A Fast Algorithm for Detecting, Labeling and Tracking Volleyball Players in Sport Videos

Hananeh Salehifar
Faculty of Electrical, Computer, and IT
Islamic Azad University, Qazvin Branch
Qazvin, Iran
h.salehifar@qiau.ac.ir

Mohammad Mahdi Dehshibi
Faculty of Electrical, Computer, and IT
Islamic Azad University, Qazvin Branch
Qazvin, Iran
mohammad.dehshibi@yahoo.com

***Abstract*—this paper aims at successful tracking of volleyball athletes during competition using only a single camera. Due to the wide range of possible motions and non-rigid shape changes, the tracking task becomes quite complex. We propose a novel method based on adaptive background subtraction by two concurrent frames. The proposed method detects and labels with high precision. It is a reliable tracking in cases of facing conditions such as multiple player occlusions.**

Keywords-human motion;tracking; volleyball athletes

I. INTRODUCTION

In recent years, new trends have aroused for computer vision and image processing. These new trends are more practical. One the remarkable application is analyzing the players' activity in different sports. Locating, labeling, and tracking players have broad application, especially for a team analyzer, and in the broadcast sports videos industry. There are several issues which make this kind of video analyzing a challenging task. Players' occlusion, similar player's appearance, various numbers of players, unexpected camera motion, noise, and video blur are the samples of this difficulty. Many algorithms have been proposed to deal with the multiple target-tracking (MTT) problem in which Particle Filter [1], [2], Joint Probabilistic Data Association Filter (JPDAF) [3], Multiple Hypothesis Tracking (MHT) [4], MCMC data association [5], [6] and track linking [7], [8], [9] are used. Several researchers also study the specific problem of labeling and tracking of players in sports video [2][10][11]. In [10], a clustering based trajectory matching method is proposed to track players in a soccer video. In this work, labeling of individuals is achieved through a supervised classification method. Comaniciu et al [11] build a track graph, and take the tracking problem as an inference problem in a Bayesian network. In both works, a multi-camera system is used to get a fixed, high-resolution and wide-field view of a soccer game. These pre-defined settings ensured a reliable background subtraction system, which are not very practical approach.

This paper proposes a new solution for detecting, labeling, and tracking of players in Volleyball videos. This algorithm performs adaptive background subtraction on two concurrent frames in the sequence.

This paper is organized as follows: Labeling and Tracking algorithms are described in section II. Experimental results and implementation are reported in section III. Finally, this paper is concluded in section IV.

II. TRACKING ALGORITHM

In most cases, different movie, which should be analyzed, have different number of frames; thus, the proposed method first determines the number of frames related to the input film and extracts frames as still images. As the proposed method contains an iterative procedure, in order to save memory space and speed up the algorithm, just one frame loads into memory during each iteration. Table 1 shows the functionality of the proposed method:

Table 1: Adaptive Background Subtraction procedure.

Adaptive Background Subtraction Algorithm
1. Initialize step
/* x and y are the pixels' position */
B(x, y) = First Frame;
2. Subtract Next Frame From *B*
D(x, y) = (B(x, y) +1) –(B(x, y));
3. Filtering step
/* α is threshold */
$I(x,y)=\begin{cases}0, & D(x,y)<\alpha\\ 1, & D(x,y)\geq\alpha\end{cases}$
4. Assign new frame:
B(x, y) = B(x, y) +1;
While there is any frame
GOTO Step 1;

In order to enhance highly connected "blobs" in the image, dilation operator is applied to the filtered image. Then, each blob is ANDed with the subtracted background image to isolate the connected components. The objective of this step is allowing the algorithm to draw the smallest possible rectangle around the moving object. In this algorithm, the blobs that their size is smaller than 100 pixels assume as noise and do not consider by the object's rectangle.

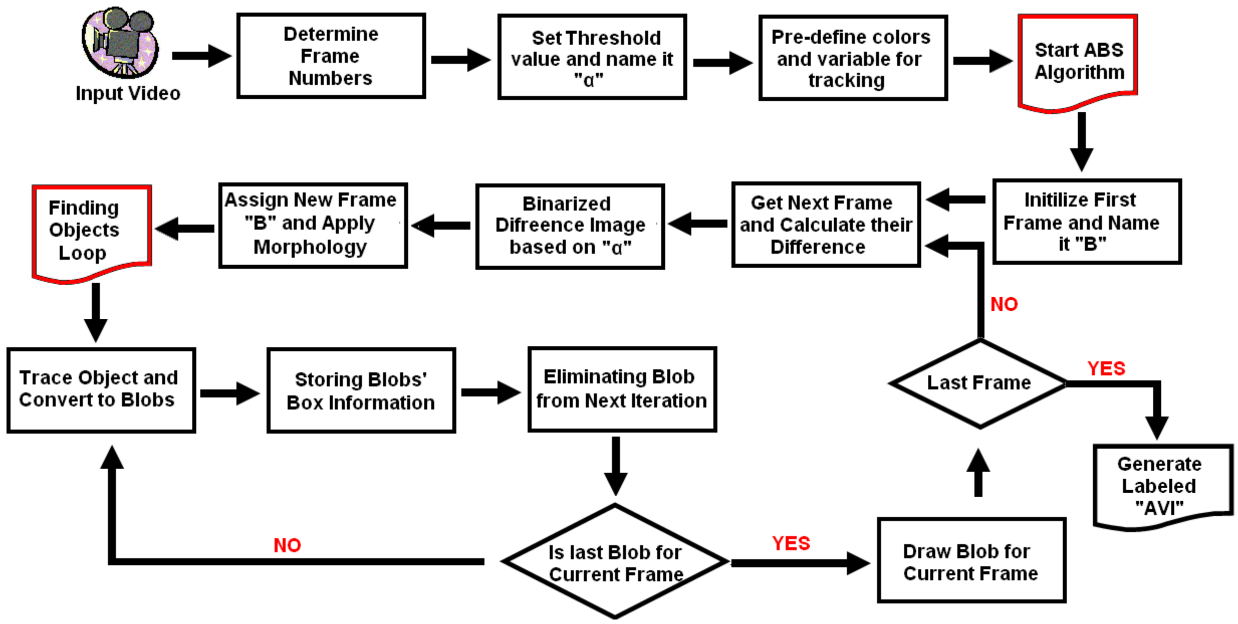


Figure 1.the proposed tracking algorithm scheme.

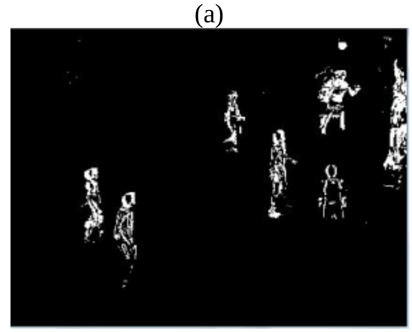
Finally, cross correlation between the normalized blob and a list of known blobs from previous fames is calculated to find an adaptation. If any matching can be found then the color corresponds to the known patch is used as the rectangle color.

The known patch is then updated with the new image by cutting out the middle of the blob. If there are too many failures to track an individual object, i.e. if the known image list gets too large, the algorithm will fall back into initial step. However, in this step yellow color rectangles are utilized. In order to detect each blob once, a clever method should be defined. To achieve this, every blob is removed from the binary image before the next iteration of the blob detection. In this manner, blobs are never detected more than once and as a result, only one yellow rectangle is drawn around each moving object.

In order to find out boundary of the moving objects and to draw rectangles around them, a sole algorithm is developed. The aim of the proposed algorithm is calculating the bounds of the connected components which identified as a single entity. For each frame, bounds for multiple boxes are stored and then drawn on the original frame using a pre-defined color. After applying the whole procedure to the last frame of video, the output labeled video is generated and saved with “avi” format. Overall procedure of the proposed method is illustrated in Figure 1.

The structure of the proposed method is such designed that any changes in motion across the sequence of videos can be detected. Therefore, if there is no moving in the sequence of the frames, we will lose the still object. Figure 2 shows since the player in the right side of frame has no movement, the algorithm cannot detect and track it.

(a)



(b)

Figure 2.Changing in image detection process for (a) original image, (b) the difference image.

III. IMPLEMENTATION

In order to implement the proposed algorithm (ABS), two kinds of data related to the Volleyball sport, where the cameras are fixed, are used. The main reason that we chose fixed camera traced back to the nature of the proposed method; because the method should be apply on difference of two consecutive frames. As it is evident in Figure 3, utilizing ABS players are detected and are labeled in consecutive frames. Consequently, tracking players in two concurrent frames in the sequence is achieved.

Figure 3 .players and ball are successfully tracked by using the proposed system, color rectangle show the location where the object detected.

There are two major issues should be considered during implementation, which we describe them in detail in this section. Initially we design our program to load all of the image frames at the start of the program. This worked fine during the early development of our algorithm, taking only twenty seconds or so to load all of the images. However, as we further developed our code to perform Adaptive Background Subtraction (ABS), loading all of the images at the start of the program consumes too much memory. To solve this problem, we made the decision to revise how we loaded the image frames. To achieve this, loading frames left on the fly as the adaptive background subtraction was being performed. In this manner we saved memory and the proposed method is executed significantly faster.

Another issue is choosing a correct and accurate threshold for the adaptive background subtraction process. When the background template was subtracted from the current image frame, we had to find a value for the threshold that would eliminate most of the outliers. We settled on a threshold value of 15. Although this value eliminates a lot of outliers caused by noise in the image frames while at the same time depicting the moving objects which we were interested in, it was the optimum value. As conducted experiments are shown, those values that were less than 15 remove both detected noisy pixels as well as other moving objects in the scene. On the other hand, threshold values higher than 15 cause rendering real moving objects in the scene hardly distinguishable.

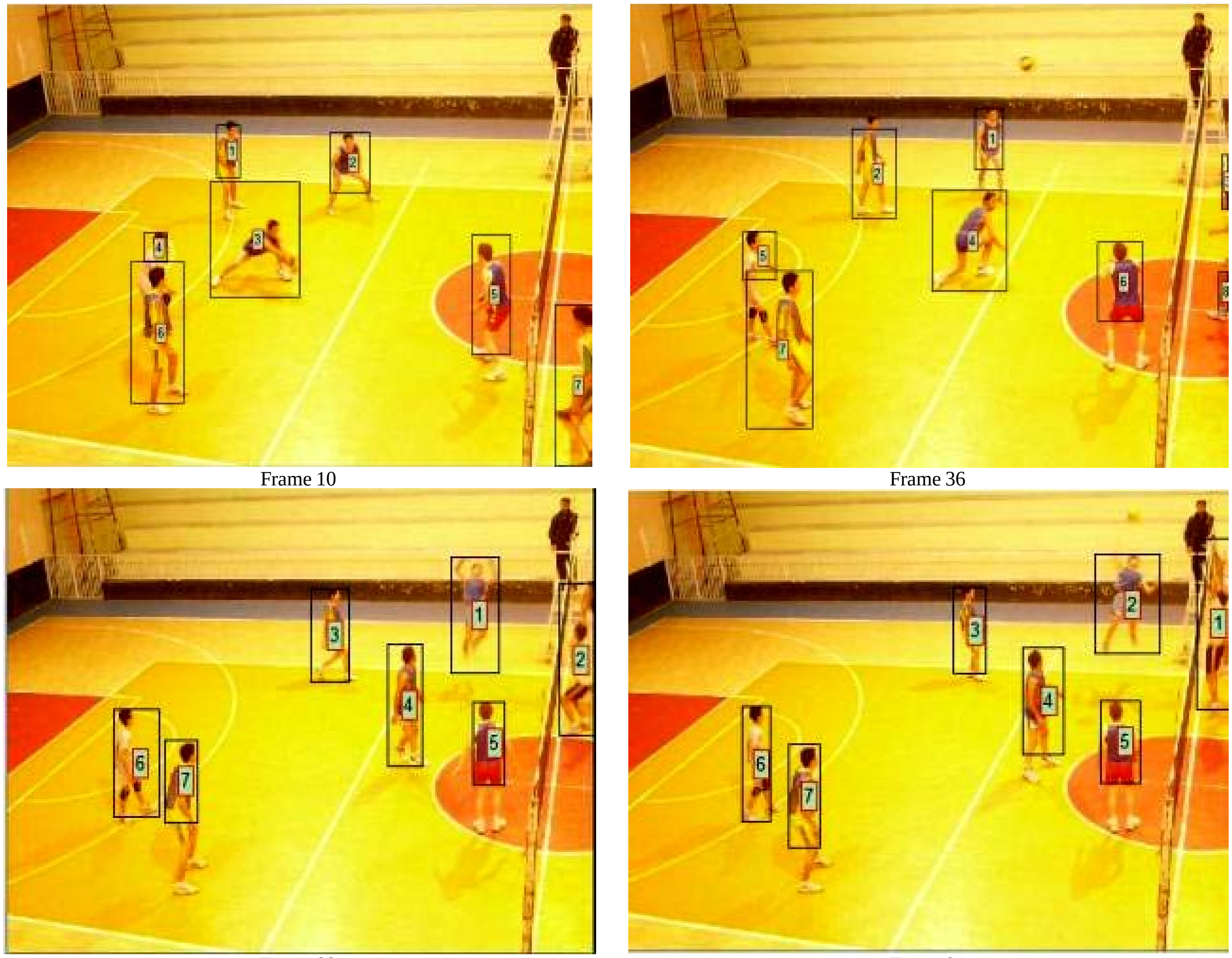
The last issue that we face was determining a threshold value for blob detection process. In our blob detection loop, blobs with less than 100 pixels are discarded and yellow box are not drawn around them. We use this threshold value because experiments show that this value can successfully exclude much of the noisy pixel in the video frames that was enhanced by applying dilation process.

IV. CONCLUSION

This paper sorts out the problem of tracking and labeling of volleyball athletes during competition using only a single camera. Due to the wide range of possible motions and non-rigid shape changes, the tracking task becomes quite complex. We propose a novel method based on adaptive background subtraction on two concurrent frames in the sequence.

With the best of our knowledge, occlusion is a common problem encountered in tracking multiple people. This might be due to complete or partial overlap of a person by another person. Another contribution to this work is to overcome the problem of partial occlusion. Adaptive Background Subtraction (ABS) can detect players with separating connected component in the most frames. The missed detection by the proposed system referred some complex situations due to full occluded objects with each other since the proposed system deals only on the partial occlusion. Another reason for the missed detection by proposed system is the image resolution. In this work, the system deals with a low resolution video scenes to increase the accuracy and execution time of the system; hence this amount of occlusion is natural and also it is less than the existing methods.

As we discussed earlier, when the objects move, the algorithm works very well. However, when an object stops moving, the box around it goes away. This is because the differencing is adaptive and non-moving objects will become the background image. Another benefit of the proposed method is diagnosis ball in the game. The ball always moves in the consecutive frames and it can be detected very well utilizing ABS. In conclusion, this fast and efficient algorithm which we called Adaptive Background Subtraction (ABS), can detect players and ball in analyzing a Volleyball game. Figure 4 demonstrates the ability of the proposed method in tracking and labeling the athletes and other moving objects, i.e., ball.



Frame 10 Frame 36

Frame 89 Frame 91

Figure 4. this figure demonstrates the ability of the proposed to track people in complex situation

REFERENCES

[1] J. Sullivan and S. Carlsson, "Tracking and Labeling of Interacting MultipleTargets," In Proc. European Conf. on Computer Vision (ECCV), 2006.

[2] Thomas Mauthner1, Christina Koch2," Visual Tracking of Athletes in Beach Volleyball Using a Single Camera", International Journal of Computer Science in Sport,2008.

[3] M. Isard and J. MacCormick, "BraMBLe: A Bayesian Multiple-Blob Tracker," In Proc. IEEE Int'l Conf. on Computer Vision (ICCV), 2001.

[4] K. Okuma, A. Taleghani, N. de Freitas, J. J. Little, and D. G. Lowe, "A Boosted Particle Filter: Multitarget Detection and Tracking," In Proc. European Conf. on Computer Vision (ECCV), 2004

[5] D. Reid, "An Algorithm For Tracking Multiple Targets," IEEE Transaction on Automatic Control, vol. 24, no. 6, pp. 843–854, December 1979.

[6] S. Oh, S. Russell, and S. Sastry, "Markov Chain Monte Carlo Data Association for Multiple-Target Tracking," University of California, Berkeley, Technical Report UCB//ERL M05/19, June 2005.

[7] Q. Yu, G. Medioni, "Map-Enhanced Detection and Tracking from a Moving Platform with Local and Global Data Association," IEEE Workshop on Motion and Video Computing (WMVC'07), 2007.

[8] A.G.A. Perera, C. Srinivas, A. Hoogs, G. Brooksby, WenshengHu,"Multi-Object Tracking Through Simultaneous Long Occlusions and Split-Merge Conditions,"IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR), 2006.

[9] C. Stauffer, "Estimating Track Sources and Sinks", In Proc. IEEE Workshop on Event Mining in Video, 2003.

[10] P. Nillius, J. Sullivan and S. Carlsson, "Multi-Target Tracking – Linking Identities using Bayesian Network Inference," In Proc. IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR), 2006.

[11] D. Comaniciu, P. Meer, Mean shift: a robust approach toward feature space analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (5) (2002) 603–619.

[12] L. Vincent, P. Soille, Watersheds in digital spaces: an efficient algorithm based on immersion simulations, IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (6) (1991) 583–598.

[13] B. Sumengen, Variational image segmentation and curve evolution on natural images, Ph.D. Thesis, University of California.