# LOW COMPLEXITY ADAPTIVE VIDEO CROPPING SYSTEM FOR SPORTING APPLICATIONS USING INTELLIGENT OBJECT TRACKING

W.A.N.A. De Alwis[1], G.A.N. Perera[1],R.W.M.R.C.J. Wanigasekara[1], M.B. Dissanayake[1]

[1]Department of Electrical & Electronic Engineering, Faculty of Engineering, University of Peradeniya

## ABSTRACT

A tool is presented for generating a new video, which concentrates only on a user specified player, using the original video which consisted of multiple players. By using this system, coaches or players will be able to monitor individual players clearly. Operation of this tool can be lined up into two main stages. Initially, tracking of the player is done throughout the video, using mean-shift algorithm. The proposed tracking mechanism has the advanced capability of continuing the tracking even if the object is temporary lost. After the tracking process is completed, new video is generated by cropping the relevant area around the player, in each frame. This tool can be further developed for a wider range of applications such as movies, news, games, etc. to suite for heterogeneous network resources.

***Key words*****:** Mean-shift, tracking, adaptive cropping

## 1. INTRODUCTION

In many sporting events, it is important to monitor each player's actions closely to maximize his/her performance. Coaches or players view video record of an entire game, captured by the broadcasters for this purpose. Also, the recorded versions of the videos are of higher resolution and will not fit properly to a portable device like PDA. As a result, it is difficult to monitor an individual player in such a video. It will be very much useful to have a separate version of the video which concentrates only on the specified player.

Reference [1] highlights the importance of client centered multimedia content delivery in heterogeneous network environments. This involves Region of Interest (ROI) based tracking and adaptive cropping of videos to cater the client demand. There are several techniques available in literature for object tracking, for example blob tracking, mean-shift tracking and contour tracking. Further, a novel technique for tracking ROI ofsport videos is highlighted in [2]. An advance techniques for Client or Server based ROI prediction for high resolution streaming video are presented in [5]. A self-adaptive image cropping technique for small displays is presented in [4].

The tool presented in this paper tracks a user specific object (a player in this case) in a video and produce a cropped version, which concentrates only on the object specified earlier. It uses the mean-shift tracking for ROI tracking, considering its advantages [3].

## 2. METHODOLOGY

The overall system architecture is shown in Figure 1. In the approach presented in the paper, color tracking was initially used to track the player. But when background color is closer to the player's clothing, it is difficult to continue the tracking. For example, according to Figure 2(a) it is difficult to track the player as the background color is closer to the clothing of the player. But as in Figure 2(b) the follower can be tracked easily since the colors can be separated clearly in that corresponding video.

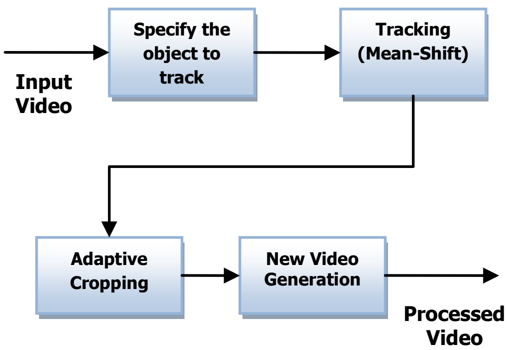



**Figure 1: Overall System Architecture**

(a) Tracking a Player (b) Tracking Flower

**Figure 2: Tracking using Color**

**Figure 3: Tracking using Mean-Shift Algorithm**

The mean-shift algorithm can be used to overcome this problem [3]. It is an iterative scheme, which is based on comparing the histogram of the original object in the current image frame and histogram of candidate regions in the next image frame. The aim is to maximize the correlation between two histograms. Object tracking for an image frame is performed by a combination of histogram extraction, weight computation and derivation of new location.

Assuming $g(x) = -K'(x)$ we have,

$$m(x) = \frac{\sum_{i=1}^{n} g\left(\frac{x - x_i}{h}\right) x_i}{\sum_{i=1}^{n} g\left(\frac{x - x_i}{h}\right)} - x \qquad (1)$$

Where K' is the kernel and the quantity $m(x)$ is called as the mean shift.

Mean-shift procedure can be summarized for each point $x_i$. First mean-shift vector $m(x_i^t)$ is calculated using the above equation. Then the density estimation window is moved by $m(x_i^t)$. This process is repeated till the convergence.

By increasing the number of iterations, accuracy of the video tracking can be increased. However incrementing of these iterations require more processing power. To suit the processing power of the portable device, the developed tool is optimized to perform five iterations. Further, the system assumed that the color distribution of the video frame has a Gaussian distribution as given in (2).

$$\phi(x) = e^{-x^2/2\sigma^2} \qquad (2)$$

## 3. RESULTS

First, the time taken for generation of new video frames is considered using videos with different qualities as the input. From the results in Table 1, it is observed that the time taken for generating a frame is proportional to the quality of the input video. However, it did not exceed 0.20 seconds for most of the tested videos. This is a significant measurement, because the server should beable to react to the client's request with as little latency as possible.

**Table 1: Video Generation times with Different Quality Input Videos**

| Input Video Quality | No of Frames | Generation Time/(s) | Time per frame/(s) |
| --- | --- | --- | --- |
| Mobile | 184 | 25.38 | 0.1379 |
| Standard Definition | 263 | 51.67 | 0.1965 |
| | 272 | 53.52 | 0.1968 |
| High Definition (HD) | 152 | 34.34 | 0.2259 |

The quality of the generated video is then tested by comparing an automatically generated frame (using the tool) with a manually cropped frame. Since objects of the same shape are considered here, color histogram comparison is used to check the similarity between automatically generated frame and the manually cropped frame.

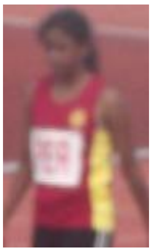
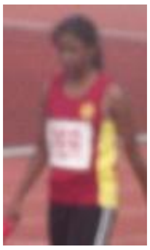
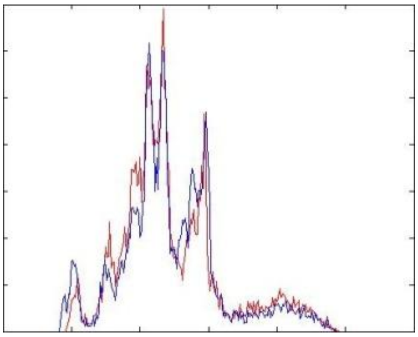
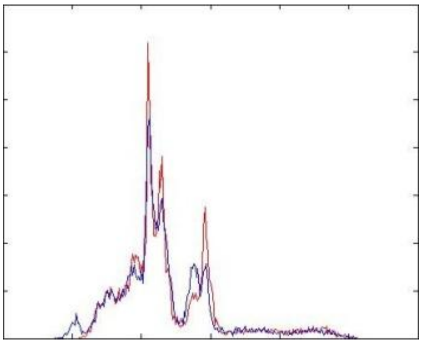
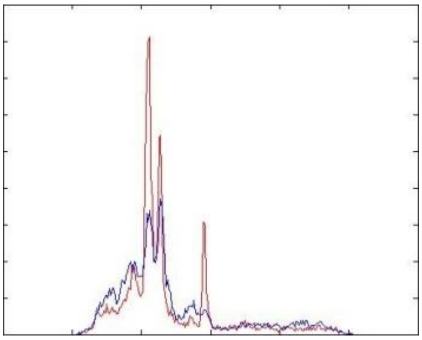
Table 2 shows the comparison of three random frames, generated automatically and manually. In the histogram plot, x-axis represents a particular value/intensity, wherey-axis represents the number of pixels of that particular value/intensity. From the histogram results, it can be seen that both of the above frames are nearly the same. So the quality of the video is preserved at a tolerable level when the video is cropped automatically using the tool, while the efficiency of the system is considerably increased compared to the manual system.

Further, the Figure 4 shows a test running of the proposed tool for football application, where the player in blue jersey is the object of interest.

## 4. CONCLUSION

In this paper, an intelligent tracking based video cropping system for sporting applicationsis proposed. With this proposed system, user can clearly observe the player he/she desire rather than watching the whole video which consists of many players and different background information.

**Table 2: Histogram Comparison between Automatic and Manual Frames**

| Automatically Cropped Frame | Manually Cropped Frame | Histogram Comparison (Red: Automatically Cropped Blue: Manually Cropped) |
|---|---|---|
|  |  |  |
|  |  |  |
|  |  |  |

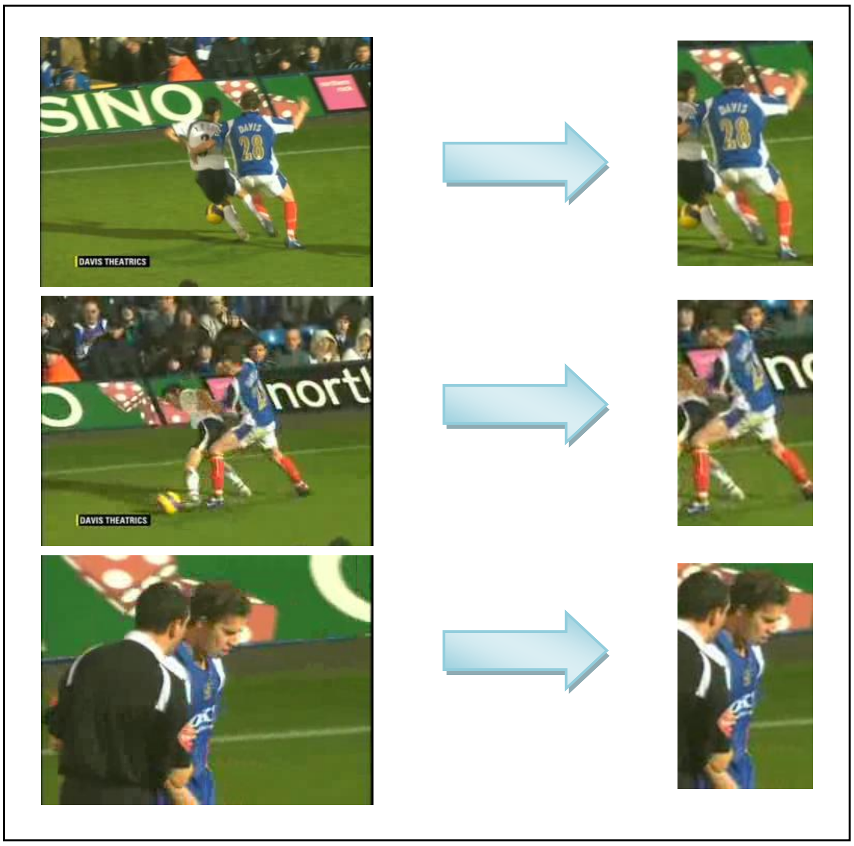



**Figure 4: Input Video Frames and the Cropped Video Frames after Tracking**

The proposed system can be extended to tracking objects in a video such as movie, news, etc., with the adaptive cropping mechanism. For an instance, when the selected object reaches the edge of the video, rather than cropping the area around the specified object, downsizing the video for that particular period of time would be proper. This system can be further improved for a real-time video, which can be useful when viewing video on different display settings, such as on mobile or on computer or on television.

## 5. REFERENCES

[1] Yong Wei, Suchendra M. Bhandarkar, Kang Li, August 2009,'Client-Centered Multimedia Content Adaptation', *ACM Transactions on Multimedia Computing, Communications and Applications*, Vol. 5, No. 3, pages 1-24

[2]António M. G. Pinheiro,"Tracking Sport Video Shots Region of Interest using the Angular Orientation Partition Descriptor", Proceedings of TrakMark2012, 21st International Conference on Pattern Recognition, 11th, November 2012, Japan

[3] Zhi-Qiang,Wen,Zi-Xing Cai, August 2006, 'Mean Shift Algorithm and Its Application in Tracking of Objects', *Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, 13-16 August 2006*

[4]Ciocca, G.; Cusano, C.; Gasparini, F.; Schettini, R., "Self-Adaptive Image Cropping for Small Displays," Consumer Electronics, 2007. ICCE 2007. Digest of Technical Papers. International Conference on, vol., no., pp.1,2, 10-14 Jan. 2007

[5]Mavlankar, A.; Varodayan, D.; Girod, B., "Region-of-interest prediction for interactively streaming regions of high resolution video," Packet Video, vol., no., pp.68,77Nov. 2007

[6] D. Comaniciu, V. Ramesh, 'Real-time Tracking of Non-Rigid Objects using Mean Shift', *Computer Vision and Pattern Recognition, Proceedings. IEEE Conference, 2000*, Vol. 2, pages 142 – 149