Feature Descriptor Based Identity Retention and Tracking of Players Under Intense Occlusion in Soccer Videos

Dippal Israni[1]*   **Hiren Mewada**[2]

[1]*U and P U Patel Department of Computer Engineering,*
Charotar University of Science and Technology, India
[2]*V T. Patel Department of Electronics and Communication Engineering,*
Charotar University of Science and Technology, India
* Corresponding author's Email: dippalisrani.ce@charusat.ac.in

Abstract: Tracking plays an important role to analyze the action of players, i.e movement of players and path corresponding to this movement. In recent years, the number of approaches to detecting players throughout the monocular image sequences has grown steadily. Severe occlusion makes this tracking challenging and standard algorithm fails to retain the identity of players. In soccer sequences, it is even more challenging due to erratic movement of players as well as players having clothes of the same color. Feature Descriptors do not require prior training to hold player's size, shape or color. Thus, this paper proposes feature descriptor based object tracking algorithm to tackle severe spatially extended and temporally short and long-term occlusions. The objective is to detect players and maintain their identities over the time even under the situation of intense occlusion. Gaussian Mixture Model (GMM) and Hungarian Assignment (HA) are used for player detection and association respectively. In the proposed approach, the feature descriptors are pre-calculated for the players having a reliable blob. In the case of demerge, these reliable feature descriptors are used to retain player identities. The proposed approach is evaluated on ISSIA Soccer dataset having six sequences having 25 frames per second (fps) for the image size of 1080 x 1920. The proposed method outperforms state of the art trackers by achieving 80% accuracy and 64% precision on ISSIA dataset.

Keywords: Occlusion and demerging, Player detection and tracking, Soccer video, GMM, HA.

1. Introduction

In the field of computer vision, human action recognition is one of the major challenges. It has various applications like sports video analysis [1, 2], surveillance scenarios [2 - 4], robots navigation [5, 6], human-robot interaction [7, 8]. In sport sequences, players are tracked to analyze player's strength, team strength and mining their movements to provide required guidelines to other players.

Tracking of players and their movements is essential to understand the tactic of the game along with automatically mining the player's movement, speed, strength and ability to instruct other players in specific situations. Significant work has been carried out to track players. For example, people can be tracked by searching in the neighborhood from last frame as proposed by S. Denman et al. [9], using model matching proposed by M. Andriluka et al. [10], using joint probabilistic methods proposed by D. Schulz [11, 12].

Occlusion limits the continuity of player to be tracked. The occlusion defines the effect of one player blocking another player in the 3D Space from captured view. This occlusion can be classified as self-occlusion, inter-player occlusion and occlusion by background. Self-occlusion mostly arises while tracking articulated players (when one part of the player's body occludes another part of the same player). Inter-player occlusion occurs when two players being tracked, occlude each other. Occlusion by the background occurs when a structure in the background occludes the tracked players. In sports,

analysis of game plan is obstructed due to the challenges like occlusion. Initial task in this paper is to detect player and assign an identity to the player. This paper proposes identification and retention of player's identity under intense occlusion using various statistical feature sets. The accuracy of the proposed model is calculated by generating the ground truth sequences using semi-automatic model provided by T. D'Orazio [13]. This dataset has six soccer videos in different scenarios which cover the whole ground.

In this paper an efficient tracker is proposed in which GMM based object detection is used. Hungarian cost assignment is used for association. This paper proposes a novel technique for detection of occlusion and for handling occluded players. This Proposed solution also uses feature descriptors to retain identity of players once demerge occurs.

This paper is further organized as follows: The next section gives some brief background of researches related to the proposed technique. Section 3 describes the proposed methodology for object detection and tracking. The experimental results and discussions of the proposed approach are depicted in Section 4. Finally, conclusions are summarized in Section 5.

2.   Related works

Presently, various models are proposed targeting surveillance video with objectives of detecting and representing people and tracking these peoples. Categorization of occlusion as non-occlusion, partial occlusion, full occlusion and long-term occlusion is proposed by L.Yong et al. [14]. Object permanence method for 5 months old infants is proposed by R. Baillargeon et al. [15] which suggested that an occluded object will re-emerge near its occluder. The same concept can be extended for surveillance and sports to retain identity under occlusion scenarios. Part matching tracker (PMT) is used to match local parts from multiple frames jointly by considering their low-rank and sparse structure information [16]. Authors applied the algorithm to surveillance video and tested the results under partial occlusion. However nonlinear and fast movement of the player in soccer video makes it challenging to detect intensive occlusion. Kanade Lucas Tomasi (KLT) tracker was proposed by B. Han et al. [17] for tracking features of the moving objects. This approach works fine for occlusion detection but identity retention is unresolved. P. Bilinski et al. used Features from Accelerated Segment Test (FAST) algorithm for computing feature points and build a Histogram of Gradient

(HOG) based descriptor for each feature point [18]. This retains the identity of the person in occlusion scenarios at the cost of high complexity raised by FAST and HOG algorithms. A new approach for continuously analyzing the occlusion situation by exploiting the spatiotemporal context information is proposed by J. Pan [19]. However, this approach does not track the erratic movements of the players which often occur in sports video like a soccer game.

The particle filter, also known as the sequential Monte Carlo method is the most popular algorithm that generates the posterior probability density function using the propagation rule of state density. A technique to extract feature points of each target object and tracking features using particle filter based approach is proposed by M M. Naushad Ali et al. [20]. This approach gives better results for tracking a person but player retention is not achieved in it. To maintain object's identity Tran et al. [21] have proposed a technique using color and Bhattacharya distance algorithm. In soccer videos, players wear t-shirts of the same color. So color based approach cannot distinct demerged players.

Yang et al. [22] proposed a technique with the ability to solve long-term occlusion without using any prior knowledge about shape and motion of objects. This algorithm can retain the identities after occlusion if apriori information about the structure is stored. A customized genetic algorithm for optimal region tracks is proposed by Huang et al. [23] to detect occlusion using object permanence. However, identity retention is not classified in this paper.

In sports, especially in soccer games analysis, a trajectory estimation of player movement is essential and in that occlusion is an obstruction. There is a need of a novel work to handle occlusion and retain player id once the players are demerged (deoccluded). To track players and ball in indoor soccer videos S. Bidhendi et al. [24] proposed ground field detection method without identity retention. Pose estimation of the player using contourlet extraction and the neural network is proposed by T. D'Orazio [25] et al. This achieves an identity of player at the cost of computational complexity. G. Thomson [26] used marking on the pitch such as arcs and lines to compute the camera pose and compute the position, orientation and focal length of a camera in a real time using image analysis to track player's relative movement.

There are two types of occlusion, short-term occlusion where players collide for a short period of time and long-term occlusion where players collide for a long period of time. To tackle spatially
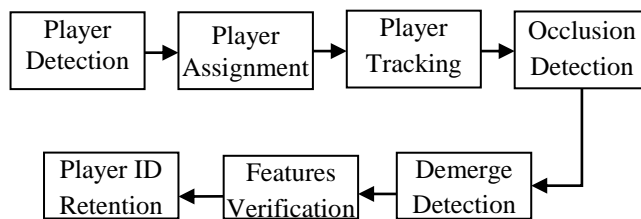
Player Detection → Player Assignment → Player Tracking → Occlusion Detection

Player ID Retention ← Features Verification ← Demerge Detection ← (Occlusion Detection)

Figure. 1 Block Diagram of the proposed framework for tracking and identity retention

Video Frame Sequence → Read Image → Blob Detection → Morphological Operations

Compute Cost and Distance ← [image] ← Identifying Centroids and bounding box
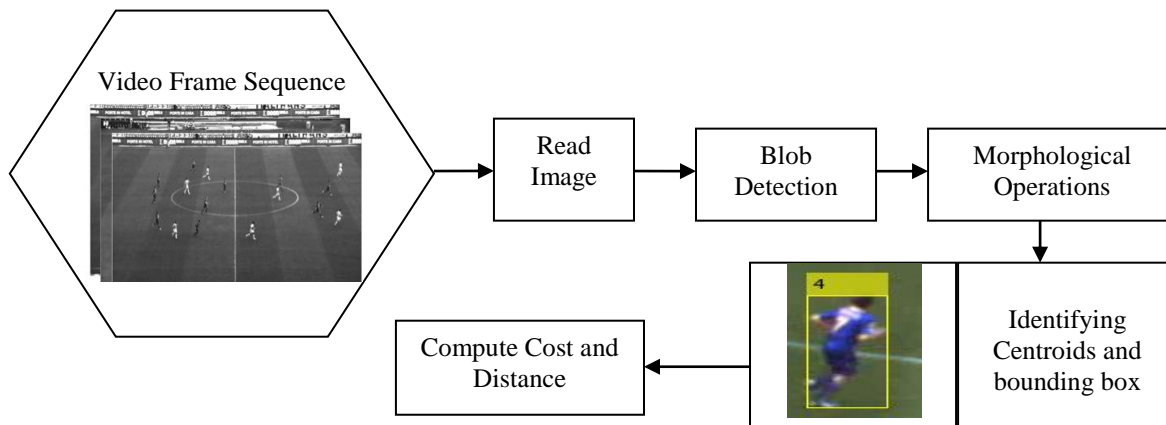
Figure. 2 Player detection process

extended and temporally short and long-term occlusions, statistical feature based player tracking algorithm is advantageous as less information is needed and there is no requirement to store each frame. In this context, S. Pertuz et al. [27] described the computation of the focus level for every pixel of an image by means of various focus measure operators. The focus measures were divided among six various families known as Gradient-based operator (GRA), Laplacian-based operator (LAP), Wavelet-based operators (WAV), Statistics-based operators (STA), DCT-based operators (DCT) and Miscellaneous operators (MIS). This paper has concluded 36 kinds of focus measures depending upon different characteristics. These features do not require any prior training. Hence proposed method uses few families from listed to tackle occlusion region and to retain the player identity in the occluded region. The features are selected based on their rotation scale and translation (RST) invariance and divergence. To identify multiple players, Gaussian Mixture Model (GMM) based background and foreground extraction method [28] is used in the proposed model. Kalman filter can be used for efficient tracking as proposed by Welch et al. [29]. However, it fails to retain the player's id when players are completely occluded. A distributed approach using particle filter and clustering is proposed by F. Previtali et al. [30]. This algorithm has higher efficiency in contrast to other algorithms. However it uses views of multiple cameras to achieve accuracy. Hungarian assignment based on bipartite graph method is proposed by [31]. This method is used in proposed approach for association. In all existing approach, very less amount of work is done for minimizing identity switches and with higher accuracy and precision. Hence the proposed approach concentrates on minimizing the identity switches and maximizing accuracy and precision.

3. Proposed approach

Most of the existing techniques are able to identify occlusion but failed to retain correct player identity due to long and intense occlusion. In case of soccer games, players occlude too often and player's clothes are of same color and pattern. Also, velocity and acceleration of player's movement vary nonlinearly. In this type of situations, occlusion retention is too intense. In the proposed approach a novel technique is proposed which retains id of players even under long and intensive occlusion using feature-based algorithms [27]. The block diagram of proposed approach consisting 7 steps is shown in Fig. 1 and explained in the following section.

3.1 Player detection

Detection of player is a pre-requisite step for any tracking applications. Player detection method is described in Fig. 2. In this method, Gaussian

Frame	Blob Detection using GMM

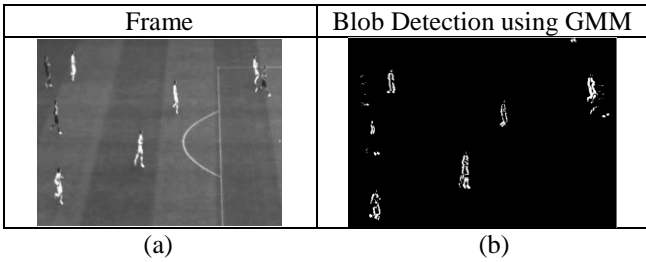(a)   (b)

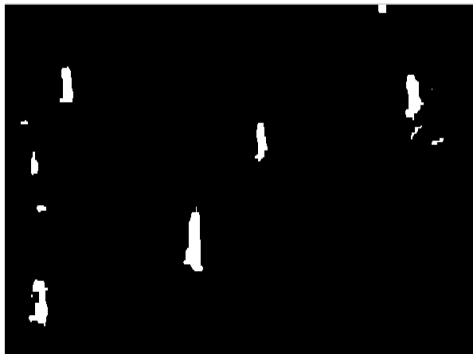Figure. 3 Blob detection: (a) Soccer frame and (b) GMM based player detection

Figure. 4 Player detection after morphological processing

Mixture model is initialized to detect foreground objects [28]. To avoid undesirable noise/blob, morphological operations are applied. Next bounding box of each connected pixel is calculated. Figure 2 shows player detection model.

GMM is used to detect player (foreground blob) [28]. GMM is a parametric probability density function represented as a weighted sum of Gaussian component densities.

It is computed as,

$$P(X_t) = \sum_{i=1}^{k} \omega_{i,t}\, \eta\big(X_t; \mu_{i,t}, \Sigma_{i,t}\big) \qquad (1)$$

Where X is a D-dimensional continuous-valued data vector ω_i, k is a number of Gaussian components, $i = 1 \ldots k$, are the mixture weights, and $\eta\,(X_t;\,\mu i,\,\sum i\,,\,t)$, $i=1\ldots.k$, are the components Gaussian densities. GMM processed frame and detected blobs (i.e. players) are presented in Fig. 3.

Morphological operations including opening and closing are applied to fill gap within blobs. Corresponding results is shown in Fig. 4. After performing morphological operations centroids are detected for each blob and cost function is computed using the Euclidian distance between each blob and all other blobs using Eq. (4).
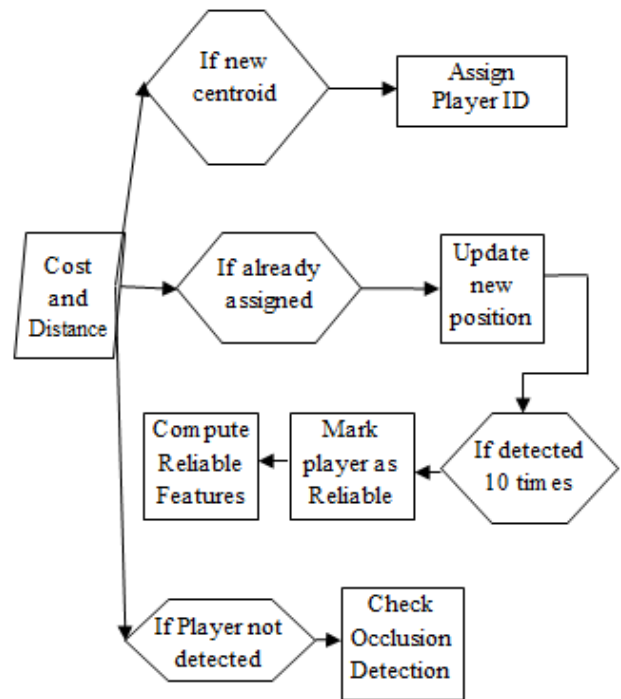

Figure. 5 Player assignment process

$$d = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2 + (x_j - y_j)^2} \qquad (4)$$

Where x_i and y_i are points and d is the distance computed between blobs, n is a total number of players. This cost function is used for assigning an identity to the player.

3.2 Player Id assignment

Once a player is detected there are chances that few players are new in frames, the same player is available from previous frame or player is invisible due to occlusion. Therefore once the cost function is computed, player id assignment is done in three categories. The flow diagram of the player assignment is as mentioned in Fig. 5.

1) First time appearance of the player

If the player is present for the very first time assign a new unique player identity to a player detected. Age is used to maintain total occurrence of the player. Invisible count determines the number of frames from when the player got invisible.

2) The player is already present

In each frame, the location of detected player is marked as the bounding box. If the player is already present update its player position, increment his age.
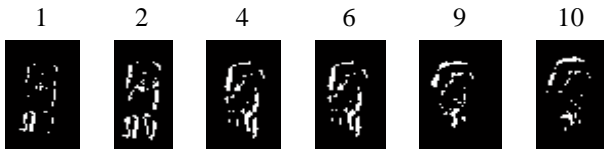
Figure. 6 Player's visibility in every frame

For the first few times, when the player is detected, it is not considered as completely visible. This is because video frame rate is 25 frames / second. The player is not considered reliable as the player has done an insufficient motion to be considered as a full blob. This can lead to many identities for a single player which is explained below in Fig. 6.

If a player is visible for ten consecutive frames, a blob of player is entirely set. Mark that player as a reliable player and reserve its properties such as bounding box, player identity. Also, five feature descriptors of player are reserved which will be useful in future to retain player identity after demerging.

The feature descriptors are used to retain an identity as there are numerous advantages for it. They are Rotation Scaling and Translation invariant, no apriori training required, requires very few computation and no high storage of features required. For every reliable player, five feature descriptors will be computed which are as follows. All five features are distinct as they are from different families proposed by Perutz et al. [27]. In Eqs. (5) and (7) * operator represents convolution operation.

Diagonal Laplacian:

In order to compute modified Laplacian of image, Thelen et al. [32] proposed vertical variations of the image.

$$LAPD = |I * L_x| + |I * L_y| + |I * L_{x1}| + |I * L_{x2}| \tag{5}$$

Where, $L_x = [-1 \ 2 - 1]$, $L_y = L_x^T$,

$$L_{X1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{bmatrix}, L_{X2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Standard deviation of player:

Standard deviation is a measure of variation or dispersion of a set of values which can be computed using Eq. (6).

$$S = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)} \tag{6}$$

Where, x_i = intensities of image and μ = mean of intensities.

Image curvature:

This was proposed for microscopy applications [33]. If the image grey levels are interpolated as a surface, the curvature of the surface may be used as a focus measure [33, 34].

$$\emptyset = |c_0| + |c_1| + |c_2| + |c_3| \tag{7}$$

$Where, c_0 = M_1 * I$

$$c_1 = M_1^T * I$$

$$c_2 = \frac{3}{2} * M_2 * I - M_2^T * I$$

$$c_3 = \frac{3}{2}(M_2^T * I) - (M_2 * Irevx)$$

$$Where, M_1 = \frac{1}{6}\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, M_2 = \frac{1}{5}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Gradient energy:

Gradient Energy is proposed as a focus measure which is the sum of squares of the first derivative in the x and y direction. [35 - 37].

$$\emptyset_{X,Y} = \sum_{(i,j)\in\Omega(x,y)}(I_x(i,j)^2 + I_y(i,j)^2) \tag{8}$$

Where, I_x and I_y are first order derivatives in x and y direction respectively.

Vollath's autocorrelation:

A focus measure based on autocorrelation of the image was proposed in [38, 39, 40]. It can be computed using Eq. (9).

$$\emptyset_{x,y} = \sum_{(i,j)=\Omega(x,y)}\big(I(i,j) \times I(i+1,j)\big) - \sum_{(i,j)=\Omega(x,y)}\big(I(i,j) \times I(i+2,j)\big) \tag{9}$$

Hence all these features are computed for every player who gets reliable using Eqs. (6), (7), (8), and (9).
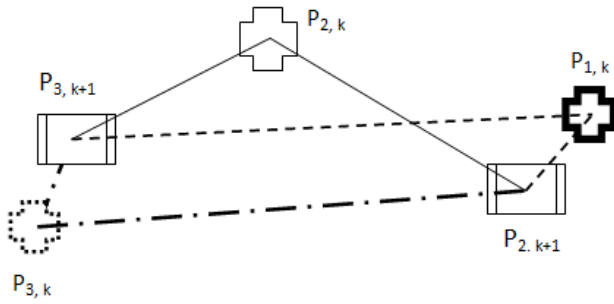
3) The player is not present

Figure. 7 Player positions update using Hungarian algorithm

Figure. 8 Occlusion detection process

If the player is not present there are two cases possible.
i)   No movement of player in particular frame
ii) Player is occluded by some other player

3.3 Player tracking

Player Tracking Phase tracks all the players in each frame as they change their positions continuously. To continuously update their positions, optimal mapping is done for all observed players using Hungarian assignment method. The Hungarian algorithm is the method to do linking process between identified objects in frame k to the unidentified objects in frame k+1 by finding the extreme solution of the Euclidean distance in the assignment matrices. The process of position update is as mentioned in Fig. 7.

$P_{i,k}$ represents an identified object in the previous frame and $P_{i,k+1}$ represents detected an object in the current frame. From the case in Figure 7, there will be generated assignment matrices illustrated below.

$$D_{k,k+1} = \begin{bmatrix} d_{1k,1k+1} & d_{1k,2k+1} & d_{1k,3k+1} \\ d_{2k,1k+1} & d_{2k,2k+1} & d_{2k,3k+1} \\ d_{3k,1k+1} & d_{3k,2k+1} & d_{3k,3k+1} \end{bmatrix} \quad (10)$$

Where, $d_{ik,jk+1}$= Euclidian distance $(P_{i,k}, P_{j,k+1})$
Hungarian method finds the mapping for each object by finding the minimum point solution of the assignment matrices. Finally the assignment is done using

$$f = Min(\{D_{k,k+1}\})$$
$$\{f: \{(P_{1,k+1} \to P_{1,k}), (P_{2,k+1} \to P_{3,k})\} | P_{i,k}, P_{i,k+1} \in R^2\} \quad (11)$$

Thus while tracking, details of players like a number of times player detected, total visible counts, consecutive invisible count, bounding box are updated.
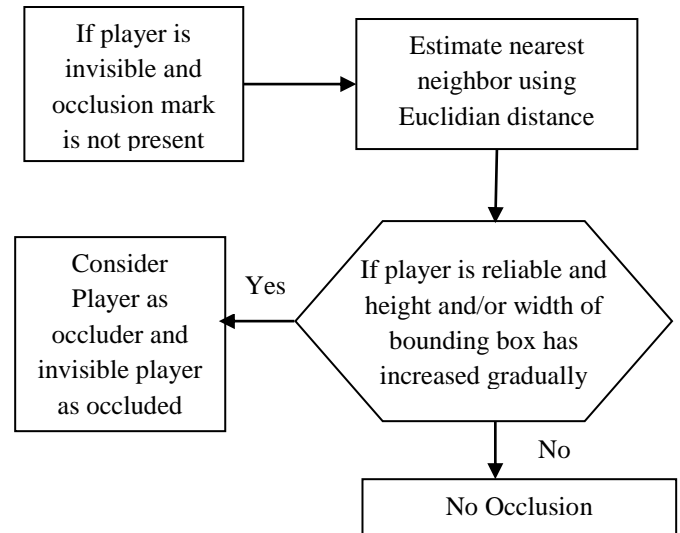
3.4 Occlusion detection

In this phase, it will be verified whether a player is occluded or it is invisible due to lack of motion in the current frame. If the occlusion is present a flag of occlusion mark will be on. The flow of verification is mentioned in Fig. 8 below.

If an object is occluded it will be occluded by its nearest object which is stated as object permanence [15]. Nearest neighbor's bounding box (height/width/both) is estimated in proposed method and compared with the reliable bounding box. If nearest neighbor's bounding box (height/width/both) is increased by threshold or more than the nearest player (front player) is marked as occluder and invisible player (Backside Occluded player) is marked as occluded. If the condition is not satisfied it will be verified in next frame and considered as no occlusion. The detected occlusion can be classified as self-occlusion (Sb) and full occlusion (f). This classification can be performed using Eq. (12) and Eq. (13) as follow.

$$S_b(x, y) = \begin{cases} \sqrt{\sum \left(A_{c(j)}(x,y) - A_{c(i)}(x,y) \right)}, PV = 0 \ and \\ \qquad\qquad OM = 0 \\ \quad - \quad , \qquad\qquad\qquad Otherwise \end{cases} \quad (12)$$

$$x = \frac{S_b(x)}{Sf_b(x)}, y = \frac{S_b(y)}{Sf_b(y)}$$

$$f(x) = \begin{cases} Mark \ Occluder \ and \ Occluded, \\ \qquad\qquad x||y > threshold \\ No \ Occlusion, \qquad\qquad otherwise \end{cases} \quad (13)$$
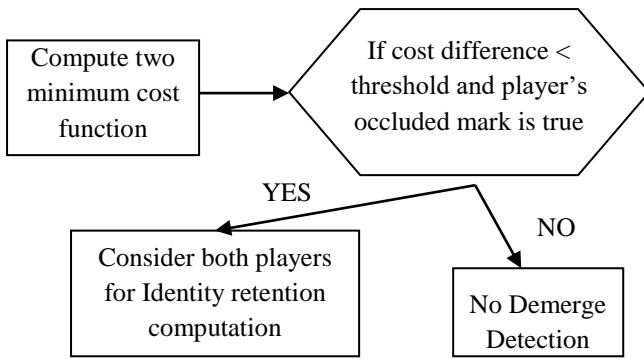
Figure. 9 Player demerge detection

Where PV = Player's visibility in current frame, $A_c(j)(x,y)$ = Centroids of Occluded Object, $A_c(i)(x,y)$ = Centroids of all detected players in current frame. $S_b(x,y)$= Horizontal and vertical position of centroids of current frame's probable occluder respectively, $Sf_b(x)$, $Sf_b(y)$ = Horizontal and vertical position of centroids of reliable frame, $f(x)$ = marking Occluder and Occluded Objects.

Object Permanence states that if a player is demerged it will be from its adjacent player only. Similarly, whenever a player is occluded it will be occluded by its adjacent player only. Therefore as mentioned in Eq. (12) whenever a player is not visible in the current frame, it's most adjacent player is verified for occluder. The most adjacent player's height/width/both are compared between the current frame and the frame when a player got reliable. If the computed values are greater than the threshold, they are marked as an occluder.

3.5 Demerge detection

Demerge detection is an important task to handle occlusion and to retain ID of players. The steps of demerge function of the proposed approach are mentioned in Fig. 9.

$$x = [A] - [B] < \mu$$
$$f(x) = \begin{cases} Demerge\ Detected, & x = 1\ and\ OM = \\ No\ Demerge\ , & otherwise \end{cases}$$

$$(14)$$

Where $[A]$ = Hungarian cost function of object1, $[B]$ = Hungarian cost function of object2, μ = threshold = 60, OM = Occlusion mark of player.

As presented in Figure 9 if a player is not occluded then it can be verified by computing two minimum cost function and if cost function of two tracks is less than a threshold and it's occluded mark is greater than 0, it states particular player is no more occluded from the current frame, it is

demerged which is as mentioned in Eqs. (13) and (14).

3.6 Features verification and player identity retention

To retain player identity occluder and occluded players features are required to be computed. These feature values of the occluder and occluded players are compared with occlude and occluded features of reliable player and the player which is matched more are retained their player identity accordingly. The flow of feature verification and player identity retention presented in Fig. 10.

$$sum1 = \sum_{i=1}^{4}[f_{d1}(i)-f_{r1}(i)]$$
$$+ \sum_{i=1}^{4}[f_{d2}(i)-f_{r2}(i)]$$
$$sum2 = \sum_{i=1}^{4}[f_{d1}(i)-f_{r2}(i)] \qquad (15)$$
$$+ \sum_{i=1}^{4}[f_{d2}(i)-f_{r1}(i)]$$

$$Identity\ retention$$
$$= \begin{cases} f_{d1} = P_1, f_{d2} = P_2\,, & sum1 < sum2 \\ f_{d1} = P_2, f_{d2} = P_1\,, & otherwise \end{cases}$$

Where $f_{d1}(i)$=Player 1's feature set in a detected frame, $f_{r1}(i)$=Player 1's feature set in a reliable frame, $f_{d2}(i)$=Player 2's feature set in a detected frame, $f_{r2}(i)$=Player 2's feature set in a reliable frame. Players are assigned same. P1 and P2 are player's identity in current frame.
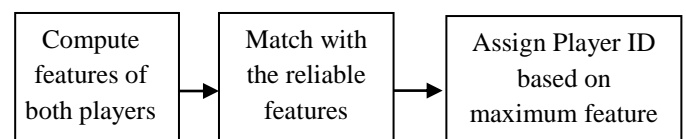

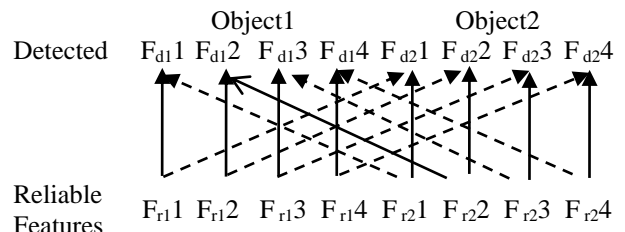Figure. 10 Feature verification and player ID retention

Figure. 11 Feature comparison with reliable player

To retain identity, demerged players positions are now compared with their own saved feature sets

and with saved feature sets of another object. If the difference is less for player's position and its feature sets then the same identity is retained otherwise identities will be swapped.

4. Analysis and comparison

To evaluate the performance of the proposed framework, experiments are carried out on a public dataset [13]. This dataset consists of six videos under challenging conditions like illumination variation, and occlusions. The video carries frame rate of 25 frames per second and resolution of 1080 × 1920.

Three videos are selected for processing (filmrole2, filmrole3, fimrole4). 10 seconds (~250 frames) are extracted from each selected video. The videos are selected based on ground truth observation of a higher number of occlusion activities.

Accuracy is measured in terms of precision and recall which can be computed using below-mentioned Eqs. (16) and (17).

$$Precision = \frac{\begin{array}{c}(PlayerIdretainedcorrect) \cap \\ (totalplayerocclusion)\end{array}}{(totalplayerOcclusion)} \quad (16)$$

$$\begin{array}{c}Recall \\ = \frac{(PlayerIdretained) \cap (totalplayerocclusio}{(playerIDretainedcorrect)}\end{array} \quad (17)$$

4.1 Occlusion retention for players of the same team

Occlusion in sports can be between players of same team or players of the different team. Intensity similarity in players of the same team forms a challenge in identity retention. In Table 2 identity retention of players having occlusion with players of the same team are computed on standard dataset videos. This retention and mismatch are compared with occlusions on ground truth of standard dataset in Table 2.

4.2 Occlusion retention of two players for different team / wearing different color clothes

Occlusion in sports can also be between players of the different team. In Table 3 identity retention of players having occlusion with different team's

player are computed on standard dataset videos. This retention and mismatch are compared with occlusions on ground truth of standard dataset in Table 3.

4.1. Comparison between Kalman based tracking and statistical features based tracking algorithm

This section evaluates and compares the result with standard motion based Kalman filter [29]. The algorithm is compared for occlusion between players of the same and different team. Occlusion and demerge scenario between players of a different team is described in Table 4. Based on the results mentioned in Table 4, in Kalman based approach as it does prediction, it results in incorrect identity retention for a different team player. In contrast proposed approach feature-based approach leads to correct retention. This is due to Kalman based tracking estimates identity based on the linear path in erratic scenarios. In contrast to that feature based proposed approach uses different feature descriptors and retains the identity of players throughout the sequence.

Occlusion and demerge scenario between players of the same team is described in Table 5. This scenario is challenging as in this case both demerged player have same color t-shirt. The feature descriptor based proposed approach has better retention results in this case in comparison with traditional Kalman filter.

Based on the results mentioned in Table 5, in Kalman based approach as it does prediction, it results in incorrect identity retention for same team player whereas feature-based approach leads to correct retention.

Table 2: Identity retention of players in same team

Sequences	Number of occlusions (Ground Truth)	Retention	Mismatch	Precision	Recall
S1	0	0	0	1	1
S2	3	3	0	1	1
S3	6	6	0	1	1

Table 3. Identity retention of players in different team

Data set	Number of occlusions (Ground Truth)	Retention	Mismatch	Precision	Recall
S1	2	1	1	1	0.5
S2	1	1	0	1	1
S3	4	4	0	1	1

Table 4. Comparison between traditional and proposed technique of different players

Dataset	Kalman [29] based tracking	Proposed Approach
Before occlusion Frame: 44th		
Mid Of Occlusion Frame :75th		
Demerge Frame: 85th		

Table 5. Comparison between traditional and proposed technique of same players

Dataset	Kalman [29] based tracking	Proposed Approach
44th Before occlusion		
Start with Occlusion Frame 50th		
Demerge Frame: 85th		

4.2 Comparison of identity retention with Kalman prediction and Hungarian cost assignment method

Based on the results in Table 6, it can be depicted that Kalman filter with Euclidian distance works well in terms of retention of Identity of player. One major drawback of using Kalman's prediction system is its velocity/acceleration doesn't vary constantly. Due to that Kalman with Euclidian and prediction gives inconsistent results. In comparison to that proposed statistical feature based algorithm gives all correct result.

Table 6. Comparison between state of the art and proposed technique on standard dataset

Parameters	Fabio Previtali* (Single camera) [30]	Proposed approach
MOTA (↑) (%)	72.3	80.1
MOTP (↑) (%)	61.2	64.4
ID-S (↓)	21	15

One most important advantage of the proposed approach is even if for one occlusion it gives an incorrect result. On another occlusion it will rectify itself as algorithm will compare with its features. Compare to the proposed approach, in Kalman based tracking algorithm once identity is changed, it cannot be retained back.

To evaluate the performance, the most widely used CLEAR MOT [42] parameters are used to verify the accuracy.

Multi Object Tracking Accuracy (MOTA) reflects the number of false positives, the number of missed detections, and the identity switches. Multi Object Tracking Accuracy (MOTP) reflects the precision of the detection. ID-S reflects the total switched identities.

Here, ↑ indicates higher the better and ↓ indicates lower the better.

The proposed algorithm is simulated on standard dataset ISSIA-Soccer on sequence 1. Based on the results depicted in Table 6, proposed approach outperforms state of the art method. It has higher accuracy and precision and it has less identity switches.

5. Conclusion

This paper proposed the algorithm to retain the identification of the players under intense occlusion. The proposed model also includes player assignments; occlusion and demerge detection. Based on the experiments proposed approach outperforms traditional algorithm. The proposed approach retains the identity of players when they have similar characteristics like same colors, shape and size. Also, it gives a high performance when players have erratic motion. Extensive experimental results compared with state of the art methods validate the robustness and effectiveness of the proposed approach. The proposed approach has higher accuracy, precision with lower identity switches. This work can be extended to player's game analysis and understanding tactics of the team.

References

[1] T. Zhang and B. Ghanem, "Robust Multi-object Tracking via Cross-Domain Contextual Information for Sports Video Analysis", In: *Proc. of International Conf. on Acoustics, Speech and Signal Processing*, pp.985-988, 2012.

[2] S. Kuppuswamy and B. Panchanathan, "Similar Object Detection and Tracking in H.264 Compressed Video Using Modified Local Self Similarity Descriptor and Particle Filtering", *International Journal of Intelligent Engineering and Systems*, Vol.10, No.5, pp. 95-104, 2017.

[3] S. Ojha and S. Sakhare, "Image Processing Techniques for Object Tracking in Video Surveillance- A Survey", In: *Proc. of IEEE International Conf. on Pervasive Computing*, pp.1-6, 2015.

[4] A. Nazib and C. Oh, "Object Detection and Tracking in Night Time Video Surveillance", In: *Proc. of 10th IEEE International Conf. on Ubiquitous Robots and Ambient Intelligence*, pp.1-6, 2015.

[5] G. Capi and H. Toda, "A Vision based Robot Navigation and Human Tracking for Social Robotics", In: *Proc. of International Workshop on Robotics and Sensor Environment*, pp.1-6, 2010.

[6] P. Benavidez and M. Jamshidi, "Mobile Robot Navigation and Target Tracking System", In: *Proc. of 6th International Conf. on System of Systems Engineering*, pp.299-304, 2011.

[7] S. Rautaray and A. Agrawal, "Design of Gesture Recognition System for Dynamic User Interface", In: *Proc. of IEEE International Conf. on Technology Enhanced Education*, pp.1-6, 2012.

[8] T. Hongyong and Yu Youling, "Finger Tracking and Gesture Interaction with Kinect", In: *Proc. of 12th IEEE International Conf. on Computer and Information Technology*, pp.214-218, 2012.

[9] S. Denman, V. Chandran, and S. Sridharan, "An Adaptive Optical Flow Technique for Person Tracking Systems", *Pattern Recognition Letters*, Vol. 28, No. 10, pp.1232-1239, 2007.

[10] M. Andriluka, S. Roth, and B. Schiele "People-Tracking-by-Detection and People-Detection-by-Tracking", In: *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp.1-8, 2008.

[11] D. Schulz, W. Burgard, D. Fox, and A. B. Cremers, "People Tracking with a Mobile Robot Using Sample-based Joint Probabilistic Data Association Filters", *International Journal of Robotics Research*, Vol. 22, No. 2, pp.99-116, 2003.

[12] D. Schulz, D. Fox, and J. Hightower, "People Tracking with Anonymous and ID-Sensors Using Rao-Blackwellised Particle Filters", In: *Proc. of the 18th International Joint Conf. on Artificial Intelligence*, pp.921-926, 2003.

[13] T. D'Orazio, M. Leo, N. Mosca, P. Spagnolo, and P. Mazzeo, "A Semi-Automatic System for Ground Truth Generation of Soccer Video Sequences", In: *Proc. of the 6th IEEE International Conf. on Advanced Video and Signal Surveillance*, pp.559-564, 2009.

[14] L. Yong, L. Liew, W. Cheah, and Y. Wang, "Occlusion Handling in Videos Object Tracking: A Survey", In: *Proc. of the 8th International Symposium of the Digital Earth IOP Conf. series earth and environmental science*, pp.12-20, 2014.

[15] R. Baillargeon, "Object Permanence in Five-Month-Old Infants", *Cognition*, Vol. 20, No. 3, pp.191-208, 1985.

[16] T. Zhang, K. Jia, C. Xu, Y.Ma, and N. Ahuja, "Partical Occlusion Handling for Visual Tracking via Robust Part Matching", In: *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp.1258-1265, 2014.

[17] B. Han and C. Paulson, "Tracking of Multiple Objects under Partial Occlusion", Proceedings of the SPIE, *Automatic Target Recognition*, Vol. 7335, No. 733515, 2009.

[18] P. Bilinski, F. Bremond, and M B. Kaaniche, "Multiple Object Tracking with Occlusions using HOG Descriptors and Multi Resolution Images", In: *Proc. of 3rd International Conf. on Imaging for Crime Detection and Prevention*, pp.1-6, 2009.

[19] J. Pan and B. Hu, "Robust Occlusion Handling in Object Tracking", In: *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp.1-8, 2007.

[20] M. N. Ali, M. Abdullah-Al-Wadud, and S. Lee, "Multiple Objects Tracking with Partial Occlusion Handling using salient feature point", *Information Sciences*, Vol. 278, pp.448-465, 2014.

[21] A. Tran and K. Harada, "Depth Aided Tracking Multiple Objects under Occlusion", *Journal of Signal and Information Processing*, Vol. 4, No. 3, pp.299-307, 2013.

[22] T. Yang, S. Li, Q. Pan, and J. Li, "Real time Multiple Objects Tracking with Occlusion Handling in Dynamic Scenes", In: *Proc. of IEEE Computer Society Conf. on Computer*

Vision and Pattern Recognition, pp.970-975, 2005.

[23] Y. Huang and I. Essa, "Tracking Multiple Objects through Occlusions", In: *Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp.1051-1058, 2005.

[24] S. Bidhendi and R. Safabakhsh, "Effective tracking of the players and balls in indoor soccer games in the presence of Occlusion", In: *Proc. of 14th International CSI Computer Conf.*, pp.524-529, 2009.

[25] T. D'Orazio, M.Leo, P.L.Mazzeo, and P.Spagnolo, "Soccer Player Recognition by a Multivariate Features Integration", In: *Proc. of Seventh IEEE International Conf. on Advanced Video and Signal Based Surveillance*, pp.32-39, 2010.

[26] G. Thomson, "Real Time camera tracking using sports pitch markings", *Journal of Real-Time Image Processing*, Vol. 2, No. 2-3, pp.117-132, 2007.

[27] S. Pertuz, D. Puig, and M. A. Garcia, "Analysis of Focus Measure Operators for Shape-from-Focus", *Pattern Recognition*, Vol. 46, No. 5, pp.1415-1432, 2013.

[28] D. Santosh, P. Venkatesh, P. Poornesh, L. N. Rao, and N. A. Kumar, "Tracking Multiple Moving Objects Using Gaussian Mixture Model", *International Journal of Soft Computing and Engineering*, Vol. 3, No. 2, pp.114-119, 2013.

[29] G. Welch and G. Bishop, "An Introduction to the Kalman Filter", *University of North Carolana, Chapel Hill-NC*, 1995.

[30] F. Previtali, D. Bloisi, and L. Iocchi, "A distributed approach for real-time multi-camera multiple object tracking", *Machine Vision and Applications*, Vol. 28 No. 3-4, pp. 421-430, 2017.

[31] J. Munkres, "Algorithms for Assignment and Transportation Problems", *Journal of the Society for Industrial and Applied Mathematics*, Vol. 5, No. 1, pp.32-38, 1957.

[32] A. Thelen, "Improvements in Shape-From-Focus for Holographic Reconstructions With Regard to Focus Operators, Neighborhood-Size, and Height Value Interpolation", *IEEE Transactions on Image Processing*, Vol. 18, No. 1, pp.151-157, 2009.

[33] F. Helmli and S. Scherer, "Adaptive shape from focus with error estimation in light microscopy", In: *Proc. of Image and Signal Processing and Analysis*, pp.188-193, 2001.

[34] R. Minhas, A. A. Mohammed, Q. M. Wu, and M. A. Sid Ahmed, "3D shape from focus and depth map computation using steerable filters", In: *Proc. of the International Conf. on Image Analysis and Recognition*, pp.573-583, 2009.

[35] M. Subbarao, T. Choi, and A. Nikzad, "Focusing techniques", *Journal of Optical Engineering*, Vol. 32, No. 11, pp.2824-2836, 1993.

[36] W. Huang and Z. Jing, "Evaluation of focus measures in multi-focus image fusion", *Pattern Recognition Letters*, Vol. 28, No. 4, pp.493-500, 2007.

[37] A. S. Malik and T. S. Choi, "A novel algorithm for estimation of depth map using image focus for 3D shape recovery in the presence of noise", *Pattern Recognition*, Vol. 41, No. 7, pp.2200-2225, 2008.

[38] Santos, C. O. de Solorzano, J. J. Vaquero, J. M. Pena, N. Mapica, and F. D. Pozo, "Evaluation of autofocus functions in moleclar cytogenetic analysis", *Journal of Microscopy*, Vol. 188, No. 3, pp.264-272, 1997.

[39] V. Hilsenstein, "Robust Autofocusing for Automated Microscopy Imaging of Fluorescently Labelled Bacteria", In: *Proc. of Digital Image Computing: Techniques and Applications*, pp.1-7, 2005.

[40] Y. Sun, S. Duthaler, and B. J. Nelson, "Autofocusing in Computer Microscopy: Selecting the Optimal Focus Algorithm", *Microscopy Research and Technique*, Vol. 65, No. 3, pp.139-149, 2004.

[41] Y. Ohno, "Tracking players and estimation of the 3D position of a ball in soccer games", In: *Proc. of 15th International Conf. on Pattern Recognition*, pp.145-148, 2000.

[42] K. Bernardin and R. Stiefelhagen, "Evaluating Multiple Object Tracking Performance: The CLEAR MOT metrics", *EURASIP Journal of Image and Video Processing*, Vol. 2008, No. 1, pp. 1–10, 2008.