# Automatic Color Object Detection and Learning using Continuously Adaptive Mean Shift with Color, Scale and Direction

Kishor S. Jeve
Department of Computer Science,
College of Computer Science and
Information Technology, Latur.

Ashok T. Gaikwad
Institute of Management Studies
and Information Technology,
Vivekanand College,
Aurangabad.

Pravin L. Yannawar
Department of C.S. & IT,
Dr. Babasaheb Ambedkar
Marathwada University,
Aurangabad.

## ABSTRACT
The tracking of motion of color object and detection of objects in real time video sequences is the important task in computer vision, video processing, image processing, intelligent system, etc. In this paper, we proposed Continuously Adaptive Mean Shift (CAMSHIFT) tracking with Color, Scale and Direction. In the CAMSHIFT the scale of the kernel is automatically altered in accordance with the object scale, and also follows the object direction accurately, when an initial adoption of the object shape is provided. The proposed method tracks the object accurately and efficiently.

## Keywords
Continuously Adaptive mean shift, object tracking, color, scale, direction kernel.

## 1. INTRODUCTION
Nowadays, Automatic Color Object Detection and Learning is used in a wide range of fields, such as intelligent machines, multimedia, video-data compression, industry production, military affairs and so on. Accordingly, it is of great real significance and application value to investigate in object tracking. Many Methods are used to track the object in videos. Mean shift tracking [1][2][3] is one of the robust methods for object tracking and learning. Its flexibly combines the technologies of image processing, auto control and information science, forms a new technology of real time detection of moving object, extraction of location information of the object and tracking of it.

Moreover, rapid progress in technologies of signal processing, image processing, sensor development, robotics, and computer vision provides reliable software and hardware for the capturing and processing of images or videos in real time.

Detecting moving object is difficult due to its Complex object shapes or change in motion, moving direction, non-rigid or articulated nature of objects, partial or full object occlusions, illumination changes in image, Real-time processing requirements, Noise in video or Audio etc. Continuously Adaptive mean shift [4][6] tracking uses mean shift tracking with adaptive filtering to detect moving objects shape and to track object throughout video scene.

## 2. SYSTEM OVERVIEW
The proposed system takes video as input from video database. Selected video is divided into consecutive image frames [9][10][12]. Select first frame as a target model and mark the object to track. Apply the proposed algorithm for detection and Learning of the object thought the video sequence. The Architecture of the proposed system is shown in fig 1.1
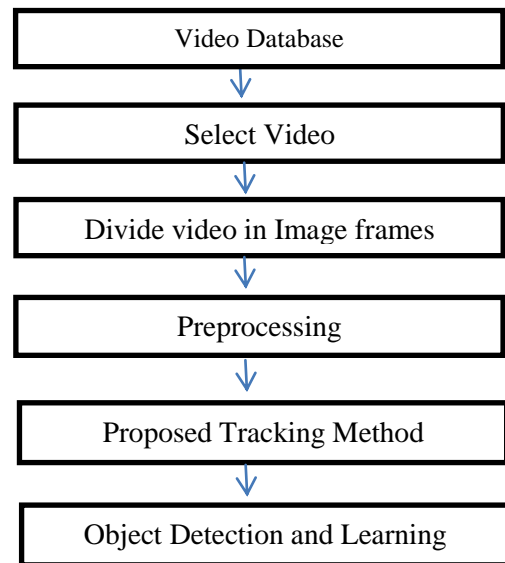


**Fig. 1.1: The architecture of proposed System**

## 3. MEAN SHIFT TRACKING
Mean shift is a non-parametric feature-space analysis technique for locating the maxima of a density function [5]. It is used for analysis of cluster in computer vision and image processing [3]. Mean shift tracking starts from the position of model in the first /current frame (Target Model), and then find the best candidate in neighborhood frame by maximizing similarity function [11] as shown in fig. 1.2.
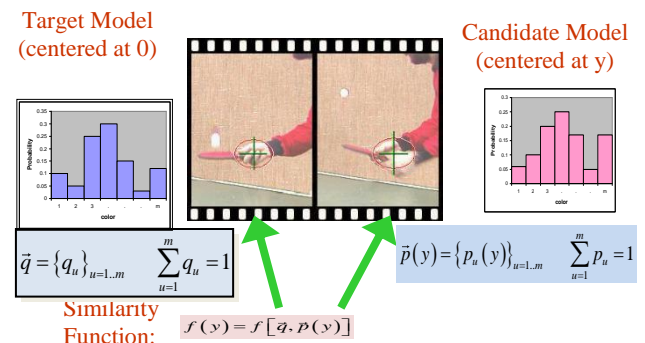


**Fig. 1.2: Mean shift tracking process**

## 3.1 Target Model

Target is usually represented as an ellipse or rectangle in object tracking. We denote by $\{x_i^*\}_{i=1,\ldots,n}$ the normalized pixels in target model, centered at the origin point. $\hat{q}$ is the target model and $\hat{q}_u$ is the probability of the color u in the target model is derived from,

$$\hat{q} = \{\hat{q}_u\}_{u=1,\ldots,m} \tag{1}$$

$$\hat{q}_u(y) = C\sum_{i=1}^{n_h} k\left(\| x_i^* \|^2\right)\delta[b(x_i^*)-u] \tag{2}$$

Where $b(x_i^*)$ the color bin index of $(x_i^*)$, $\delta$ is the Kronecker delta function, k(x) the differentiable, anisotropic, convex, monotonically decreasing kernel profile. The normalization constant C is derived by,

$$C = \left[\sum_{i=1}^{n} k[\| X_i^* \|]^2\right]^{-1} \tag{3}$$

## 3.2 Candidate Model

Let us denote by $\{X_i^*\}_{i=1,\ldots,n_h}$ pixel locations of the target candidate, centered at y in the current frame. By employing the same weighting function, k, the probability of the color u in the target candidate at position y are given by

$$\hat{P}(y) = \{\hat{P}_u(y)\}_{u=1,\ldots,m} \tag{4}$$

by

$$\hat{P}_u(y) = C_h\sum_{i=1}^{n} k\left(\| \tfrac{y-xi}{h} \|^2\right)\delta[b(x_i)-u] \tag{5}$$

Where $\hat{P}(y)$ is the target candidate model, $\hat{P}_u(y)$ is the probability of the color u in the target candidate region, h is the bandwidth of kernel density estimation function; $C_h$ is the normalization function which is independent of y.

$$C_h = \left[\sum_{i=1}^{n_h} k[\| \tfrac{y-x_i}{h} \|]^2\right]^{-1} \tag{6}$$

## 3.3 Similarity Measure

The similarity measure is based on Bhattacharya coefficient [7]. Bhattacharya coefficient is used to find likelihood of the target model and target candidate as follows,

$$\rho[\hat{p}(y),\hat{q}] = \cos\theta_y = \frac{\hat{p}(y)^T}{\|\hat{p}(y)\|\cdot\|\hat{q}\|} = \sum_{u=1}^{m}\sqrt{\hat{p}_u(y)\hat{q}_u} \tag{7}$$

Using equation (5) and (7) we obtain,

The distance between $\hat{P}(y)$ and $\hat{q}$ is defined as,

$$d[\hat{p}(y),\hat{q}] = \sqrt{1-\rho[\hat{P}(y),\hat{q}]} \tag{8}$$

## 3.4 Similarity function (Distance) Minimization

The minimization of distance $d[\hat{p}(y),\hat{q}]$ (8) being equivalent to maximizing Bhattacharya coefficient [7][8] in equation (9).

The search for a new target is initialized with the target position $y_0$. Taylor expression around $\hat{p}(y_0)$ gives,

$$\rho[\hat{p}(y),\hat{q}] \approx \frac{1}{2}\sum_{u=1}^{m}\sqrt{\hat{P}(y_0)}\hat{q} + \frac{1}{2}\sum_{u=1}^{m}\hat{P}_u(y)\sqrt{\frac{\hat{q}_u}{\hat{P}_u(\hat{y}_0)}} \tag{9}$$

$$\rho[\hat{p}(y),\hat{q}] \approx \frac{1}{2}\sum_{u=1}^{m}\sqrt{\hat{P}(y_0)}\hat{q} + \frac{C_h}{2}\sum_{u=1}^{m}w_i k\left(\left\|\tfrac{y-x_i}{h}\right\|^2\right) \tag{10}$$

Where,

$$w_i = \sum_{u=1}^{m}\sqrt{\frac{\hat{q}_u}{\hat{P}_u(\hat{y}_0)}}\delta[b(x_i)-u] \tag{11}$$

To minimize a distance (8), is to maximize the second term in equation (10). The second term represents the density estimate computed with kernel profile k at y, with the data being weighted by $w_i$ (11).

## 3.5 Color and Scale Adaption

The similarity measure we design is based on color, scale and direction information. The probability of color in a target model is given by equation (2) and the probability of color in the candidate model is given by equation (5). The problem is to find the discrete location in the candidate model whose associated density is the closest to density in the target model. We achieve this by using (8).

If the object enlarges or shrinks abruptly in the adjacent frames, then it is difficult to track object accurately. Suppose candidate region is bigger or smaller than target region, then we simply change window size of kernel profile to minimize distance.

## 4. ALGORITHM

The maximization of similarity measure based on mean shift can be achieved by following algorithm,

1) Derive Target model $\hat{q} = \{\hat{q}_u\}_{u=1,\ldots,m}$ and initialize the location $y_0$ of the target model.

2) Calculate target candidate model $\hat{P}(y) = \{\hat{P}_u(y_0)\}_{u=1,\ldots,m}$ in the current frame.

3) Calculate the weight vector $\{w_i\}_{i=1,\ldots,m}$ using equation (11).

4) Derive the new position $y_1$ of target candidate model using following,

$$y_1 = \frac{\sum_{i=1}^{n_*} x_i w_i g\left(\left\|\tfrac{y-x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_*} w_i g\left(\left\|\tfrac{y-x_i}{h}\right\|^2\right)} \tag{12}$$

Update $\{\hat{P}_u(y_1)\}_{u=1,\ldots,m}$ and evaluate

$$\rho[\hat{p}(y_1),\hat{q}] = \cos\theta_y = \frac{\hat{p}(y_1)^T}{\|\hat{p}(y_1)\|\cdot\|\hat{q}\|} = \sum_{u=1}^{m}\sqrt{\hat{p}_u(y_1)\hat{q}_u}$$

5) If $d < \varepsilon$ stop and go to step 1

Where $d \longleftarrow \|y_1 - y_0\|$.

## 5. EXPERIMENTAL RESULTS

The proposed Continuously Adaptive Mean Shift with Color, Scale and Direction algorithm is then tested by using real video sequences. The first video is an ellipse sequence (Fig. 1.3) where the object has clearly a scale and orientation (direction) changes.

The proposed algorithm tracks the ellipse well in the sequence. Scale refers to the size of the object and the orientation is direction of the object. The first frame of the sequence was used to define the target model and the rest frames were used for testing. It can be seen that the proposed method achieves good estimation accuracy.
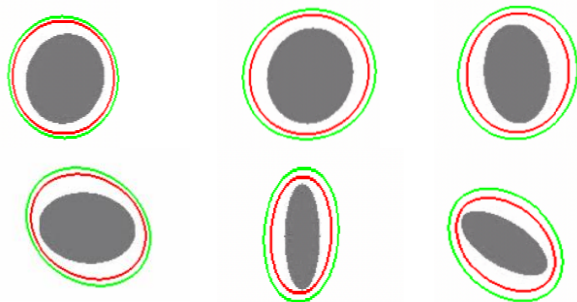


**Fig.1.3: An ellipse sequence Tracking**

The second video is a keratocyte sequence (Fig. 1.4 The algorithm tracks the keratocyte well in the sequence.
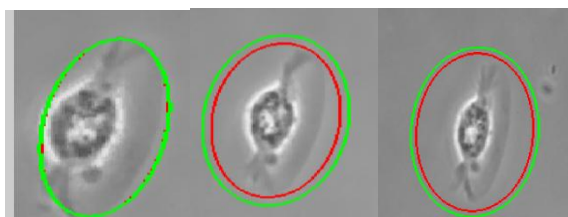


**Fig. 1.4: keratocyte sequence Tracking**

## 6. CONCLUSION

The algorithm uses the advantages of mean shift tracking, such as simplicity, efficiency, real data analysis and robustness [3][12]. The Continuously Adaptive Mean Shift algorithm with Color, Scale and Direction detect and track the object robustly. The proposed system will be used for robots or intelligent agent to detect and track moving color objects in real-time environment, such as human. The system will also use for surveillance, traffic monitoring, investigation and so on.

## 7. REFERENCES

[1] Changjiang Yang, Ramani Duraiswami and Larry Davis, 2005. Efficient Mean-Shift Tracking via a New Similarity Measure.IEEE Computer Society Conference on Computer Vision and Pattern Recognition.

[2] R. Venkatesh Babu, Anamitra Makur, 2007. Kernel-Based Spatial-Color Modeling For Fast Moving Object Tracking. IEEE International Conference on Acoustics, Speech and Signal Processing.

[3] D. Comaniciu and P Meer, 2002.Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[4] Q. Qifeng, D. Zhang, and Y. Peng, 2007. An adaptive selection of the scale and orientation in kernel based tracking. IEEE Conference on Signal-Image Technologies and Internet-Based Systems.

[5] Cheng, Yizong, 1995. Mean Shift, Mode Seeking, and Clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[6] Katharina Quast and Andr´e Kaup, 2010. Shape Adaptive Mean shift Object Tracking Using Gaussian Mixture Model. Image Analysis for Multimedia Interactive Services (WIAMIS).

[7] M. Sohail Khalid, M. Umar Ilyas, M. Saquib Sarfaraz, M. Asim Ajaz, 2006. Bhattacharyya Coefficient in Correlation of Gray Scale Objects. Journal of Multimedia.

[8] Nwe Nwe Soe, 2015.Image Matching Scheme by using Bhattacharyya Coefficient Algorithm. International Journal of Innovative Research in Computer and Communication Engineering.

[9] M. J. Black and A. D. Jepson, 1998. Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. International Journal of Computer Vision.

[10] S. Avidan, 2001. Support vector tracking. In Proc. IEEE Conf. Comp. Vision Pattern Recognition, volume I.

[11] A. Elgammal, R. Duraiswami, and L. Davis, 2003. Probabilistic tracking in joint feature-spatial spaces. In Proc. IEEE Conf. Comp. Vision Pattern Recognition, volume I.

[12] Changjiang Yang, Ramani Duraiswami and Larry Davis, 2005. Efficient Mean-Shift Tracking via a New Similarity Measure. IEEE Computer Society Conference on Computer Vision and Pattern Recognition.