# Video Object Extraction by Using Background Subtraction Techniques for Sports Applications

R. Manikandan and R. Ramakrishnan

***Abstract*---**Segmenting out foreground object from its background is an interesting and important research problem in the video based applications. It has more importance in the field of computer vision due to its applications such as sports, security systems, video surveillance, etc. The background subtraction algorithms are used to analyze the player's activity in sports, to improve the performance of player by detecting the motion of the players in video sequences. The various algorithms like frame difference, approximate median, mixture of gaussian are compared and analyzed with real time sports videos. Mixture of Gaussian turns out to be best in reliability of extraction of moving objects, robust to noise, whereas the conventional algorithms result in noise and poor extraction of objects. The parametric analyzes of metric such as recall, precision, etc., gives the complete behavior of player.

***Keywords*---**Approximate Median, Frame difference, Mixture of Gaussian, Motion Detection

## I. Introduction

IN sports video analysis, nowadays people paying more attention to the development of high-speed digital cameras and video processing. It has found many applications in sports such as ball/player tracking, game highlight extraction, computer-assisted refereeing, etc. In all these applications fixed cameras are used with respect to static background, a common approach of background subtraction is used to extracting the foreground object (players) in a running sequence of images. The background subtraction is achieved by taking absolute difference between each incoming frame and a background model of the scene [1]-[3].

### *A. Challenges of the Background Subtraction*

Since the human body is non-rigid in nature and it has various un-predicted shapes, the extraction of players and their motion analysis is a challenging issue. In the video based object extraction, the fast and accurate background subtraction technique plays an important role. More time is needed for extracting the features as well as the computational complexity increases when the entire image is computed by computer vision techniques. So many practical factors [12] are influenced, while capturing the real time sports videos, such as illumination changes, camera noise, camera calibration, camera misfocus, reflections, occlusions, lighting conditions and shadows which affect the extraction of players from the video.

### *B. Motivation*

The sports dynamic analysis requires extracting the foreground object (players) through background subtraction, for analyzing the player's activity. Even though, the analysis of this process is difficult to obtain using video/image processing techniques, a wide range of research possibilities are present in the sports application. Marker based sports player analysis is a simple approach but it needs markers on the human body or some markers are used in the event field at every time. So, the marker-less human motion analysis is very important in the automated analysis.

The paper is organized as follows: methodology of background subtraction system can be found in Section II. The various algorithms of Background Subtraction techniques are discussed in section III. The performance analysis parameters, results and discussions are presented in sections IV and V, followed by conclusions and future work on section VI.

## II. Methodology of Background Subtraction System

To obtain background subtraction, the background has to model first. Then, the incoming frame is obtained, and subtract out from the background model. With the background model, a moving object can be detected. This algorithm is called as 'Background Subtraction'. The efficiency of a background subtraction technique correlates with three important steps: modelling, thresholding and data validation as shown in fig.1.

Background modeling [10], is the backbone of the Background Subtraction algorithm. Background model defines the type of model selected to represent the background, and the model representation can simply be a frame at time (t-1) formula such as the median model. Model Adaption is the procedure used for adjusting the background changes that may occur in a scene. Thresholding is a procedure that eliminates an unwanted range of pixels in the scene with respect to certain threshold values. Data validation is involved with the collection of techniques to reduce the misclassification of pixels.

Manuscript received September 21, 2013, review completed on October 05, 2013, and revised on October 11, 2013.

R. Manikandan is doing research at Tamilnadu Physical Education and Sports University, Chennai, India. Mobile: +91-9444007821. E-Mail : money_kandan2004@yahoo.co.in

R. Ramakrishnan is with Department of Advanced Sports Training and Technology, Tamilnadu Physical Education and Sports University, Chennai, India.

Digital Object Identifier: DIP092013007.

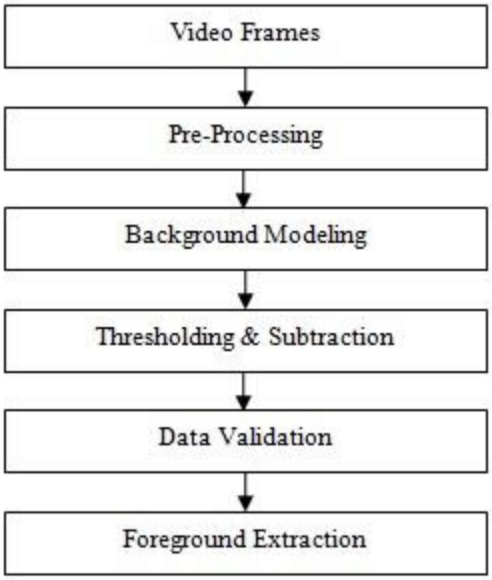



Fig.1 General Flow Diagram of BGS System

In the recent papers, many background subtraction algorithms are proposed, because no single algorithm is able to cope with all the challenges in the sports applications. There are several problems that a good background subtraction algorithm must resolve. Therefore in the next section most commonly used, three different types of background subtraction algorithms are discussed.

## III. Various Algorithms of Background Subtraction Techniques

### *A. Frame Difference*

Frame difference [4, 13] is the simplest form of background subtraction. The current frame $F_i$ is subtracted from the previous frame $F_{i-1}$, refer to "(1)", and if the difference in pixel values for a given pixel $D_i$ is greater than a threshold $T_s$, the pixel is considered as a part of the foreground $B_i$.

$$| \text{frame}_i - \text{frame}_{i-1} | > T_s \qquad (1)$$

*Algorithm*

1. Convert the incoming frame '$F_i$' to grayscale
2. Subtract the current frame from the background model '$F_{i-1}$' (in this case it's just the previous frame)
3. For each pixel, if the difference between the current frame '$F_i$' and background '$F_{i-1}$' is greater than a threshold $T_s$, the pixel is considered as part of the foreground '$B_i$'.

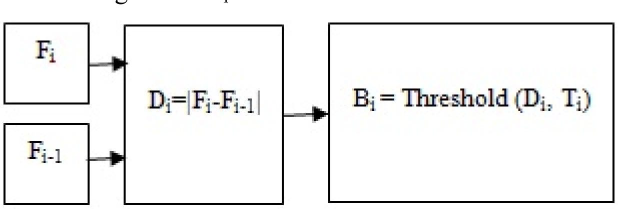



Fig. 2 Block Diagram of Frame Difference Method

A major disadvantage of this method is that for objects with uniformly distributed intensity values, the interior pixels are interpreted as part of the background. Another disadvantage is that objects must be continuously moving. If an object stays still for more than a frame period (1/frames per second (fps)), it becomes part of the background.

This method does have two major advantages. One advantage is the modest computational load. Another is that the background model is highly adaptive. Since the background is based solely on the previous frame, it can adapt to changes in the background faster than any other method (at 1/fps to be precise).

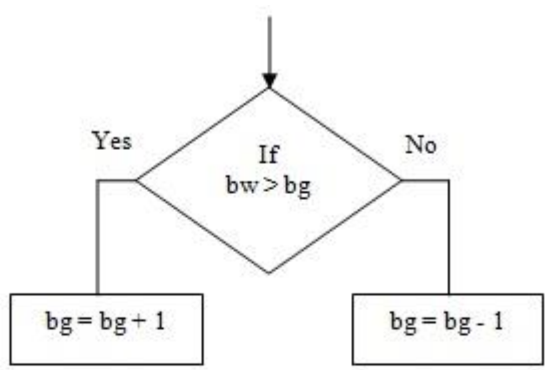
The frame difference method (FDM) subtracts out the background noise, much better than the more complex approximate median (APM) and mixture of Gaussians methods (MoG). A challenge with this method is determining the threshold value. The threshold is set too low, will get every object pass through it, and also it cannot be set too high because it will block the foreground.

### *B. Approximate Median*

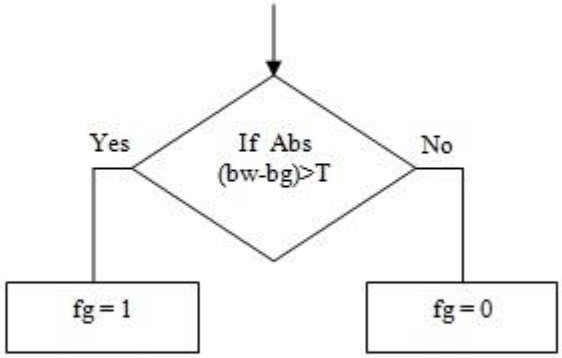
Approximate median [5, 14] method uses a recursive technique for estimating a background model. Each pixel in the background model is compared to the corresponding pixel in the current frame, to be incremented by one if the new pixel is larger than the background pixel or decremented by one if smaller. A pixel in the background model effectively converges to a value where half of the incoming pixels are larger than and half are smaller than its value. This value is known as the median. The Approximate median method has been selected, for it handles slow movements, which are often the case in our environment, better than the Frame differencing Method. The Approximate median foreground detection compares the current video frame to the background model, refers to "(2)", and identifies the foreground pixels. For this it checks if the current pixel bw(x, y) is significantly different from the modeled background pixel bg(x, y).

$$|bw(x, y) - bg(x, y)| > T \qquad (2)$$

A simplified pixel-wise implementation of the approximate median background subtraction method in flow chart and pseudocode is given in Fig.3 where, T represents the threshold value.




(a)




(b)

```
/* Adjust background model */
if (bw > bg) then bg = bg + 1;
else if (bw < bg) then bg = bg - 1;
/* Determine foreground */
if (abs (bw - bg) > T) then fg = 1;
else fg = 0;
```

(c)

Fig.3 Approximate Median Method (a) Flow Chart for Adjust background model (b) Flow Chart for Determine foreground and (c) Pseudocode

### C. *Mixture of Gaussian*

A Gaussian mixture model (GMM) was proposed for the background subtraction in Friedman and Russell, [6] and efficient update equations are given in Stauffer and Grimson, [7]. In Power and Schoonees, [8] the GMM is extended with a hysteresis threshold. This method uses a Gaussian probability density function to evaluate the pixel intensity value. It finds the difference of the current pixel's intensity value and cumulative average of the previous values. So it keeps a cumulative average (μ) of the recent pixel values. If the difference of the current image's pixel value and the cumulative pixel value is greater than the product of a constant value and standard deviation then it is classified as foreground [11]. That is, at each t frame time, the I pixel's value can then be classified as foreground pixel if the inequality: $|I_t - \mu_t| > k\sigma$ holds; otherwise, it can be considered as background, where k is a constant and σ is standard deviation.

Here background is updated as the running average:

$$\mu_{t+1} = \mu_t * I_t + (1-\alpha) * \mu_t \qquad (3)$$

$$\sigma^2_{t+1} = \alpha (I_t - \mu_t)^2 + (1-\alpha)\, \sigma^2\, t \qquad (4)$$

where,

σ, the learning rate, is typically 0.05,

$I_t$ is the pixels current value and

$\mu_t$ is the previous average.

The Table I show that the various tuning parameters of the proposed background subtraction algorithms. By means of changing every pixel in the image, the tuner has to evaluate their detection results.

TABLE I
DIFFERENT TUNING PARAMETERS USED IN ALGORITHMS

| Algorithms | Fixed Parameters | Test Parameters |
|---|---|---|
| FDM | None | Foreground threshold, Ts |
| APM | None | Foreground threshold, Ts |
| MoG | Number of components, K = 3<br>Initial variance, $\sigma^2_e = 36$<br>Initial weight, $\omega_e = 0.1$ | Adaptation rate, α<br>Weight threshold, T<br>Deviation threshold, D |

## IV. PERFORMANCE ANALYSIS PARAMETERS

In this section, we compare the performance of three popular background subtraction techniques [9, 15]. Our goal is to evaluate the ability of each method to correctly detect motion, a ground truth is available for the videos constituting the database allowing the evaluation of true positives (TP), false positives (FP), true negative (TN) and false negatives (FN) numbers. Definitions for the evaluation parameters as follows:

1. True Positive (TP): which represents the number of foreground pixels correctly detected by the algorithm.
2. False Positive (FP): is responsible for the number of pixels which are incorrectly classified as foreground objects.
3. True Negative (TN): indicating the number of background pixels which are correctly detected as background scene by the algorithm.
4. False Negative (FN): stands for the number of pixels corresponding to foreground objects which are misclassified as part of background image (also referred as misses).

Those values are combined into a (Precision/Recall) couple defined as:

### A. *Recall*

Recall is measure of completeness and is defined as number of true positives divided by the total number of elements that actually belong to the foreground objects. (i.e. some of both true positives and false negatives).

$$Recall = \frac{TP}{TP + FN}$$

In other words it can be rewritten as:

$$Recall = \frac{\textbf{number of correctly foreground pixels}}{\textbf{number of foreground pixels in ground truth}}$$

### B. *Precision*

Precision can be considered as a measure of exactness or fidelity and is evaluated through dividing the number of items (foreground objects) correctly detected by the total number of pixels classified as foreground by algorithm. In fact we are evaluating if the algorithm shows that a certain pixel is foreground and how reliable that statement would be.

$$Precision = \frac{TP}{TP + FP}$$

In other words it can be rewritten as:

$$Precision = \frac{nmber\ of\ correctly\ identified\ foreground\ pixels}{number\ of\ foreground\ pixels\ detected\ by\ algorithm}$$

Recall and precision values are both within the range of 0 and 1. When applied to the entire sequence, the recall and precision reported are averages over all the measured frames. Typically, there is a trade-off between recall and precision (recall usually increases with the number of foreground pixels detected, which in turn may lead to a decrease in precision). A good background algorithm is one producing simultaneously a small number of false positives and false negatives, i.e. both a high Precision and Recall value.

## V. Results and Discussion

The proposed work has been developed using MATLAB 7.10(R2010a) on Intel dual core processor, 2GB RAM and Windows XP SP2. The real time video sequences are acquired at the rate of 30 frames/second with the frame size of 640×360 pixels resolution. Figure 4 and 5 shows the simulation results of proposed work for frame 41 and frame 71.

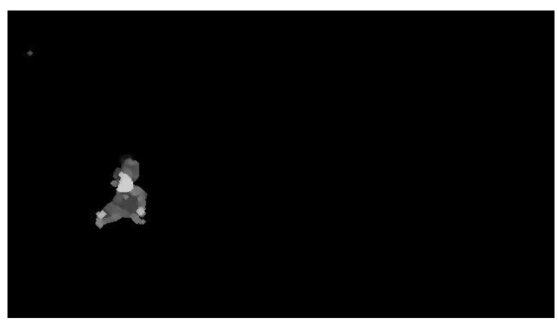
(a)

(b)

(c)

(d)

(e)



(f)

Frame No: 41

Fig.4 Foreground objects detected by using different algorithms (a) Original Image (b) Grayscale Image (c) Background Model (d) Frame Difference algorithm for threshold $T_s = 60$ (e) Approximate Median algorithm for threshold $T_s = 60$ (f) Mixture of Gaussians algorithm for threshold 0.25, learning rate $\alpha = 0.01$, and Positive deviation threshold, D = 2.5.

(a)

(b)

(c)

(d)

(e)

(f)

Frame No: 71

Fig.5 Foreground objects detected by using different algorithms (a) Original Image (b) Grayscale Image (c) Background Model (d) Frame Difference algorithm for threshold $T_s = 60$ (e) Approximate Median algorithm for threshold $T_s = 60$ (f) Mixture of Gaussians algorithm for threshold T = 0.25, learning rate $\alpha = 0.01$, and Positive deviation threshold, D = 2.5.

The Table II shows that the Recall and Precision results for three background estimation algorithms. Based on the measurements shown in Fig 6 and visual examination on the resulting foreground masks, we make the following observations regarding the background algorithms are tested in this paper: With the appropriate parameters, MoG achieves the best precision and recall. FDM is significantly worse than the AMF and MoG schemes. Even though AMF is not as good as MoG, it produces good performance with an extremely simple implementation. Since the amount of background update (+1 or -1) is independent of the foreground pixels, it is very robust against moving object. The only drawback is that it adapts slowly toward a large change in background. Mixture of Gaussian algorithm gives very well at separating out objects and suppressing background noise.

TABLE II
RECALL AND PRECISION RESULTS FOR THREE BACKGROUND ESTIMATION ALGORITHMS

| Recall | Precision | | |
|---|---|---|---|
| | FDM | APM | MoG |
| 0.1 | 0.91 | 0.92 | 0.97 |
| 0.2 | 0.85 | 0.93 | 0.97 |
| 0.3 | 0.83 | 0.93 | 0.97 |
| 0.4 | 0.8 | 0.9 | 0.97 |
| 0.5 | 0.76 | 0.92 | 0.97 |
| 0.6 | 0.72 | 0.88 | 0.97 |
| 0.7 | 0.67 | 0.87 | 0.97 |
| 0.8 | 0.63 | 0.84 | 0.97 |
| 0.9 | 0.59 | 0.78 | 0.87 |
| 1 | 0.15 | 0.23 | 0.31 |

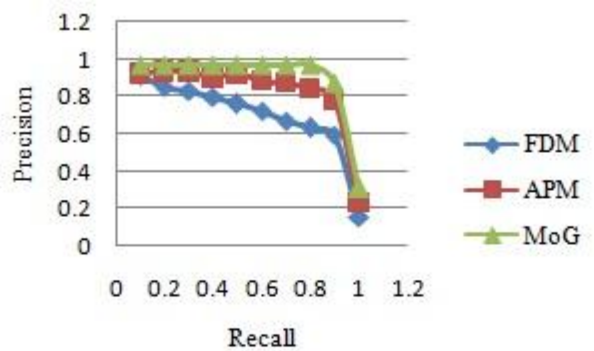



Fig.6 Precision-recall plots for (a) Frame Difference (FDM) (b) Approximate Median Filter (APM) and (c) Mixture of Gaussians (MOG)

## VI. Conclusions and Future Work

In this paper, we are discussed three different types of background subtraction algorithms for to analyze the player's activity in sports, to improve the performance of player by detecting the motion of the players in video sequences. We analyze them based on the performance parameters such as recall and precision. The following algorithms are tested on real time sports video sequences: frame differencing, approximate median filtering, and mixture of Gaussians. Mixture of Gaussians produces the best results, while approximate median filtering offers a simple alternative with competitive performance. More research, however, is needed to improve robustness against environment noise, sudden change of illumination, and to provide a balance between fast adaptation and robust modeling.

## Acknowledgement

We thank our Secretary and Correspondent Dr.P.Chinnadurai, Panimalar Engineering College, Chennai, India for providing financial assistance and facilities for doing this research.

## References

[1] M.Piccardi, "Background subtraction techniques: a review," in Proc. IEEE Int. Conf. Systems, Man, Cybernetics, pp 3099 – 3104, 2004.

[2] A.McIvor, "Background subtraction techniques," in Proceedings of Image and Vision Computing , Auckland, New Zealand, 2000.

[3] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting moving objects, ghosts and shadows in video streams," IEEE Transactions on Pattern Analysis and Machine Inteligence , vol. 25, no. 10, pp. 1337–1342, 2005.

[4] Zhan Chaohui Duan Xiaohui, Xu Shuoyu Song Zheng Luo Min, "An improved Moving Object Detection Algorithm Based on Frame Difference and Edge Detection", IEEE Fourth International Conference on Image and Graphics , 2007.

[5] Sen-Ching S. Cheung, Chandrika Kamath, "Background Subtraction with Foreground Validation for Urban Trafic Video", EURASIP Journal on Applied Signal Processing, 14: 2330-2340, 2005.

[6] Friedman N., Russell, S, "Image segmentation in video sequences: a probabilistic approach", In: Proc. 13th Conf. on Uncertainty in Artificial Intelligence, 1997.

[7] C.Stauffer, E.Grimson, "Adaptive background mixture models for real-time tracking", IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2:246-252, 1999.

[8] Power, P.W., Schoonees, J.A., "Understanding background mixture models for foreground segmentation", In: Proc. of the Image and Vision Computing New Zealand, 2002.

[9] Y. Benezeth, P.-M. Jodoin, B. Emile, H. Laurent, C. Rosenberger "Comparative study of background subtraction algorithms" Journal of Electronic Imaging, SPIE, vol. 19, 2010.

[10] M. Hedayati, Wan Mimi Diyana Wan Zaki , Aini Hussain, "A Qualitative and Quantitative Comparison of Real-time Background Subtraction Algorithms for Video Surveillance Applications" Journal of Computational Information Systems ", pp 493 – 505, 2012.

[11] Z.Zivkovic,"Improved adaptive Gausian mixture model for background subtraction", IEEE International Conference on Pattern Recognition (ICPR), pp 28-31, 2004.

[12] Sebastian Brutzer, Benjamin Hoferlin, Gunther Heidemann," Evaluation of background subtraction techniques for video surveillance", CVPR, *pp 1937-1944*, IEEE, *2011*.

[13] Xiong Weihua , Xiang Lei, Li Junfeng , Zhao Xinlong,"Moving object detection algorithm based on background subtraction and frame differencing", IEEE Control Conference (CCC), pp 3273–3276, 2011.

[14] Ahmad, K.A, Saad. Z, Abdullah. N, Hussain.Z,"Moving vehicle segmentation in a dynamic background using self-adaptive kalman background method", IEEE 7th International Colloquium on Signal Processing and its Applications (CSPA), pp 439-442, 2011.

[15] White. B, Shah, M, "Automatically tuning background subtraction parameters using particle swarm optimization", IEEE International Conference on Multimedia and Expo, pp 1826-1829, 2007.

**R. Manikandan** received B.E degree in Electronics and Instrumentation Engineering from Annamalai University, Chidambaram in the Year 2002. He obtained his M.E degree in Applied Electronics from Anna University, Chennai. Currently, he is pursuing PhD in the Department of Advanced Sports Training and Technology at Tamilnadu Physical Education and Sports University, Chennai. His research areas are Image/Video Processing based measurements and Sensors & Transducers. He is working as Associate Professor in the Department of Electronics and Instrumentation Engineering at Panimalar Engineering College, Chennai. He has published research papers in national and international journals and conferences.

**Dr.R. Ramakrishnan** received AMIE degree in Mechanical Engineering from the Institution of Engineers (India), Calcutta. He obtained his M.E degree in Mechanical Engineering from Bharathiar University and PhD degree from Anna University, Chennai. He is working as Professor and Head in the Department of Advanced Sports Training and Technology, Tamilnadu Physical Education and Sports University, Chennai. He worked as Principal, Panimalar Engineering College, Chennai, for two years. His areas of interest are metal cutting, micro machining, modeling and optimization and nano materials. Currently he is guiding six PhD scholars. He has published more than 40 papers in various national and international journals and conferences.