CAMSHIFT-Based Algorithm for Multiple Object Tracking

Sorn Sooksatra and Toshiaki Kondo

School of Information, Computer, and Communication Technology
Sirindhorn International Institute of Technolgy, Thammasat University, Bangkok, Thailand
ohm.sorn@gmail.com, tkondo@siit.tu.ac.th

Abstract. This paper presents a technique for object tracking by using CAMSHIFT algorithm that tracks an object based on color. We aim to improve the CAMSHIFT algorithm by adding a multiple targets tracking function [1].When one object is selected as a template, then it will search objects that have the same hue value and shape by shape recognition. So,the inputs of the algorithm are hue values and shape of the object. When all objects are absent in the frame, the algorithm will search whole frame to find most similar-looking objects and track them. The important task of the object tracking is to separate a target from background or frame that in some cases where the noise is present in the tracking frame. Then object identification method was added to the algorithm for filtering the noise and counting numbers of objects to make decide how many targets track.

Keywords: CAMSHIFT, Multiple target, Object tracking, Shape recognition.

1 Introduction

Nowadays, the technology is growing fast. The application of the technology will be diverse and then interfere of machine is not only keyboard or mouse any more. There is more advance technology that supports more applications at present and in future. Object tracking is one of them. Therefore, object tracking is one of important research topics in computer vision applications.

There are a lot of applications related to object tracking. For example Meanshift tracking for Tracking of Moving Images on an Autonomous Vehicle Testbed Platform in [2] and Armed robot for selecting object based on color and shape of object in [3]. There are other application such as auto-mobile driver assistance, vehicle navigation, robotics, human-computer interaction, video surveillance, biometric, video games, industrial automation and security. There are several techniques of object tracking. A certain technique can be usable and suitable for each application. The most popular one is CAMSHIFT that has been used for this research paper.

The CAMSHIFT (Continuously Adaptive Mean Shift) that is detailed in [4] is derived from the Mean Shift algorithm and it is one of popular techniques

P. Meesad et al. (Eds.): *IC²IT2013*, AISC 209, pp. 301–310.
DOI: 10.1007/978-3-642-37371-8_33

that are used for object tracking. This algorithm is based on object hue value from HSV image. It has been applied as basis and core component in many applications.

This technique is a robust nonparametric technique for finding the peak probability distribution. Then, the target color will have high probability and other will be almost zero. We use this for separate target from background. Therefore, CAMSHIFT will be most efficient when target's color is totally different to background's color.

There are five main drawbacks of traditional CAMSHIFT. First, its tracking failure when color of background and target are similar. Second, object can change their appearance anytime (rubric example in [5]) that makes it difficult to track a target by using only one histogram. Third, traditional CAMSHIFT cannot track more than one target. We can use multiple tracker for tracking multiple object same in [1]. Fourth, traditional CAMSHIFT can track similar color object that overlap real target instead. Lastly,when saturation is low. Hue will becomes quite noisy, because of small range in Hue, the small number of discrete hue pixels cannot adequately represent slight changes in RGB [4].

The main aim of this paper is to modified the CAMSHIFT to be able to track more than one target, identify the shape of object and make CAMSHIFT work every situation even if the color of a target is similar to that of background and background's color being for quite same color.

2 CAMSHIFT Algorithm

In the CAMSHIFT process as shown in Fig. 1, the process will collect the data or parameters of each step for update data in each frame since video consists of many frames. The important parameter is location of search window and image order moment. The process is following

2.1 Select Region of Interesting (ROI)

This step will find the initial location of target located in the frame and obtain data related to hue value in the target for making color histogram in the next step. The initial location will be used for calculation in tracking process that starts in this region to avoid other object in the frame.

2.2 Create Color Histogram

Histograms of each color will be created. The height of each column represents number of pixels in a frame region having that "hue." Hue is one of the three values describing a pixel's color in the HSV (Hue, Saturation, and Value) color model as shown in Fig. 2(b).

This will take number of pixels with the Hue value within ROI. These data will be used for generating probability map process.

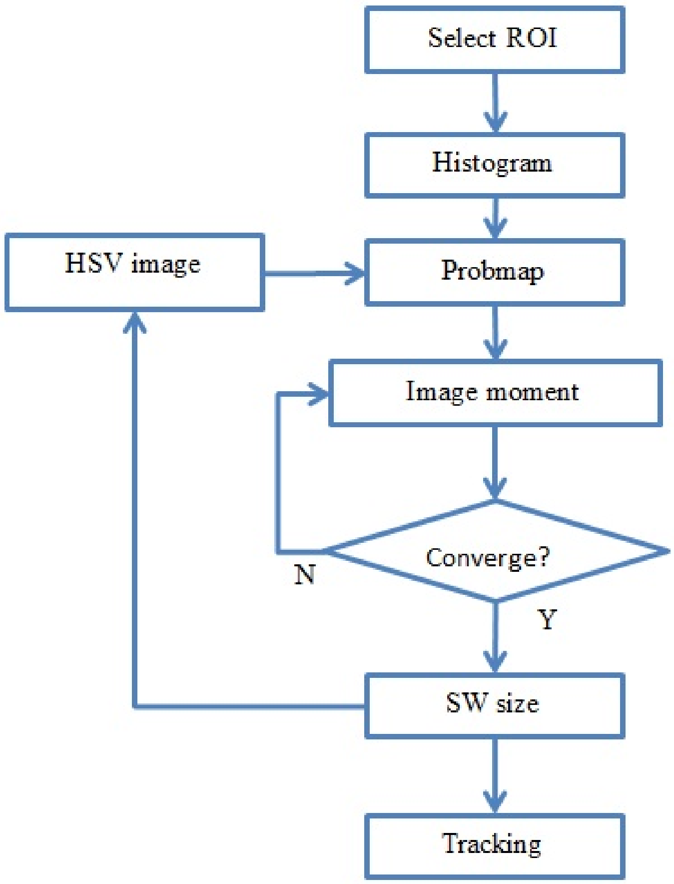


Fig. 1. CAMSHIFT Flowchart

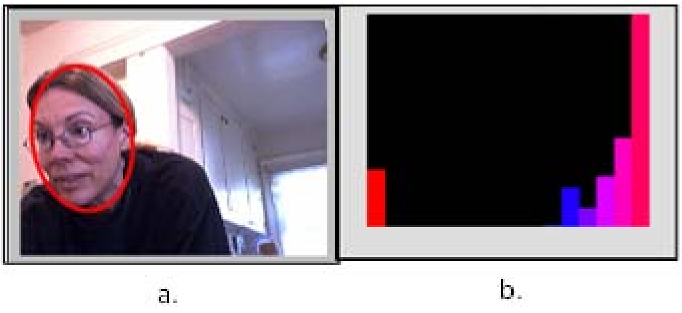


Fig. 2. (a) input image and region of interest (ROI) and (b) Histogram of hue within the ROI [6]

2.3 Generating Probability Map (probmap)

This is the first step in the loop of CAMSHIFT technique. The probability map shows the probability of each pixel in each frame used for separating target and background. Therefore, probability map is the heart of this technique.

For probability, it is from number of colors that obtained from previous step by following equation.

$$probmap(row, col) = numberof(hue(row, col)) \tag{1}$$

$$probmap = probmap/max(probmap) \tag{2}$$

$$probmap = probmap \times 255 \tag{3}$$

where row is vertical location of the frame and col is horizontal location in the frame. Since this algorithm is based on target that has colourful color or high saturation in HSV model as shown in Fig. 3. So, it ignore pixel that has low saturation for calculating in probability map (i.e. make it to be zero).

2.4 Shift New Location

Since the target can always move, we must find the new centroid within the search window (SW) by computing 0th (M_{00}), 1st (M_{10}) and 2nd (M_{01}) order image moment within the ROI by following equation

$$M_{00} = \sum_{row}\sum_{col} probmap(row, col) \tag{4}$$

$$M_{10} = \sum_{row}\sum_{col} col \times probmap(row, col) \tag{5}$$

$$M_{01} = \sum_{row}\sum_{col} row \times probmap(row, col) \tag{6}$$

$$X_C = \frac{M_{01}}{M_{00}}; Y_C = \frac{M_{10}}{M_{00}} \tag{7}$$

where X_C and Y_C is location of the new centroid of SW. This step will be repeated until it converges.

2.5 Calculating the Size of Tracking Window

The size of tracking window depends on the size of the target or zeroth order image moment (M_{00}) calculated from previous step. Then size of SW is equal to function of M_{00} that in following equation.

$$H = \sqrt{\frac{M_{00}}{255}}; W = 1.2 \times H \tag{8}$$

where H and W are the height and width of tracking window respectively. This equation comes from Ref. [4] that are adjusted for our implementation.

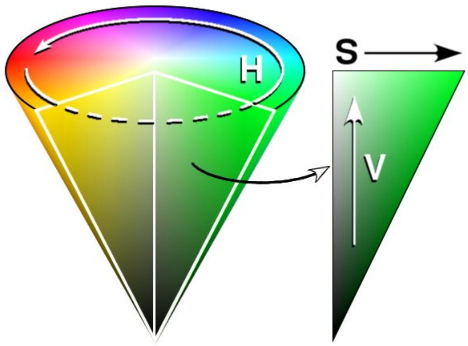


Fig. 3. HSV color model

2.6 Tracking Process

For the tracking window, it will locate the tracking window by the location that gets from previous step. When the target is disappear on the frame. It will track whole frame instead of SW again same as in [5] until target enter in the frame again. After this step, the process will get back to step C again.

3 CAMSHIFT Improvement

In this paper, we tried several ways to improve CAMSHIFT to support more applications. We tried to solve the noise problem in the probability map of CAMSHIFT and add more feature with multiple target tracking and object identification. The detail of improvement is following.

3.1 Multiple Tracking Technique

The main purpose of this paper is to add multiple tracking ability to CAMSHIFT technique. According to today's applications. The object tracking does not track single or only one target anymore. There are application that use multiple tracking such as counting the object, track many human faces in the camera and other.

The technique that we use is making all parameters used for tracking into array. As mentioned before, object tracking needs data about related to location and size of the SW. That information comes from minimum row (rmin), maximum row (rmax), minimum column (cmin), maximum column (cmax) and 3 image order moment .The format of array is following:

$$P[i] \in P[1], P[2], P[3], ..., P[n] \tag{9}$$

$$T[i] \in T[1], T[2], T[3], ..., T[n] \tag{10}$$

where P is one of parameters , T is one of target and n is number of target in the frame.The initial search area of T[1] is in the ROI and other T[i] is the whole frame and it will be reduced if found the target.The process is same as traditional CAMSHIFT for each T[i]. The calculation of P[i] of T[i] is processed in order of array (10) (i.e. order T[1] to T[n]) not at the same time. When T[i] is finished to obtain the P[i], it will set the region of SW of T[i] in the probmap to be zero. This process will help for protecting repeated or same location of T[i+1] and so on.For the order of the target, it depends on size or density (M_{00}) of the target. If target has the largest size in the probmap, it will be tracked first. However, the first target (T[1]) is normally be the object that in the ROI on the initial frame of video.Tracking window appearance is depending on the size (M_{00}) of each T[i] . If size of T[i] is very small (i.e. target is far away from camera or disappear in the frame), SW will not appear in the frame and T[i] will search whole frame again.There are some rule for this multiple tracking technique as following.

1. First target that enter to the frame or in ROI region will be MT, Second will be ST-1, third will be ST-2 and so on. If two or more ST has already in the frame, it will choose order for ST randomly.
2. When MT disappears in the frame, it will choose one of ST to be MT.
3. When center of each tracking window are close enough, it will merge 2 or more target to be one target.

3.2 Object Identification Method

This method is used as noise filter in probability map. This method is concern about the size of object both noise and target. In most situations, size of noise is less than target size. We can eliminate noise concerning size of object by using [7] in this process as follows

1. The algorithm will search from top-to-bottom scan order in probability map.
2. When it found connected or only one pixel that has intensity more than zero. They will be counted as one object.
3. The size of each object is equal to number of pixel of that object.
4. The algorithm will filter out the object that has size less than threshold (noise object).

This method can be used as object counting by counting the object after filter out.We must adjust size of object for this method to make most efficacy for object tracking.

3.3 Shape Recognition

For the shape recognition, this process begins after tracking process. It help identify the shape of the target (T[i]) by using compactness theorem in [8] and [9]. Eq.(11) show the equation of compactness calculation.

$$C = \frac{P^2}{4\pi A} \tag{11}$$

where C is compactness, P is perimeter and A is area of the target. We use region of SW of each target in probmap as image input because probmap shows the target clearly form the background. Then, we use this to check whether shape of T[i] is same as template that we select in ROI or not. If difference of compactness between those two is too much, it will identify that this target is not target that we are interest due to the shape.

4 Experiment Results

At this point, this paper implement algorithm that can track more than one target at that time. Since we add more feature to this algorithm, then the flow chart is changed as in Fig 5. Fig. 4 shows the result about tracking window and probability map of traditional CAMSHIFT and improved CAMSHIFT. It shows that improved CAMSHIFT has less noise but less density of the target when compare with traditional CAMSHIFT and It can track more than one target. However, it has the problem when two target overlap to each other, then the order of the two targets are switched.

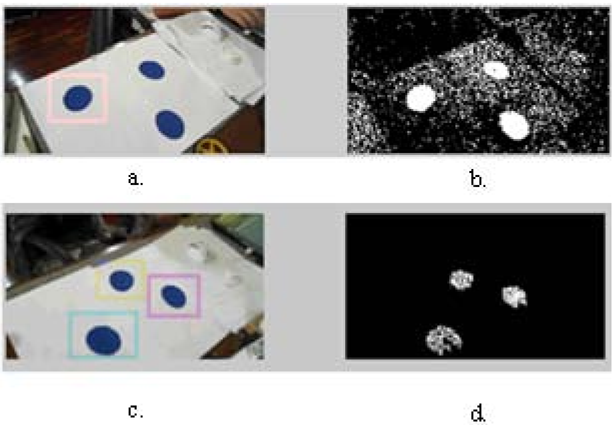


Fig. 4. (a) Input frame,(b) probmap of traditional,(c) input frame, and (d) probmap of improved

Fig. 6 shows the comparison between this algorithm without shape recognition (Fig. 6.a)and with shape recognition (Fig. 6.b). It shows that when a target is found, it checks the shape of the target if the target matches to template or not (Template in this case is circle shape). Therefore, square and triangle shaped objects are not be tracked. However, this tracker is quite sensitive when object's shape is changed due to environment. For example, saturation changed by the light reflection or object is shown only some part that causes the shape of object in probmap be changed.

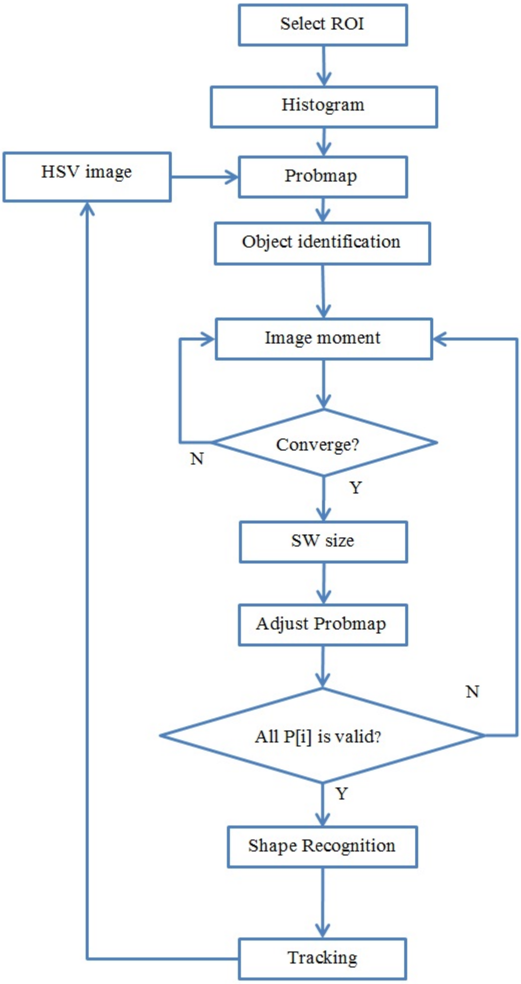
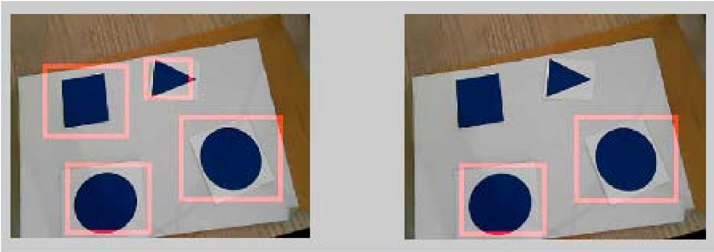


Fig. 5. Improved CAMSHIFT flowchart



a. b.

Fig. 6. (a) Improved CAMSHIFT without shape recognition and (b) improved CAMSHIFT with shape recognition

5 Conclusion and Future work

The improved CAMSHIFT in this paper still uses the same concept of the traditional CAMSHIFT that tracking is based on color. We added a multiple objects tracking function and shape recognition for finding more applications. In addition, it has object identification for eliminating noise in the frame. One possible weakness of the proposal method is that the probabilities of a small object may be reduced in the noise elimination step.

As future work, we plan to introduce a more advanced shape recognition technique method in order to handle more complicated shapes. We will also work in speed-up the proposal method.

Acknowledgements. This research is financially supported by Center of Excellence in Biomedical Engineering, Sirindhorn International Institute of Technology (SIIT), Thammasat University (TU) and the National Research University Project, Thailand Office of Higher Education Commission.

References

1. Hidayatullah, P., Konik, H.: CAMSHIFT Improvement on Multi-Hue and Multi-Object Tracking. In: International Conference on Electrical Engineering and Informatics, Bandung, Indonesia (2011)
2. Gorry, B., Chen, Z., Hammond, K., Wallace, A., Michaelson, A.: Using Mean-Shift Tracking Algorithms for Real-Time Tracking of Moving Images on an Autonomous Vehicle Testbed Platform. World Academy of Science, Engineering and Technology (2007)
3. Amores, P.M.M., Gagwis, J., Jamora Marie, A.: PIC – Based Color And Shape Recognition: Using 3-Link Robotic Arm (2011)
4. Bradski, R.G.: Computer Vision Face Tracking For Use in a Perceptual User Interface. Microcomputer Research Lab. Santa Clara, CA. Intel Corporation (1998)
5. Exner, D., Bruns, E., Kurz, D., Grundhofer, A.: Fast and Robust CAMSHIFT Tracking. Bauhaus-University Weimar, Germany (2010)

6. Cognotics. How OpenCV's Face Tracker Works (2007), http://www.cognotics.com
7. Mathwork. Remove small objects from binary image (2012), http://www.mathworks.com/help/images/ref/bwareaopen.html
8. Pomplun, M.: Compactness (2007), http://www.cs.umb.edu/~marc/cs675/cv09-11.pdf
9. Zhongwan, L.: Compactness theorem. Mathematical Logic for Computer Science 2, 147–158 (1998)