Journal of Engineering Sciences
Vol 11, Issue 3, March/2020
ISSN NO:0377-9254

# A FRAMEWORK FOR OBJECT SEGMENTATION BASED ON ACTIVE CONTOUR AND COLOR RECOGNIZING MODEL

Dr. R. Matheswari
*M.E.S College of Arts, Commerce and Science, Bangalore*
*matheswarir@gmail.com*

***Abstract* - Object detection and tracking of moving objects play a vital role in digital image processing. Segmentation of moving objects from dynamic background conditions play an important role in visual surveillance system. An object moving in a particular boundary needs to be detected and tracked for several reasons. Though many existing works addresses these issues effectively, there is still a need to develop more optimized algorithm for segmenting objects in dynamic background conditions. As a way of addressing these issues, a new algorithm for object segmentation in the presence of challenging dynamic background conditions is presented in this paper. This work presents a combined framework of segmenting objects based on active contour and color recognizing model with a set of fuzzy aggregated multifeature similarity measures model.**

***Keywords* - Color detection, Image tracking, Active Contour Model, ACM, statistical local texture features, model feature vectors, model level fuzzy similarity, neighborhood supported model initialization.**

## 1. INTRODUCTION

Image processing is a rapid growing research area. The applications of image processing methods requires quantitative and qualitative analysis of the effective methods suitable for the various circumstances. Segmentation of objects is an essential step in image processing. Different methods are available to segment an object from the reference background. An object recognition and tracking system, here known as SRCCA (Software de Reconhecimento de Cor e Contorno Ativo or Active Contour and Color Recognizing Software) by using color recognition in the HSV color space and active contour models (ACM) developed in [1] has a greater potential for applications involving the tracking of moving targets, with processing time and convergence-suited to real-time operation and robustness to interference intrinsic to the scenarios and environments.. However, changes in the luminance or in the saturation of the colors also not committed to the system due to the intrinsic characteristics of the color space HSV.

A set of fuzzy aggregated multifeature similarity measures applied on multiple models corresponding to multimodal backgrounds. Advanced Fuzzy Aggregation based Background Subtraction (AFABS) developed in [2], integrates multi-feature multi-model with neighborhood assisted initialization into a model level fuzzy aggregation framework, where each pixel is associated with weighted feature vectors, composed of intensity (I) and statistical texture (ST). The algorithm is enriched with a neighborhood-supported model initialization strategy for faster convergence. A model level fuzzy aggregation measure driven background model maintenance ensures more robustness. Similarity functions are evaluated between the corresponding elements of the current feature vector and the model feature vectors. Concepts from Sugeno and Choquet integrals are incorporated in our algorithm to compute fuzzy similarities from the ordered similarity function values for each model. Model updating and the foreground/background classification decision is based on the set of fuzzy integrals.

## 2. RELATED WORKS

The history of video-based object detection starts from detection of moving objects in videos captured by a stationary camera. Jain and Nagel [2] proposed the frame difference scheme to detect a foreground object. Wren et al. [3] proposed the use of a Gaussian model, Stauffer and Grimson [4] proposed a GMM-based approach, and Elgammal et al. [5] applied kernel density estimation for background modeling. Unfortunately, the above methods cannot

serve well for scenarios in which the camera is moving (even with nominal motion). Recent researchers focus more on foreground object detection in videos captured by freely moving cameras.

In [6], Sheikh and Shah proposed to build foreground and background models using a joint representation of pixel color and spatial structures between them. In [1], Patwardhan et al. decomposed a scene and used maximum-likelihood estimation to assign pixels into layers. From their experimental results, only moving foreground objects with the average velocity up to 12–15 pixels per frame can be detected. As a result, their approach is only capable of handling videos captured by a camera with mild camera motions. In this paper, we address automatic video foreground object detection problems under arbitrary camera motion (e.g., panning, tilting, zooming, and translation). Prior methods focusing on this type of problem can be classified into two categories. The first category (e.g., Meng and Chang's method [7]) is to detect moving foreground object as the outliers, and thus to estimate the global motion of the camera [8]. Irani and Anandan [9] proposed a parametric estimation method for detecting the moving objects, and Wang et al. [10] also approached this problem in a similar setting.

## 3. RESEARCH METHODOLOGY

The research work proposes a combined framework of segmenting objects based on Active Contour and Color Recognizing model (ACCR) for effective segmentation of objects.

The segmentation method is based on an online multilayer background modeling technique known as Multi-background registration (MBReg). The key concept in this algorithm is the fact that it models the background with N layers of background images instead of a single background layer. For each pixel position, the corresponding pixel in each layer of the background image represents one possible background pixel value. As shown in Fig.1. , the background model is established and maintained in the MBReg and background update and release blocks.

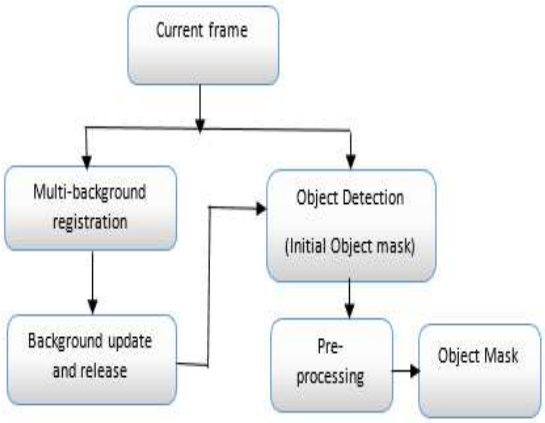


Fig 1. Multi background registration

### 3.1 THE COLOR DETECTION IN HSV

Digital video cameras often have the capture data in the format of YCbCr, a pattern that is widely-adopted for to be able to ensure the highest quality for the storage of data of luminance, in comparison with the data of chrominance. However, despite the YCbCr to be a color space appropriate to the capture and storage of images, if compared to other formats available, it is shown to be less suitable for color processing [13].

### 3.2 FILTERING AND SELECTION OF OBJECTS

The bit array generated by the converter color displays noises from objects with similar colors to the colors of the target of the trace. Become necessary, therefore, screens that may provide the tracked object to be the target desired. The first filter applied on the bit-map is a filter morphological, which performs the operations of dilation and erosion. The operation of erosion eliminates small white regions in the map, considered a noise. The expansion eliminates the small black regions within the white regions, reducing the noise the internal of the objects identified.

The Fig. 2, illustrates the operation of erosion.

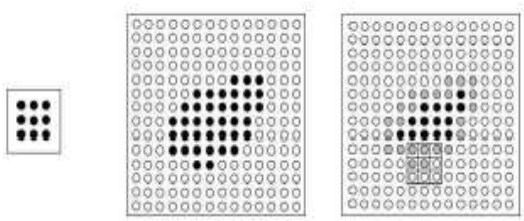

Fig 2. The operation of erosion on an image

The coefficients that establish the weight of that average are defined by the relation of the gaussian-established as equation (1) mentioned below.

$$g(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{\frac{x^2+y^2}{2\sigma^2}} \quad \textbf{(1)}$$

where x and y are the coordinates within the window of the filter, and$\sigma$ is a value of variance. The larger the window of the filter, the greater is its ability to reduce the noise present in the image. By applying the filter to a window of 3x3 can eliminate more noise than the filter of the simple average using a window of the same dimension, while maintaining a greater integrity of the details.

### 3.3 SEGMENTATION OF OBJECTS

Among the several available techniques, the targeting of objects in an image can be performed by means of the models contour active (ACM). In this method, the object is tracked from lines to splines that seek, by means of iterations, the regions of lowest energy along the contour of the object. These splines are also called snakes (snakes) because of your behavior "trembling" while the suit the regions of low energy [1].

$$E_{snake} = \int_0^1 (E_{int}(V(S)) + E_{image}(V(S)) + E_{field}(V(S)))ds \quad (2)$$

the energy of the image is calculated as the weighted sum of the three tranches denominated energy of line, edge and termination

$$E_{image} = W_{Line}E_{Line} + W_{Edge}E_{Edge} + W_{term}E_{term} \quad (3)$$

### 3.4 ADVANCED FUZZY AGGREGATION-BASED METHOD OF SEGMENTATION OF OBJECTS

Multi-feature based background subtraction techniques require efficient fusion of the information contributed by the individual features. Information fusion is the process of combining these features into a single datum and the fusion is achieved through aggregation operators, which are mathematical functions.

### 3.5 COMPUTATION OF FEATURE VECTOR

For every frame of the video sequence and for every pixel, we compute a feature vector, which comprises of four elements – the intensity (*I* ) and three local ST features, namely, mean, local homogeneity, and energy, derived from the co-occurrence matrix, computed over a neighborhood region *R* of size *M* × *N*, centered at that pixel. The ST features — local homogeneity (*L*), energy (*E*), and the *x*-directional mean (*Mx* ) — are calculated using normalized symmetric co-occurrence matrix as follows:

$$L = \sum_{i=0}^{l-1}\sum_{j=0}^{l-1} \frac{1}{1+(i-j)^2} c(i,j) \quad (4)$$

$$M_x = \sum_{i=0}^{l-1}\sum_{j=0}^{l-1} i\,c(i,j) \quad (5)$$

$$E = \sum_{i=0}^{l-1}\sum_{j=0}^{l-1} c(i,j)^2 \quad (6)$$

where, *c(i, j )*'s are the elements of the normalized symmetric co-occurrence matrix. A feature vector **X** with a set of four features (*I* , *L*, *E*, *Mx )* is formed for each pixel.

$$X = [I\ L\ E\ M_x] \quad (7)$$

### 3.6 MODEL LEVEL FUZZY AGGREGATION OF FEATURES

For every pixel, a set of feature vectors is maintained as models and each is compared with the

feature vector computed at that pixel in the current frame. The similarity values obtained for such element by element comparisons for each model are ordered and then the fuzzy integrals are computed on these ordered similarities using the membership values of the features. In this subsection, we present the definitions of the fuzzy measures and the fuzzy integrals, which would be necessary for a proper understanding of our algorithm.

Let X= _$x1$, $x2$, . . . , $xn$_ be the finite feature set and **B** be the Borel field of **X.** $\mu(xi) \in (0, 1)$ be the importance given to the feature $x_i$ .

### 3.7 ADVANCED FUZZY AGGREGATION BASED BACKGROUND SUBTRACTION (AFABS)

In AFABS, each pixel is modeled with a feature vector, composed of intensity and ST features, a combination of pixel and region-based features, to inherit the advantages of both types of features. By giving an importance value to each feature and fusing those by a fuzzy integral, correlations or interactions between the features can be considered. In this approach, multiple models are constructed for each pixel, where the models are initialized with neighborhood support, thereby achieving faster convergence of the background model to the background variations. Model level fuzzy similarity, calculated between each model and the current by the fuzzy integral, represents the amount of matching between those, and the model is updated accordingly. The schematic diagram of the proposed approach is shown in Fig. 1. The algorithm consists of five steps — model initialization, background models selection, fuzzy integral calculation for all the models, background model updating, and the foreground detection. The last subsection deals with the optimization of the parameter values used in the proposed algorithm.

```
Step: 1 for each pixel at t=0, initialize r models and
initialize model weights to the same values
for t>0 to the end of sequence do
For each pixel do
Step -2 form a feature vector X
Step -3 normalize the models weights
Step -4 select top b high weighted models
Step -5 evaluate the similarity function for each
model with X.
Step -6 evaluate the fuzzy integral set (F1,....Fi)
Step -7 [val,s]=max(F1,....Fi)
Step -8 if(val<Ti) do
a. Pixel label=foreground
b. Update the model using case-1
        Else do case-2
c. If (s>b) pixel label=foreground
d. Else pixel label=background
e. Update the matched model
f. Update models weights
end
end
```

Algorithm 1. Segmenting objects based on Active Contour and Color Recognizing model (ACCR)

### 3.8 MODEL UPDATING

The model of the pixels should be updated in order to cope up with the changes that have taken place in the background, as follows:

**Case-1:** If the $max(F1, . . . , Fr) < TP$, the current feature vector replaces the model feature vector having the lowest weight.

**Case-2:** If the $max(F1, \cdot \cdot \cdot , Fr) \geq TP$, the best matching model that is the model feature vector having the maximum integral value is updated with the current feature vector.

### 4. EXPERIMENTAL EVALUATION

The research work is implemented in MATLAB using 168VJ Clip dataset. Video frames ranging from 10 to 100 frames were taken as input to the proposed system. The evaluation of the proposed method is performed using the following factors viz, segmentation time(ST), Peak signal-to noise ratio(PSNR), segmentation accuracy(SA) and object detection rate(ODR).

$$ST = Number\ of\ frames\ taken * Time\ taken\ in\ milliseconds \quad (8)$$

$$SA = \frac{Number\ of\ objects\ segmented}{Number\ of\ frames\ considered} * 100 \quad (9)$$

$$PSNR = 20\log 10(MAX/(MSE)\hat{}(\tfrac{1}{2})) * 100 \quad (10)$$

$$ODR = \frac{Number\ of\ objects\ detected}{Number\ of\ frames\ considered} * 100 \quad (11)$$

From the above expressions segmentation time, segmentation accuracy, PSNR ration and object detection rate are calculated and the results obtained from proposed and existing methods in [1] and [2] are tabulated as given below.

Table I clearly shows that the proposed method ACCR has ST of 19 % and 34% lesser than ACM in [1] and AFABS in [2] and also improved SA by 10% and 22% when compared to existing ACMs and AFABS method developed in [1] and [2] respectively. Likely the proposed method ACCR has 18% and 44% higher PSNR than ACM in [1] and AFABS in [2] and also improved ODR rate by 13% and 29% than ACM in [1] and AFABS in [2] respectively.

## 5. CONCLUSION AND FUTURE WORK

The research presented a new way of integration techniques of recognition of colors and active contour models ACMs with Advanced Fuzzy Aggregation based Background Subtraction (AFABS) for the tracking of objects, usually used separately and in different applications. Their versatility in the recognition of colors, by means of the adjustment of parameters by the user, causes the system to be suitable for the tracking of moving targets of any colors, in varied scenarios. Qualitative and quantitative experiments are carried out to show the effectiveness of this proposed method of segmenting objects based on Active Contour and Color Recognizing model (ACCR) in handling various challenging situations. Future work should address improved performance on heavily dynamic background situations.

Table I. Performance evaluation with respect to ST, PSNR, SA and ODR

| No. of Video Frames /Sec | ST (ms) | | | PSNR (db) | | | SA(%) | | | ODR(%) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ACCR | ACM | AFABS | ACCR | ACM | AFABS | ACCR | ACM | AFABS | ACCR | ACM | AFABS |
| 10 | 10.9 | 13.8 | 18.4 | 48.12 | 40.85 | 33.54 | 84.23 | 74.23 | 66.26 | 71.34 | 62.44 | 53.51 |
| 20 | 12.5 | 16.1 | 21.6 | 50.85 | 43.2 | 35.61 | 87.34 | 78.52 | 70.33 | 74.48 | 64.35 | 56.34 |
| 30 | 15.2 | 18.7 | 23.3 | 54.35 | 45.86 | 37.42 | 88.21 | 80.31 | 72.34 | 76.47 | 67.15 | 58.63 |
| 40 | 16.7 | 20.9 | 25.6 | 58.76 | 49.23 | 40.35 | 89.15 | 81.35 | 73.34 | 78.33 | 68.31 | 60.34 |
| 50 | 18.8 | 22.6 | 28.4 | 63.28 | 53.61 | 43.03 | 90.61 | 82.41 | 74.33 | 80.64 | 70.31 | 62.25 |
| 60 | 20.1 | 24.8 | 29.1 | 66.97 | 56.44 | 46.09 | 91.12 | 83.27 | 75.08 | 82.41 | 72.22 | 64.09 |
| 70 | 21.3 | 26.3 | 31.5 | 72.43 | 60.28 | 49.88 | 92.44 | 84.18 | 76.28 | 84.11 | 74.81 | 66.01 |
| 80 | 22.6 | 27.2 | 32.5 | 75.44 | 63.81 | 52.46 | 93.34 | 85.64 | 77.46 | 85.34 | 76.05 | 67.31 |
| 90 | 23.4 | 28.1 | 34.6 | 77.37 | 66.12 | 54.31 | 94.25 | 86.34 | 78.27 | 86.64 | 78.61 | 69.15 |
| 100 | 24.8 | 29.3 | 35.1 | 78.49 | 68.33 | 56.43 | 95.34 | 87.46 | 79.22 | 87.11 | 80.22 | 70.34 |

**REFERENCES**

[1] A.S. Silva, F.M.Q. Severgnini, M.L. Oliveira, V. M. S. Mendes and Z. M. A. Peixoto and, "Object Tracking by Color and Active Contour Models Segmentation", IEEE Latin America Transactions, Volume 14, Issue No.3, March 2016, Pages 1488 - 1493.

[2] Pojala Chiranjeevi and Somnath Sengupta, "Neighborhood Supported Model Level Fuzzy Aggregation for Moving Object Segmentation", IEEE Transactions on Image Processing, Volume 23, Issue No.2, February 2014, Pages 645 - 657.

[3] K. Patwardhan, G. Sapiro, V. Morellas, Robust foreground detection in video using pixel layers, IEEE Trans. Pattern Anal. Mach. Intell. 30 (4) (2008) 746–751.

[4] R.C. Jain, H.H. Nagel, On the analysis of accumulative difference pictures from image sequences of real world scenes, IEEE Trans. Pattern Anal. Mach. Intell. 1(2) (1979) 206–213.

[5] C. Wren, A. Azarbayejani, T. Darell, A. Pentland, Pfinder: real-time tracking of human body, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 780–785.

[6] C. Stauffer, W. Grimson, Learning patterns of activity using real time tracking, IEEE Trans. Pattern Anal. Mach. Intell. 22 (8) (2000) 747–767.

[7] A. Elgammal, R. Duraiswami, L.S. Davis, Efficient kernel density estimation using the fast gauss transform with applications to color modeling and tracking, IEEE Trans. Pattern Anal. Mach. Intell. 25 (11) (2003) 1499–1504.

[8] Y. Sheikh, M. Shah, Bayesian modeling of dynamic scenes for object detection, IEEE Trans. Pattern Anal. Mach. Intell. 27 (11) (2005) 1778–1792.

[9] J. Meng, S.F. Chang, CVEPS – a compressed video editing and parsing system, in: Proc. ACM Intl. Conf. Multimedia, 1997, pp. 43–53.

[10] Y. Tan, S.R. Kulkarni, P.J. Ramadge, A new method for camera motion parameter estimation, Proc. IEEE Intl. Conf. Image Process. 1 (1995) 23–26.

[11] M. Irani, P. Anandan, A unified approach to moving object detection in 2D and 3D scenes, IEEE Trans. Pattern Anal. Mach. Intell. 20 (6) (1998) 577–589.

[12] R. Wang, H.J. Zhang, Y.Q. Zhang, A confidence measure based moving object extraction system built for compressed domain, Proc. IEEE Intl. Sympos. Circ. Syst. (5) (2000) 21–24.

[13]. P. Sebastian, Y. V. Voon, R. Comley, "The effect of colour space on tracking robustness", Industrial Electronics and Applications, v.1, 2008.

[14] A.C. Bovik. *Handbook of Image and Video Processing*. Academic Press, Inc. Orlando, FL, USA, 2005.

[15]. Mignotte, M., Konrad. I.: Statistical background subtraction using spatial cues. Circuits Syst.

Video Technol. IEEE Trans. 17(12), 1758–1763 (2007).

[16]. Neji, M., Wali, A., Alimi, A.M: Towards an intelligent information research system based on the human behavior: recognition of user emotional state. In: 12th IEEE/ACIS International Conference on Computer and Information Science, Japan (2013).

[17]. Ralph. O.M, Kong, S.G.: Senior member. In: Visual Analysis of Eye State and Head Pose for Driver Alertnes Monitoring, vol. 14, pp. 1462–1469. IEEE, USA (2013).

[18]. Wierwille, W.W, Ellsworth, L.A., Wreggit. S.S, Fairbanks. R.J and Kirn. C.L: Research on vehicle based driver status/performance monitoring: development, validation and refinement of algorithms for detection of driver drowsiness. National Highway Traffic Safety Administration, Technical report. DOT HS 808 247 (1994).

[19]. Marques, J.S., Jorge, P.M., Abrantes, A.J., Lemos, J.M.: Tracking groups of pedestrians in video sequences. In: IEEE International Conference on Pattern Recognition, vol. 9 (2003).

[20]. Chang, T.-H., Gong, S., Ong, E.-J.: Tracking multiple people under occlusion using multiple cameras. In: Proc. 11th British Machine Vision Conference (2000).