (1) to (9).
The main reason to convert RGB to HSL is that in HSL color model L (lightness) represents change in color due to change in brightness, so if we neglect the effect of brightness on color any color can be represented by only two values i.e. H(hue) and S(saturation) while in RGB color model it requires three values R(red),G(green) and B(blue) to represent any color. Representing color in HS plane is advantageous in the formation of color cluster. If color is represented by RGB color model color cluster will be three dimensional which will increase the computational complexity of clustering process. But if only HS plane of object is considered for color identification then it may lead to poor localization of an

$$R' = R / 255 \qquad (1)$$

$$G' = G / 255 \qquad (2)$$

$$B' = B / 255 \qquad (3)$$

$$Cmax = max\left(R', G', B'\right) \qquad (4)$$

$$Cmax = min\left(R', G', B'\right) \qquad (5)$$

$$\Delta = Cmax - Cmin \qquad (6)$$

$$H = \begin{cases} 0^0 & , \Delta = 0 \\ 60^0 * \left(\dfrac{G' - B'}{\Delta} mod6\right) & , Cmax - R' \\ 60^0 * \left(\dfrac{B' - R'}{\Delta} mod6\right) & , Cmax - G' \\ 60^0 * \left(\dfrac{R' - G'}{\Delta} mod6\right) & , Cmax - B' \end{cases} \qquad (7)$$

$$S = \begin{cases} 0 , \Delta = 0 \\ \dfrac{\Delta}{1 - |2L - 1|} , \Delta <> 0 \end{cases} \qquad (8)$$

$$L = (Cmax + Cmin) / 2 \qquad (9)$$

objects having black and white as a prominent colors as in HSV and HSL color space each color is represented by variation of H (hue) and S (saturation) except black and white colors. In this context we introduced the concept of black and white cluster which represents amount of black and white color present in an object image respectively. Pixels belonging to black color are assigned to black cluster while pixels belonging to white color are assigned to white cluster. To identify whether pixel belongs to white color or black color we used L(lightness) as threshold instead of V(value).HSL color model is preferred over HSV as in a HSL color model light can range from white to black while in HSV color model it can range from the desired color to black as shown in figure11 [13].

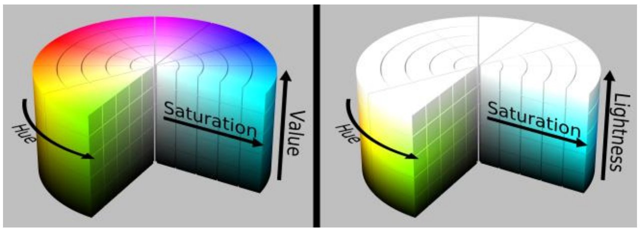

Fig. 1. HSV (left) and HSL (right) color model representations

Therefore it is obvious that HSL color model is having edge over HSV color model if one has to identify black and white colors using single threshold value. In HSL color model representation of white color is only dependent on L (lightness) and not on H(hue) and S(Saturation) while in HSV color model white color representation depends on V (value) along with H and S. therefore if HSV color model is used V(value) only cannot be considered as threshold to identify white color, along with V one has to also consider H and S as shown in figure1 Figure 2 shows that use of lightness (L) to segment white color in query object gives better result than segmenting white color using value (V) as a threshold.

In figure 2 pixels having L>.7 and H >.7 are considered as white color. But while deciding threshold for white color it should be kept as low as possible because as shown in figure 3 change in illumination can change the appearance of white color significantly which may lead to false object detection.

Fig. 2 White color segmentation of object(left) using HSV color model (middle) and HSL color model (right)

Figure 3 shows significant change in white color of the object and identified white color using (L>.7) as a threshold and (L>.6).From figure 3 it is clear that if L>.7 is considered as threshold for white color identification then it will lead to poor result while L>.6 as threshold is providing considerably good results.

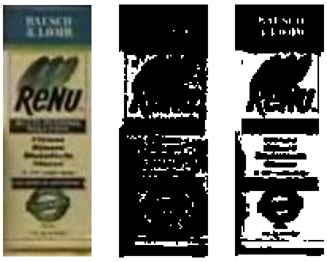

Fig. 3 White color segmentation of object(left) using L>.7 as threshold(middle) and L>.6 as threshold

Similarly black color can be identified using L as threshold. In figure 4 all the pixels having L<.8 are considered as black color.

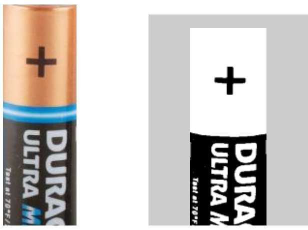

Fig. 4 Black color segmentation using L<.2 as threshold of object (right)

B. Clustering of HS plane

After formation of black and white clusters which provide us amount of black and white color present in an image, clustering of other pixels in HS plane is done. HS space which does not contain pixels representing black and white Cluster undergoes K-means clustering. K-means clustering is one of the popular unsupervised learning algorithm, which is used to solve the well known clustering problem. If d-dimensional space Rd is given, K-means clustering determines k points in Rd, called as cluster centers, so as to minimize the mean squared distance from each data point to its nearest center. If X = {x1,x2,...,xn} be the set of data points in HS plane of objects and V = {v1,...,vc} be the set centers of clusters formed in HS plane, then finally, this algorithm aims at minimizing squared error function E given by (10).

$$E = \sum_{i=1}^{c} \sum_{j=1}^{ci} (\| xi - vj \|)^2 \qquad (10)$$

||xi - vj|| is the Euclidean distance between xi and vj,‘ci’ is the number of data points in ith cluster.‘vi’ is the cluster center of ith cluster.‘c’ is the number of cluster centers.

After completion of this step HS plane of object is divided into black cluster, white cluster and K clusters as given in (11). Cluster centers of K clusters are given by (12).K means clustering of the HS plane of object to be detected along with their cluster centers is shown in figure 5. ‘HSobject’ represents percentage amount of pixels belonging to each cluster and is given by (11).

$$\text{HSobject} = \begin{pmatrix} \text{cluster1,cluster2,…,cluster K,black cluster,} \\ \text{white cluster ,nonobject cluster} \end{pmatrix} \qquad (11)$$

$$\text{Ctrs} = (\text{ctrs1}, \text{ctrs2}, \text{ctrs3}, \ldots\ldots, \text{ctrsK}) \qquad (12)$$

In equation (11) cluster 1 to cluster K is the percentage of pixels belonging to the respective cluster. Similarly black cluster and white cluster is percentage of black and white color present in object respectively. While computing descriptor percentage of pixels belonging to non object cluster will be zero as all the pixels in HS plane belongs to the object to be detected.

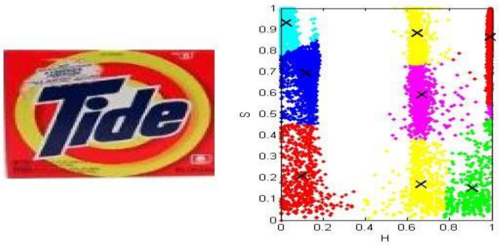

Fig. 5 Object and Clusters in its HS plane

C. Find out Dmax for each cluster

Each cluster will be having N numbers of pixels at different distances from cluster center. For all K clusters Dmax is calculated for each cluster. Dmax for a purticular cluster is equal to maximum distance between cluster center and pixel belonging to it. It is obvious that if value of K is increased, size of cluster will be smaller which will lead to smaller value of Dmax . From definition of Dmax it is clear that if d is distance between pixel in HS plane and cluster center, pixel belongs to cluster only if d is less than Dmax . This concept of Dmax is used while assigning pixels to the non object cluster in next section. If there are n pixels in cluster then vector D represents Euclidian distances of n pixels with respect to cluster center.

$$D=(d1,d2,d3,d4...d\,n) \quad (13)$$

d*n* is Euclidian distance between n^{th} pixel and cluster center of a cluster for which Dmax is to be calculated.

$$D\max = \max(D) \quad (14)$$

Detection of Object in Target Image

Process of object detection in target image is as shown in figure 6.To detect an object in target image, a rectangular search window of size w1*w2 is defined, where w1 and w2 are determined by the size of the object. Exhaustive scan of target image using rectangular window is performed to detect an object anywhere in an image. The rectangular window is shifted by z pixels at a time in one direction during scan. If size of object to be detected is unknown dimensions w1, w2 of rectangular window is changed and scan of target image is performed with resized window. This process of resizing of rectangular window continues until the window becomes greater than target image.

D. Processing of rectangular search window

If we consider a rectangular window as a image of size w1*w2, it undergoes two steps i.e. conversion of RGB image into HSV color space and formation of black and white cluster which is already explained in previous section. Having done this next step is clustering of pixels in HS plane of rectangular search window. Every pixel from HS plane of search window is assigned to one of the k clusters. To assign a cluster to any pixel its Euclidian distance with respect to the all k cluster centers given by (12) are calculated. The pixel will be assigned the particular cluster only if it is having minimum Euclidian distance with respect to cluster center of that cluster and that distance is less than Dmax for that cluster otherwise the pixel is assigned to non object cluster. After assigning cluster to each pixel of HS plane of rectangular search window, the HS plane of rectangular search window is now dividend into following cluster. Figure 7 shows cluster count for product 1.

$$\text{HSwindow}=\begin{pmatrix}\text{cluster1,cluster2,...,cluster K,black cluster,}\\ \text{white cluster ,nonobject cluster}\end{pmatrix} \quad (15)$$

Pixels belong to Non object cluster represents those colors which are present in search window but not in object to be detected.

In equation (15) cluster 1 to cluster K is the percentage of pixels belonging to the clusters respectively. Similarly black cluster and white cluster represents percentage amount of black and white color present in search window respectively

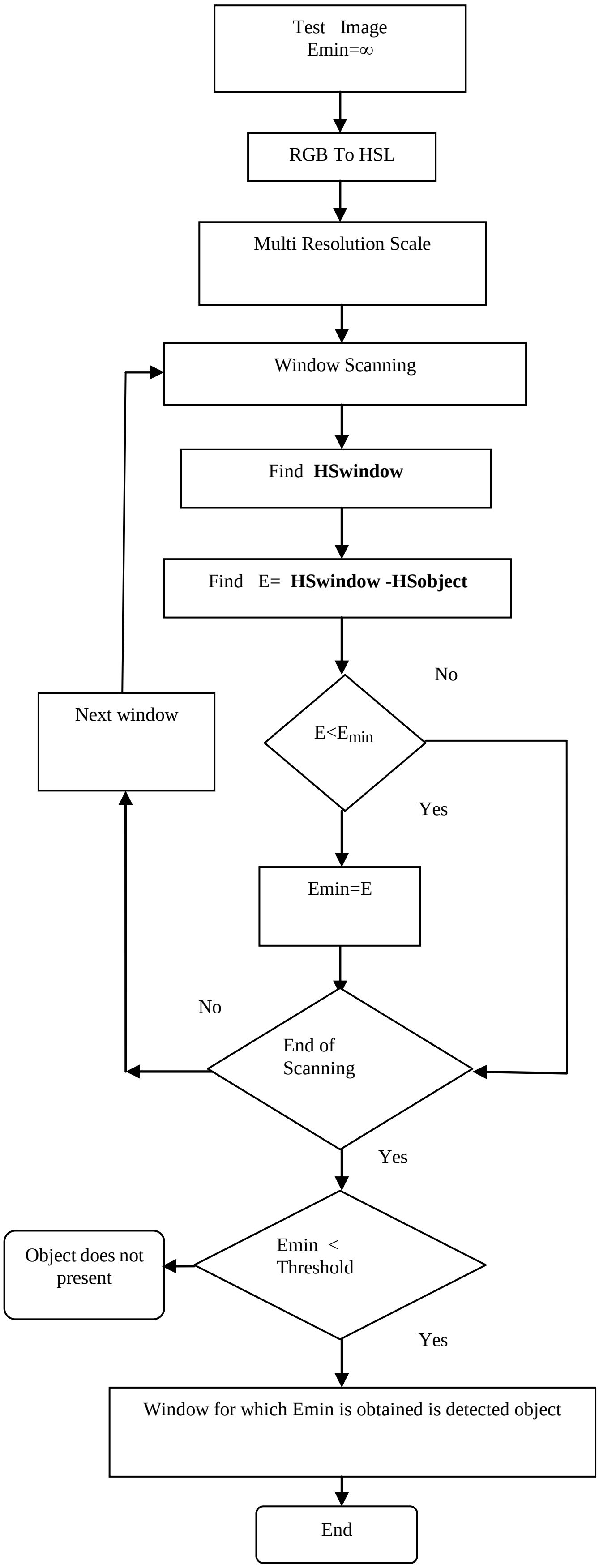

Fig. 6 Flowchart of proposed object detection process

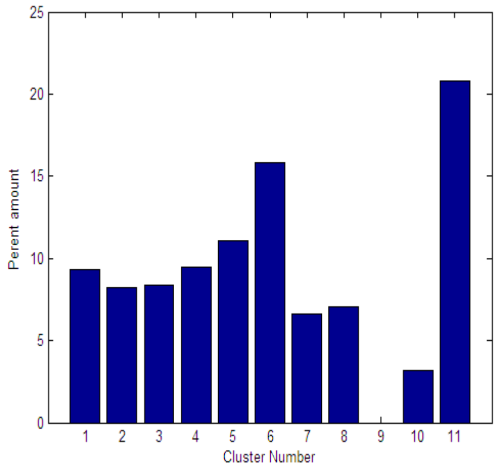

Fig. 7 Cluster count for Product 1(RENU)

To measure similarity between color distribution of object to be detected and color distribution of search window ,error 'E' is calculated.

$$E = \sum_{i=1}^{n} |\,(HS\text{object})i - (HS\text{window})i| \quad (16)$$

n is total number of clusters.n=K+3. Value of K is decided while K-mean clustering of object. While experimenting value of K is taken as 8. Smaller is the value of E, greater is the degree of similarity between search window and object to be detected.

E. Object Detection

Having calculated E for each search window it is found that for which search window from target image we are getting minimum error. If total j search windows are processed in exhaustive scan of target image we will be having vector Er representing error obtained in all search windows given by (17).

$$E_r = \left(E_1, E_2, E_3, \ldots\ldots, E_j\right) \quad (17)$$

$$Emin = \underset{1 \le i \le r}{\arg} \; minEr \quad (18)$$

If E_{min} is greater than threshold then search window for which we get minimum error E_{min} is considered as detected object else it is considered as object is not present in target image. Threshold can be decided from proper experimentation.

IV. RESULT AND DISCUSSION

In first phase of experimentation algorithm was tested on Dura images [14,15].Dura images are synthetic images and colors are distributed homogeneously with ideal illumination conditions. Figure 10, 11 shows result obtained on Dura images.

Fig. 8 Object to be detected

Fig.9 Detection results of object in Fig.8 for some of the test images

Fig.10 Object to be detected

Fig. 11.Detection results of object in Fig.10 for some of the test images.

In next phase of experimentation we compared our method, color cluster analysis (CCA), with three state-of-the-art object detection algorithms including scale-invariant feature transform (SIFT), color histogram matching (CHM) and Boosted Haar-like features(BHaar).Experimentation is performed on GroZi-120 database[11]. The GroZi-120 is a database of 120 grocery products. In order to address the problem of creating descriptor using training data that differs in quality from the testing data for object recognition and localization tasks this dataset was created. The objects belonging to dataset vary in size, color and shape. Every product in dataset is represented in two different ways, one captured in vitro and another in situ. As shown in figure *12 in vitro* images come from ideal imaging conditions while in situ images which are shown in figure 13 are captured under natural environments (real world). These images vary in color, illumination and rotation and provide cluttered background for object to be detected. Hence, performance of different algorithms for object detection and localization can be tested using this dataset.

For descriptor formation images from in vitro database is used while images from in situ are used as test images. Cluster counts obtained from images belonging to the same product were averaged, cluster per cluster, in order to obtain a final template (HSobject) representative of the object.

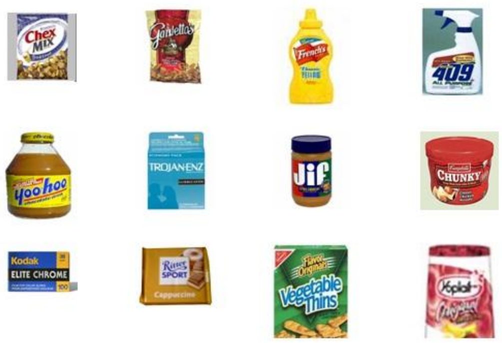

Fig. 12 Sample images of in vitro database for different products.

Fig. 13 Sample images of in situ database for different products.

Figure 14 shows ROC curves for product 1. In case of product 1, CCA outperforms CHM and Bhaar and performance of CCA is closer to SIFT. Because of distinctive text and symbols on product box SIFT has got more interest points for detection and it is performing well. Performance of CHM is poor as product 1 is having white color as prominent color and white color cannot be distinguished properly in chrominance plane.CCA is having edge over CHM as it is using concept of white and black cluster to represent amount of white and black color present in training image and shows better localization of object than CHM.

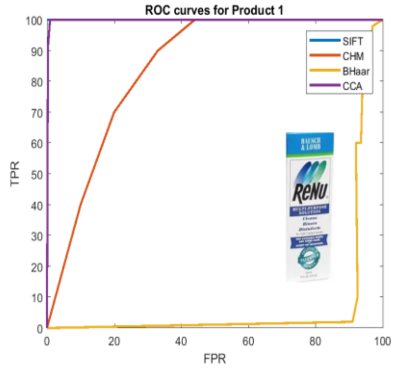

Fig.14 ROC curves for pruduct 1

Figure 15 shows ROC curves for product 15.In case of product 15 CCA and CHM are performing well as compared with SIFT and BHaar as object is having orange, green and yellow as prominent color . Figure 15 shows that performance of CCA is better than CHM ,SIFT and BHaar. Figure 16,17 considers different cases for performance analysis of four algorithms. Figure 16,17 shows that CCA provides better localization of product 52 and 34 than remaining three methods. In figure 16 BHarr and SIFT gets misled by text written on neighbourhood product while CCA is having advantage that product colors are distinguishable from background color and it shows better localization than other methods. Figure 17 shows good performance by all the methods while CCA outperforms all the remaining three methods.

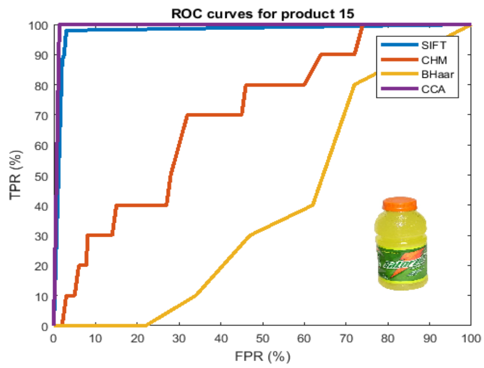

Fig. 15 ROC curves for pruduct 15

Fig. 16 Product 52 and its localization

Fig.17 Localization example of product 34 shown in Fig.5

V. CONCLUSION

In this paper, we have proposed simple but effective approach for object detection based on color distribution of object. We have shown how the descriptor derived from clustering of HS color space of the object can be used for object detection. The results have indicated that CCA shows significant improvement over state of art algorithms like CHM, BHarr and SIFT for object detection. CHM computed histogram of total 32 bins, 16 bins per channel, a and b from Lab color space while CCA shows improved results by using only 11 bins. This indicates that CCA also shows improvement in terms of computational complexcity.CHM shows very poor localization in case of object having white and black color as prominent color while CCA has successfully localized these objects using concept of black and white clusters. Background clutter because of natural conditions in situ images is handled successfully by CCA and shows better localization as compared with other methods. This approach is recommended for detection of object having non-homogeneous color distribution and difficult to segment. Advantage of this approach is detection is scale invariant and insensitive to object rotation.

REFERENCES

1. W. Chang, S. Lee, "Description of shape patterns using circular arcs for object detection", IET Computer Vision, Vol. 7, No. 2, 2013, pp. 90–104.
2. E.Gonzalez-Sosa, R.Vera-Rodriguez , J.Fierrez and V.M.Patel, "Exploring Body Shape From mmW Images for Person Recognition", IEEE Transactions on Information Forensics and Security, Vol. 12, No. 9, 2017, pp.2078-2089.
3. D.K.Panda and Meher S," Detection of Moving Objects Using Fuzzy Color Difference Histogram Based Background Subtraction", IEEE Signal Processing Letters, Vol. 23, No. 1, 2016,pp. 45-49
4. J.Li ,H. Chen , G.Li , B. He, Y. Zhang and X. Tao "Salient object detection based on mean shift filtering and fusion of color information", IET Image Processing, Vol. 9, No. 11, 2015,pp. 977-985
5. T.Lei,Y. Fan,Y. Wang ," Colour edge detection based on the fusion of hue component and principal component analysis ", IET Image Processing, Vol. 8, No. 1, 2014 , pp. 44–55
6. G.Tzanidou ,T. Zafar ,E.A. Edirisinghe ."Carried Object Detection in Videos in Videos Using Color Information",IEEE Transactions on Information Forensics and Security. Vol. 8, No. 10, 2013, pp. 1620-1631.
7. J.Stöttinger,A. Hanbury,N. Sebe and Gevers,T ," Sparse Color Interest Points for Image Retrieval and Object Categorization", IEEE Transactions on Image Processing, Vol. 21, No. 5,2012,pp.2681-2692
8. N.Habili, C.Chew and L.A.Moini ," Segmentation of the Face and Hands in Sign Language Video Sequences Using Color and Motion Cues" IEEE Transactions on Circuits and Systems for Video Technology , Vol. 14, No. 8 , 2004 , pp.1086-1097
9. J.Luo ,D. Crandall ." Color Object Detection Using Spatial-Color Joint Probability Functions",IEEE Transactions on Image Processing, Vol. 15, No. 6 , 2006 , pp.1443-1453
10. G.Chen ,C. Juang ,"Object Detection Using Color Entropies and a Fuzzy Classifier", IEEE Computational intelligence Magazine,2013,pp.33-45
11. M.Michele, G.Carolina, and B.Serge ," Recognizing Groceries in situ Using in vitro Training Data" ,in proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition,2007 ,pp.1-8
12. < Source: http://en.wikipedia.org/wiki/HSL_and_HSV >
13. <Source: https://www.instazu.com>
14. < Source : http://www.statix-electrical.co.uk>

AUTHORS PROFILE

Chetan S.Gode received the B.E. degree in Electronics from the Shri Guru Gobind Singhji Institute of Engineering and Technology,Nanded,in 1999.He received his Mastersof technology in Electronics Engineering from Rashtrasant Tukadoji Maharaj Nagpur University of India in 2007.He is currently doctoral research fellow in Rashtrasant Tukadoji Maharaj Nagpur University of India.He is currently working as an Assistant Professor in Rashtrasant Tukadoji Maharaj Nagpur University of India. He has over 14 years of teaching experience.His research interests include image processing and embedded systems. He is member of Indian society for technical education (ISTE). He has published his research work in many internationals conferences and journals

Atish S.Khobragade received the B.E. degree in Electronics from the Walchand College of Engineering, Sangli, India in 1992 and M. Tech degree in Electronics Engineering discipline from Electronics Engineering Department,Visvesvaraya Regional College of Engineering, Nagpur, India, in 2001 and the Ph.D. degree in Electronics Engineering at University of Nagpur in 2013. He is currently working as Professor in Rajiv Gandhi College of Engineering, Nagpur. He has over twenty three years of teaching experience. He has published more than twenty papers in International Conferences and ten papers in International Journals. He is a Life Member of (ISTE) India, Society of Technical Education, Member of IEEE and also a Member of Institute of Engineers Association. His research interest includes image processing and pattern recognition