Indian Statistical Institute, Kolkata

M. Tech. (Computer Science) Dissertation Series

On The (Soccer) Ball

A dissertation submitted in partial fulfillment of the requirements
for the award of Master of Technology
in
Computer Science
July, 2016

Author:
Arnab Kundu
Roll No: CS1422

Supervisor:
Prof. Dipti Prasad Mukherjee
ECSU Unit, ISI

Declaration

I, **Arnab Kundu** (**CS1422**), registered as a student of **M. Tech** program in **Computer Science**, **Indian Statistical Institute**, **Kolkata** do hereby submit my Dissertation Report entitled "**On The (Soccer) Ball**". I certify

1. The work contained in this Dissertation Report is original and has been done by me under the guidance of my supervisor.
2. The material contained in this Dissertation Report has not been submitted to any University or Institute for the award of any degree.
3. I followed by guidelines provided by the Institute in preparing the report.
4. Whenever I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the text of report and giving their details in the bibliography.

....

Place : ISI, Kolkata **Arnab Kundu**

Date : July, 2016 **(CS1422)**

M.Tech(CS) DISSERTATION THESIS COMPLETION CERTIFICATE
Student: Arnab Kundu (CS1422)
Topic: On The (Soccer) Ball.
Supervisor: Dr. Dipti Prasad Mukherjee

This is to certify that the thesis titled "***On The (Soccer) Ball***" submitted by **Arnab Kundu** in partial fulfillment for the award of the degree of Master of Technology is a bonafide record of work carried out by him under my supervision. The thesis has fulfilled all the requirements as per the regulations of this Institute and, in my opinion, has reached the standard needed for submission. The results contained in this thesis have not been submitted to any other university for the award of any degree or diploma.

Date: Dr.Dipti Prasad Mukherjee

Acknowledgments

First of all, I would like to thank Dr. Dipti Prasad Mukherjee, Professor and Head, Electronics and Communication Science Unit, Indian Statistical Institute, Kolkata for guide me and helping me to undertake the work over last year. Thank you sir, thank you for all the patience you had with me.

I would like to thank all the faculty members of Computer Science division in ISI for their unparallel help and doubt clearance to pursue my career here.

I would like to specially thank Samriddha Sanyal, research scholar at ECSU and NBHM scholar, for carrying out the work altogether enthusiastically.

I sincerely thank all of my lab mate in ECSU specially Bikash Santra, Swapna Agarwal, Angshuman Paul, Archan Ray, Nishant Kumar and all others for their inspiration and their immense help. At the same time I would like to thank all my classmates, seniors and juniors in ISI to bear with me for the last two years.

I am also grateful to Dr. Swagat Kumar for giving me an exposure to work at TCS Innovation Lab, Delhi during last summer. I am specially thankful to Sourav Garg of TCS Innovation Lab. He has been truly inspiring to bring out the best of me.

I sincerely thank Indian Statistical Institute, Kolkata for providing me an opportunity to pursue my master's study in such an emerging field and bestowing us with unmatched facilities of libraries and labs. I would also thank everyone whoever impacted my life in these two years of study.

Finally it would have been incomplete without my special acknowledgment to my parents, grandmother, brother Arup Kundu and all my family members whoever are with me from my childhood. It is for your constant support in all thicks and thins that I exist. This dissertation is dedicated to you.

Abstract

The problem of football tracking from live video is a difficult one. Successful tracking depends critically on the ability of the algorithm to balance prior constraints continuously against evidence garnered from sequences of images. Exact, deterministic tracking algorithms, based on discretized functional, suffer from severe limitations on the form of prior constraint that can be imposed tractably. This paper proposes a particle filter based algorithm that enables to track the ball when it changes its direction suddenly. This algorithm can also track the ball when it takes high speed suddenly and stops. Our algorithm has shown efficient result for partial occlusion and small time full occlusion. Our algorithm has ability to track the ball in spite of Camera movement.

Contents

Chapter 1

Introduction

There are limited applications of computer vision techniques for analysis of sports video [1]. Finding statistics of ball possession by two teams in a soccer match is a challenging problem where computer vision has a large role to play. As a part of our initiative to calculate ball possession statistics, in this paper we present a technique to track soccer ball. Of course the basic tracking scheme is same as it is in any other tracking applications [2] and [3]. However, the major challenge for tracking soccer ball is partial occlusion, and sudden appearance and disappearance of the ball while it is being filmed by multiple camera. This results in two other sub-challenges. The sudden change in soccer ball resolution while it is being filmed using a camera closer to the ball. And then in the next instant, the ball being tracked using a camera far away compared to the previous frames. The other major issue is when the ball is being kicked or passed by a player. The ball can suddenly change its direction of motion and between-frame motion can be significant and most of the conventional trackers fail in such situations. The high speed of ball movement often blurs the image of the ball causing failure of the tracking proposal.

1.1 Objective

Determining ball possession statistics plays the key role for motivating my work. There are very limited number of applications in computer vision techniques for analysis of live sports video according to the best of my knowledge. In order to facilitate that one have to track the players and the ball very robustly and reliably subjected to a variety of challenges mainly due to multiple moving camera and high speed of the ball. However the small size of the ball and blurring lead to poor features to build any deterministic model to track the ball reliably over a long sequence. As a part of determining ball possession statistics as shown in fig 1.1, I take the challenge of ball tracking in a live video.

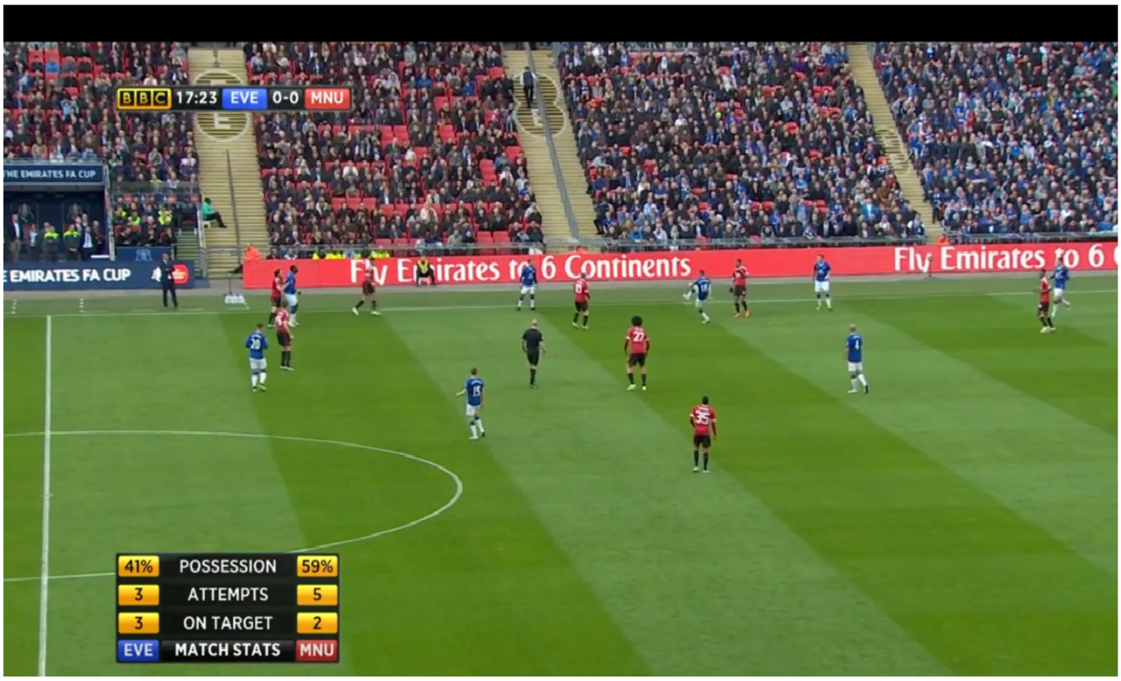

Figure 1.1: Ball Possession Statistics in the course of play

1.2 Related Works

There have been many object detection and tracking framework proposed over last two decades. Yu et al.[4] classify all the algorithms mainly in four categories: 1) feature based 2) model based 3) motion based and 4) data association based.

In [5] Huang *et al.* computed ball position in a frame by integrating segmentation based detection and particle filter-based tracking. However, detection was sometimes failed and then false positive was tracked in successive frames. This is a common issue of ball tracking. Imaged ball cannot be differentiated from other objects robustly based only on its appearance because of their size and the poor features. Yu et al.[4] tried to do ball detection, they estimate hitting points based on the players position and the hitting sound. This method, however, is not always applicable.It is really a difficult problem to extract the sound generated from the targeted object. Junliang Xing [6] tracked multiple players in a sports video by dual-mode two-way Bayesian inference approach which dynamically switches between an offline general model and an online. But they don't get success for ball which is very small comparing to players.

Seo et al. in [7]use Kalman filter based template matching procedure to track the ball. They use backprojection to predict possible occlusion. However color-distribution based template matching suffers from the problem of similar backgound color due to players' jersey color and shoe. Yamada et al. in [8] use the idea of trajectory, but restricted to small trajectories over a small number of neighboring frames used for confirmation. Ohno et al. [9] estimate the 3D position of ball in soccer game by tracking the players based on their shirt and pant regions. To extract ball position, position of the player holding the ball is estimated first by fitting a physical model of their movement in 3D space to the observed ball trajectory. But this model is limited to successful tracking in any video sequence filmed by static camera.

Comaniciu et al. [10] propose mean shift based tracking for non rigid objects seen from a moving camera. It find the most suitable position in current frame by computing the similarity between the target model(its color distribution) and target candidates by means of Bhattacharyya coefficient.In feature based algorithms, some features of the objects are used to discriminate targets from other objects within a frame. Some algorithms make use of reference image of the background called the background frame. All objects in the difference frame are obtained by subtracting the background frame from the current frame, which are the targets. In [11] the features with the manually labelled targets are used to train a neural classifier and then the trained neural classifier is used to differentiate the target from other objects.

Yogesh et al. [12] try to track deforming objects. They formulate a particle filtering based algorithm in the geometric active contour framework that can be used for tracking moving and deforming objects. Geometric active contours provide a framework which is parametrization independent and allow for changes in topology. Jia et al. [13] uses structural local sparse appearance model which exploits both spatial and temporal information of the target based on a novel alignment-pooling method.

1.3 Contribution

Due to abrupt motion or sudden change in direction, we rely on a proposal based approach. Based on information in the current and the previous frame, we put forward a set of dynamic proposals for possible location of the soccer ball in the next frame. The proposals are made based on a concept similar to particle filters. We then propose a novel strategy to select the winner particles among the different proposals. The proposal is based features collected from the ball, namely, the color, edge gradient and shape measure. An iterative scheme weighs each proposal point based on the likelihood of the proposal to lie on the contour of the ball in the next frame.

Chapter 2

Methodology of the proposed approach

Mainly, our soccer ball tracking algorithm consists of following four steps: (a) initialization of first two frames, (b) prediction of the measurement space, (c) weight assignment of the potential points, and (d) resampling of the potential points. The prediction step defines a dynamic model that generates a set of potential points, which are called particles. These particles are possible potential points on the edge of the soccer ball in the frame where we are searching for the soccer ball. In the weight assignment step, we are assigning weights to the predicted particles. Finally, in the resample step a collection of winners are selected from the generated particles at the last stage based on the likelihood ratio of the particles to fit a soccer ball model.

2.1 Initialization

In the proposed tracking model, we assume that at any instance, to predict the ball position in the current frame, we have the information about the position of the ball in last two frames. In order to facilitate that, ball position at first two frames in any video sequence need to be initialized. We initialize the ball position at any frame by choosing the centre of the ball and any point on the contour of the ball. These two informations are sufficient to capture the shape of the circle-like ball.

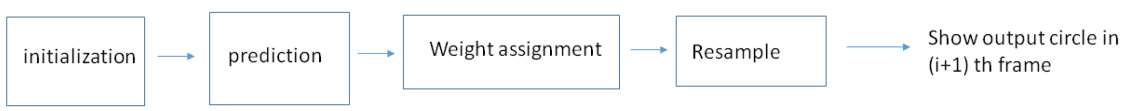

Figure 2.1: Block diagram of the proposed football tracking algorithm

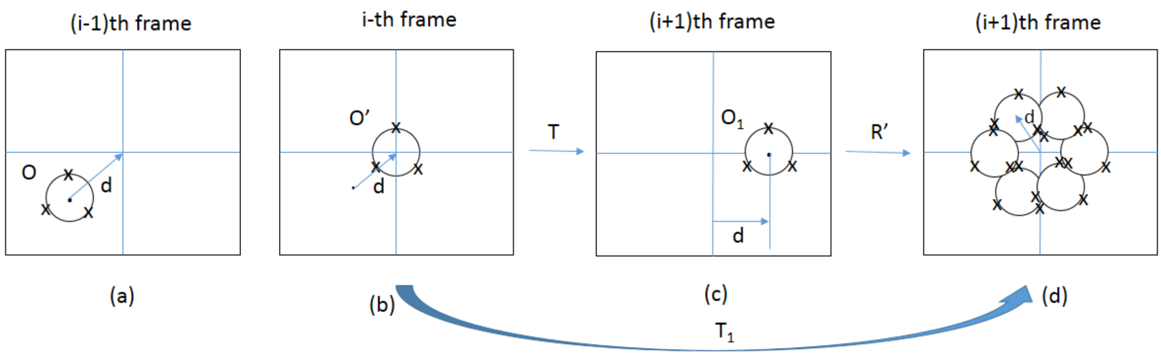

Figure 2.2: Construction of measurement space: (a) and (b): Circle representing the ball at i-1 and $i$the frame. The centre of the ball is at O and O' respectively. (c) The position of the ball in frame i+1 after translation by $(d, 0)$. The translated circle is at center O_1. (d) The *measurement space* constructed by rotating the translated circle with center at O_1 with radius d.

2.2 Prediction

In the proposed tracking model, we assume that at any instance, the information of the soccer ball for previous two frames are available. The points which are located on the contour of the circle-like ball, are termed as landmark points. Our main objective at any instance is, given a set of landmark points on the known ball positions on previous two frames, we predict a set of landmark points on the current frame. Let Figs. 2.2(a) and (b) represent the ball position at frame i-1 and i respectively. Let the distance between the centres of the ball positions, O and O' in frame i-1 and i respectively, be d. Let us assume a *measurement space* at frame i+1 centred at a distance d from position O'. Let us define this centre of *measurement space* be O_1 as shown in Fig. 2.2(c). We are assuming that the radius of the ball at O_1 is same as that of at O' and the radius of the ball is known a priori. The locus of O_1 in a circular path centred around O' with radius d gives the *measurement space*. For discrete representation of the *measurement space*, we have selected l number of locations on the locus of point O_1. Therefore, at every b location out of total l positions, we can define a circle whose center is rotated at an angle θ_b with respect to positive x-axis. Finally, we term the dynamics of giving translation by $(d, 0)$ and rotation by angle θ_b as $T^1_{\theta_b}$. These circles are shown in Fig. 2.2(d).

We now have to select a set of particles or landmark points. For ease of understanding, let us assume that there are only three landmark locations. The soccer ball locations upto frame i are known and these points are shown in Fig. 2.2(a) and (b) with $\times$. Let (x_1, y_1), (x_2, y_2) and (x_3, y_3) are these three landmark points in frame i. Our task is to select landmark points from the measurement space of the i+1 frame. Given a discrete locations b on the locus of O_1, angle θ_b, particles or landmark points on the ball positions at b can be calculated after applying appropriate rotation transform to (x_1, y_1), (x_2, y_2) and (x_3, y_3). Again, these landmark points are shown in Fig. 2.2(d). Landmark points (x_i, y_i) for a possible ball location in Fig. 2.2(d) at an angle θ_b is stored in a configuration matrix $M^{\theta_b}_{(i+1)}$. Each row of $M^{\theta_b}_{(i+1)}$ represents each landmark point location on circle rotated at an angle θ_b in measurement space corresponding to frame $(i+1)$. Thus for all the possible ball locations in Fig. 2.2(d), the universal configuration matrix Ω is formed after concatenating $M^{\theta_b}_{(i+1)}$ for all b discrete locations.

The proposed *measurement space* captures movement of the ball in every orientation. Suppose the centroid of the ball is moving at the direction $\vec{v}$. Then corresponding unit vector is $\vec{v}/\|(\vec{v})\|$. Along this direction we get position of the center of the circle(the translated circle we are rotating) which is $d\vec{v}/\|(\vec{v})\|$. We also get new landmark point locations on the circle of r radius around the transformed center($d\vec{v}/\|(\vec{v})\|$). So we can capture the boundary of the circle containing moving ball in the third frame by this measurement space. The proposal above is subjected to an assumption. Given, video frame dimension as $r \times c$, $d \leq \min(\frac{r}{2}, \frac{c}{2})$, which is very much practical.

In the next section, we utilize universal configuration matrix $M_{(i+1)}$ to weight the selection of likely ball location.

2.3 Weight Assignment

Before presenting the Weight Assignment strategy, we need to understand the tracking framework. The tracking model is based on second order Markov model which is detailed next.

2.3.1 Tracking Framework

Suppose we have information on the first n frames. That is, we know the centre and radius of the soccer ball till nth frame. We denote the ball as a circle at the nth frame as X_n. To predict the position of the ball in $(n+1)$th frame, the aim is to grow such a sequence of circles based on a prior dynamics $p(X_{i+1}|X_{0:i})$ where $i \in \{1, 2, \dots, n\}$. The prior dynamics is expected to retain the properties of the circular region representing the ball. Based on homogeneous second-order dynamics with some probability function q, we define

$$p(X_{i+1}|X_{0:i}) = q(X_{i+1}|X_{i-1:i}), \forall i \geq 2. \tag{2.1}$$

Now we will discuss about q.

2.3.2 Dynamics q

Suppose we are given information(i.e. the circle containing the ball) of i consecutive frames. Based on these circles we have to locate the ball in $(i+1)$th frame i.e we have to fit the best circle around the ball. In the proposed algorithm, based on dynamics q, we predict the $(i+1)$ th frame circle based on the information of i, $(i-1)$ th frame information. We take the dynamics of step length 2. The definition of the second-order dynamics $q(X_{i+1}; X_{i-1:i})$ amounts to specifying an a priori probability distribution on direction change $\theta_b \in (0, 2\pi]$.

One of our aim is to find the curve of the ball i.e. the edge of the ball. The smoothness of the curve(edge of ball) can be simply controlled by choosing this distribution as Gaussian with variance $\sigma^2_{\theta_b}$ per length unit. For steps with length 2(As our dynamics is second order) the resulting angular variance is $2\sigma^2_{\theta_b}$. In order to allow for abrupt direction changes at the few locations where curve of edge of the ball have been detected, we mix the normal distribution with a small proportion ν of uniform distribution over $[-\frac{\pi}{2}, \frac{\pi}{2}]$. The dynamics finally reads :

$$T^1_{\theta_b}(V) = V_{\theta_b} \forall V \in S^1,$$

where $T^1_{\theta_b}$ has been defined in section (2.2) of particle creation. q is :

$$q(\theta_b) = \frac{\nu}{\pi} + (1-\nu)N(\theta_b; 0, 2\sigma^2_\theta) \tag{2.2}$$

Statement: Given information (that is, centers and radii of all circles that contain the ball) up to n frames at a time, the probability $p(X_{0:n})$ satisfies following relation.

$$p(X_{0:n}) = p(X_{0:1}) \prod_{i=2}^{n} q(X_i; X_{i-2:i-1}). \tag{2.3}$$

The proof of the above relation is as follows:

$$\begin{aligned}
p(X_{0:n}) &= p(X_0 \wedge X_1 .... \wedge X_n) \\
&= p(X_n \wedge X_0 .... \wedge X_{n-1})[p(A \wedge B) = p(B \wedge A)] \\
&= p(X_n | X_0 .... \wedge X_{n-1}) p(X_0 \wedge .... \wedge X_{n-1})[p(A \cap B) = p(A|B)p(B)] \\
. &= p(X_n | X_0 .... \wedge X_{n-1}) p(X_{n-1} | X_0 \wedge .... \wedge X_{n-2}) .... p(X_2 | X_{0:1}) p(X_{0:1}) \\
&= q(X_n; X_{n-1:n}) q(X_{n-1}; X_{n-3:n-2}) .... q(X_2; X_{0:1}) p(X_{0:1}) \\
&= p(X_{0:1}) \prod_{i=2}^{n} q(X_i; X_{i-2:i-1})
\end{aligned}$$

2.3.3 Data Model

The evidence supporting the measurement that a predicted contour is in the vicinity of the “true” contour containing the ball, is defined as the data model. The data model is represented as $p(Y(u)|X_{0:n})$. Intuitively, Y represents collection of features derived from pixel at image location u.

Formally, $Y : \Omega \to \Re^3$. $Y(u)$ is defined as $Y(u) = (Y_1(u), Y_2(u), S'(u))$, where u is a point in Ω (universal configuration matrix).

$Y_1 : \Omega \to \Re$, $Y_1(u) = I(u)$, where $I(u)$ represents the intensity value at gray scale at $u \in \Omega$.

$Y_2 : \Omega \to \Re$, $Y_2(u) = |\nabla I(u)|$.

S' is the shape measure defined as $S'(u) = S(M^{\theta_b}_{(i+1)})$, where

$$S(M^{\theta_b}_{(i+1)}) = \sqrt{\sum_{p=1}^{k}\sum_{q=1}^{2}(M^{\theta_b}_{(i+1)}[pq] - M^{\bar{\theta_b}}_{(i+1)}[q])^2}. \tag{2.4}$$

The $M^{\theta_b}_{(i+1)}$ is a $(k \times 2)$ is a configuration matrix which contributes the point u in universal matrix Ω, $M^{\theta_b}_{(i+1)}[pq]$ represents the element of pth row and qth column of the matrix $M^{\theta_b}_{(i+1)}$. $M^{\bar{\theta_b}}_{(i+1)} = \frac{\sum_{p=1}^{k} M^{\theta_b}_{(i+1)}[pq]}{k}$. The above defined S is a *size measure* [14].

Definition(Size Measure): Let $A = \{\, X \mid X \text{ is a matrix} \,\}$. The size measure $g : A \to \Re^+ \bigcup \{0\}$ defined by $g(aX) = ag(X)$ where a is any positive scalar.

Statement:

S which is defined above is a size measure.

Proof: To prove the above statement we have to show that for any scalar $a \geq 0$ we will have $S(aM^{\theta_b}_{(i+1)}) = aS(M^{\theta_b}_{(i+1)})$ where $M^{\theta_b}_{(i+1)}$ is a configuration matrix. We have $S(M^{\theta_b}_{(i+1)}) = \sqrt{\sum_{p=1}^{k}\sum_{q=1}^{2}(M^{\theta_b}_{(i+1)}[pq] - M^{\bar{\theta_b}}_{(i+1)}[q])^2}$; where $M^{\bar{\theta_b}}_{(i+1)} = \frac{\sum_{p=1}^{k} M^{\theta_b}_{(i+1)}[pq]}{k}$. a is a positive scalar.Then consider the matrix $aM^{\theta_b}_{(i+1)}$. So each entry of the matrix $aM^{\theta_b}_{(i+1)}$ is $aM^{\theta_b}_{(i+1)}[pq]$.

$$\begin{aligned}
S(aM^{\theta_b}_{(i+1)}) &= \sqrt{\sum_{p=1}^{k}\sum_{q=1}^{2}(aM^{\theta_b}_{(i+1)}[pq] - aM^{\bar{\theta_b}}_{(i+1)}[q])^2} \\
&= \sqrt{a^2\sum_{p=1}^{k}\sum_{q=1}^{2}(M^{\theta_b}_{(i+1)}[pq] - M^{\bar{\theta_b}}_{(i+1)}[q])^2} \\
&= |a|\sqrt{\sum_{p=1}^{k}\sum_{q=1}^{2}(M^{\theta_b}_{(i+1)}[pq] - M^{\bar{\theta_b}}_{(i+1)}[q])^2} \\
&= |a|S(M)
\end{aligned}$$

Thus S is a size measure.

We are assuming:

$$\begin{aligned}
p(Y(u)) &\propto [p(Y_1(u)) + p(Y_2(u)) + p(S'(u))] \\
&= C[p(Y_1(u)) + p(Y_2(u)) + p(S'(u))]
\end{aligned}$$

Where C is normalization constant.
Hence $p(Y(u)|X_{0:n})$=$C[p(Y_1(u)|X_{0:n}) + p(Y_2(u)|X_{0:n}) + p(S'(u)|X_{0:n})]$.
We also approximate measurements on Ω as an independent spatial process

$$p(Y|X_{0:n}) = \prod_{u\in\Omega} p(Y(u)|X_{n-1:n}) \tag{2.5}$$

Each individual likelihood in the product of (2.5) is either termed as p_{on}, if we consider the given condition (given $X_{0:n}$) or is termed as p_{off}, if the given condition is not there.

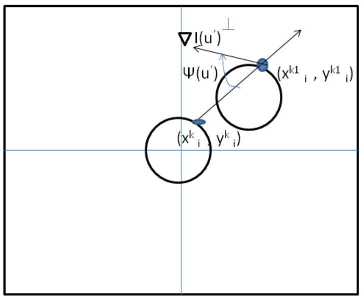

Figure 2.3: (x_i^k, y_i^k) is position of the landmark point in i th frame. After giving dynamics we get its new position $(x_i^{k_1}, y_i^{k_1}) = u$. $\nabla I(u)^{\perp}$ has made angle $\psi(u)$ with the line joining (x_i^k, y_i^k) and $(x_i^{k_1}, y_i^{k_1})$

So, we can write (2.5) as:

$$
\begin{aligned}
p(Y \mid X_{0:n}) &= p(Y|X_{n-1:n}) \\
&= \prod_{u\in\Omega - X_n} p_{off}(Y(u)) \prod_{u\in(\Omega\bigcap X_n)} p_{on}(Y(u)|X_{n-1:n}) \\
&= \prod_{u\in\Omega} p_{off}(Y(u)) \prod_{u\in(\Omega\bigcap X_n)} \frac{p_{on}(Y(u)|X_{n-1:n})}{p_{off}(Y(u))}
\end{aligned}
$$

we consider $l = \frac{p_{on}}{p_{off}}$, which denotes point-wise likelihood ratio similar to the likelihood ratio defined in [15]. Based on l, we resample the particle set which will be discussed in the next subsection. So the above equation becomes

$$
p(Y \mid X_{0:n}) = \prod_{u\epsilon\Omega} p_{off}(Y(u)) \prod_{u\epsilon(\Omega\bigcap X_n)} l(Y(u)). \tag{2.6}
$$

2.3.4 Likelihood ratio estimation

Most data terms for our football tracking algorithm are based on the spatial gradient in intensity or RGB space, and shape measure. To use before mentioned attributes as part of our measurements, we must capture its marginal distributions both off the circle containing football in previous frames (p_{off}) i.e when we are not considering the given circles containing football and p_{on} when we are considering the given circles containing the football.

The first marginal p_{off} can be can be empirically captured by the distribution of the norm of the gradient, pixel value and shape measure of the measurement space Ω. In our experiments, these empirical distributions were always well approximated by an exponential distribution with parameter λ (amounting to the average norm for the attribute gradient over the measurement space Ω), which we take as p_{off}. Let $u \in \Omega$.

$$
p_{off}(u) \propto e^{-\frac{\nabla I(u)}{\lambda}} + \frac{e^{-I(u)}}{\sqrt{2\pi}\lambda} + e^{-\|(u-c)\|}.
$$

Where c is the center of the last frame.

For p_{on}, it is difficult to learn it a *priori.* We assign normal distribution for color, i.e. Y_1

$$p_{on}(Y_1(u)|X_{i-1:i}) = \frac{e^{-(d^2)/2\sigma^2}}{\sqrt{2\pi\sigma^2}}$$

where d is the euclidean distance between the intensity (i.e. color) of u intensity of the ball in gray scale. Thus the minimum the difference of the intensities will be, the probability would be maximum.

Next we discuss on the probability distribution assigned for gradient of the pixel.The empirical distribution of the gradient over an outline of interest appears as a complex mixture filling the whole range of values from 0 to a large value of gradient norm.In the absence of an appropriate statistical device to capture adaptively this highly variable behavior, it seems better to keep the data likelihood p_{on} as less informative as possible.Now,Our aim is to give two different distribution function for different potential points, which are most likely to lie on the edge of ball and very less likely to lie on the edge of the ball.
In order to facilitate that, we define another function $\rho : \Omega \rightarrow \{0, 1\}$ to assign different probablity distribution to a point in Ω. Thus,

$$p_{on}(Y_2(u)|X_{i-1:i}) = \frac{\rho(u)}{\pi} + (1 - \rho(u))N(\psi(u); 0, \frac{\sigma_\psi^2}{|\nabla I(u)|}).$$

Depending on the value of ρ whether it is 0 or 1 the above equation will assign distribution to the corresponding point. Thus, if the point is more likely to to lie on the edge, uniform distribution will be assigned to it, otherwise a normal distribution will be assigned, where $\psi(u)$ is the angle referred in Fig.2.3.

Now to define the function ρ, first we have to define a threshold for each frame. This threshold for current frame is calculated based on the average gradient difference for all landmark points on the contour of the ball on last two frames. Suppose we have M number of mathematical landmark points set for frame $(i - 1)$ and i. We denote the mathematical landmark point set by $\{(x_{i-1}^k, y_{i-1)}^k\}_{k=1}^M$ for the frame $(i - 1)$ and $\{(x_i^k, y_i^k)\}_{k=1}^M$ for the frame i. Thus the threshold for frame $(i + 1)$ is calculated as :

$$TH_{i+1} = 1/M \sum_{k=1}^{M} ||\nabla I(x_{i-1}^k, y_{i-1}^k| - |\nabla I(x_i^k, y_i^k)||$$

Let (p, q) be a point of configuration matrix in the $(i + 1)$th frame corresponding to the same point (x_i^k, y_i^k) in the ith frame. Now $\rho(p, q) = 1$, i.e. (p,q) point is more likely to lie on the edge of the ball in the $(i + 1)$th frame if

$$||\nabla I(p, q)| - |\nabla I(x_i^{k_1}, y_i^{k_1})|| \leq TH_{i+1}.$$

Otherwise, $\rho(p, q) = 0$.

Now we assign exponential distribution for shape, i.e. S', as

$$p_{on}(S'(u)|X_{i-1:i}) = e^{-\frac{d_1}{max(r_{i-2},r_{i-1})}}$$

where d_1 is the euclidean distance between u and center of last frame circle; r_{i-2} and r_{i-1} are the radius of the circles in $(i-2)$ and $(i-1)$th frames respectively.

Finally the p_{on} of the data model is combined as

$$p_{on}((Y_1(u), Y_2(u), S'(u))|X_{i-1:i}) \propto \frac{\rho(u)}{\pi} + (1-\rho(u))N(\psi(u); 0, \frac{\sigma_\psi^2}{|\nabla I(u)|}) + \frac{e^{-(d^2)/2\sigma^2}}{\sqrt{2\pi\sigma^2}} + e^{-\frac{d_1}{max(r_{i-2},r_{i-1})}}$$

Thus the likelihood ratio for a point in the configuration matrix can be computed as $l = \frac{p_{on}}{p_{off}}$, which is deduced up to a multiplicative constant.

Now we derive how likely our predicted circle can circumscribe the real ball based on $Y(u)$, the collection of features derived for pixel at image location u. We call it posterior density $p(X_{0:n}|Y(u))$.

Statement:

$$p_n(X_{0:n}|Y) \propto p(X_{0:1}) \prod_{i=2}^{n} q(X_i; X_{i-2:i-1}) \prod_{u\epsilon(\Omega \bigcap X_n)} l(Y(u)) \tag{2.7}$$

Proof: Here we are concluding the above statement:

$$\begin{aligned} p(X_{0:n}|Y) &= \frac{p(X_{0:n} \wedge Y)}{p(Y)} \\ &= \frac{p(Y|X_{0:n})p(X_{0:n})}{p(Y)} \\ &\propto p(Y|X_{0:n})p(X_{0:n}) \end{aligned}$$

Substituting the expression of $p(X_{0:n})$ from (2.3) and expression of $p(Y \mid X_{0:n})$ from (2.5) we get:

$$\begin{aligned} p(X_{0:n}|Y) &= \frac{p(Y|X_{0:n})p(X_{0:n})}{p(Y)} \\ &\propto p(Y|X_{0:n})p(X_{0:n}) \\ &\propto p(X_{0:1}) \prod_{i=2}^{n} q(X_i; X_{i-2:i-1}) \prod_{u\in\Omega} p_{off}(Y(u)) \prod_{u\in(\Omega \bigcap X_n)} l(Y(u)) \\ &\propto p(X_{0:1}) \prod_{i=2}^{n} q(X_i; X_{i-2:i-1}) \prod_{u\in(\Omega \bigcap X_n)} l(Y(u)) \end{aligned}$$

So the statement is proved.

2.3.5 Weight(probability) giving strategy

Our football tracking algorithm is based on recursive computation of posterior densities of interest. From (2.6) we construct the following recursive relation :

$$p_{i+1}(X_{0:i+1} \mid Y) \propto p_i(X_{0:i}|Y)q(X_{i+1}; X_{i-1:i}) \prod_{u\in(\Omega \bigcap X_{i+1})} l(Y(u)), i \in \{1, 2, .., n\}. \tag{2.8}$$

Though we have analytical expression for l and q but this recursion cannot be computed analytically. There is no closed form expression of the posterior distributions p_i. We approximate posterior p_i by a finite number of points(Say M number of points) $\{x_i^m\}_{m=1}^M$, here x_i^m are of the form $(r\cos(\theta), r\sin(\theta))$ and their probability or weight are $\{p_i^m\}_{m=1}^M$. Now we have to generate particles(possible landmark points on the contour of the ball so that we can create the $(i+1)$ th frame circle) from the distribution p_{i+1}. The generation of samples from p_{i+1} is then obtained in two process.

Now we discuss the ***prediction*** step. Suppose we give $X_1, X_2, .., X_i$. Based on these information we have to predict X_{i+1}. At first we have chosen landmark points. Then we create measurement space in $(i+1)$th frame around the location of center of ball in ith frame. Every point of the universal configuration matrix Ω is chosen by sampling from the proposal density function $f(x_{i+1}; x_{0:i}^m, Y)$ over Ω.

If the landmark points $(x_{0:i}^m)_m$ are fair sample from the distribution p_i over Ω then the extended landmark point $(x_{0:i}^m, x_{i+1}^m)_m$ are fair samples from distribution fp_i over Ω. Since we are seeking samples from distribution p_{i+1} indeed, so we will go for *importance sampling* principle. Now we are describing the principle. [16] Suppose $p(x) \propto \pi(x)$ is a probability density from which it is difficult to draw samples but for which $\pi(x)$ can be evaluated [as well as $p(x)$ up to proportionality]. In addition,let $x^i \propto q(x)$ $i = 1, 2...N_s$ be the sample that are easily generated from a proposal $q(.)$ called an importance density. Then, a weighted approximation to the density $p(.)$ is given by:

$$p(x) \approx \omega_i \sum_{i=1}^{N_s} \delta(x - x^i);$$

where $\omega_i \propto \frac{\pi(x^i)}{q(x^i)}$. δ is Delta-Dirac measure. We use this to approximate a probability mass function as probability density function. In our case, the sample landmark points are weighted according to ratio p_{i+1}/fp_i [normalized over the number of rows in configuration matrix]. The resulting weighted landmark point set now provides an approximation of the target distribution p_{i+1}. This discrete approximating distribution is used in the second step of *selection* where we will draw some landmark points with replacements from the previous weighted set. The new set will be distributed according to p_{i+1}. The landmark points with smallest weights are likely to get discarded by this selection process, whereas the ones with large weights are likely to get duplicated.

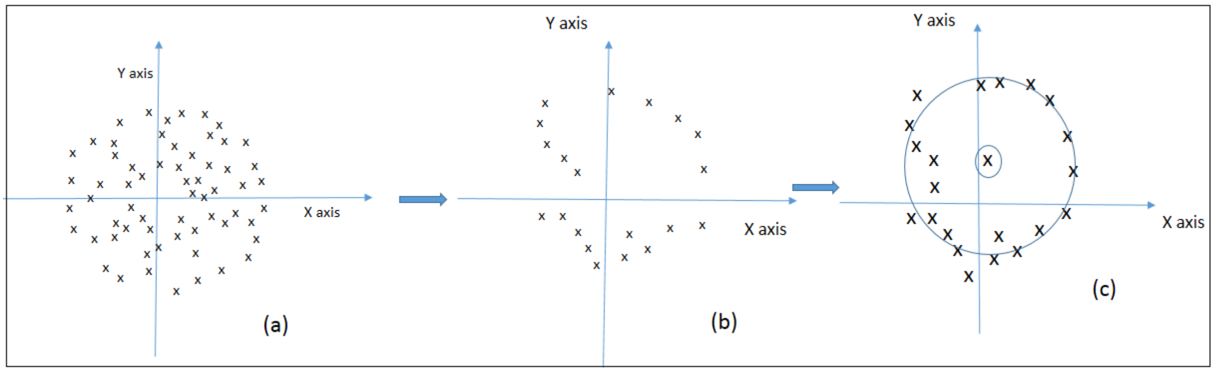

Figure 2.4: Resample Prediction of ball position in $(i+1)$th frame (a) Measurement space as a set of sample points. (b) Resampling k points from each quadrant ($k = 5$). (d) Predicted circle as the ball in $(i+1)$th frame.

Using the expression (7) of p_{i+1}, the ratio $\frac{p_{i+1}}{fp_i}$ becomes $\frac{ql}{f}$. Now the weight reads:

$$\pi_{i+1}^m \propto \frac{q(x\prime_{i+1}^m; x_{i-1:i}^m)l(Y(x\prime_{i+1}^m))}{f(x\prime_{i+1}^m; x_{0:i}^m, Y)} \tag{2.9}$$

with $\sum_m \pi_{i+1}^m = 1$. It can be shown that the optimal proposal pdf is $f = ql/\int_{x_{i+1}} ql$ from [17], whose denominator cannot be computed analytically in our case. The chosen proposal pdf must then be sufficiently "close" to the optimal one such that the weights do not degenerate (i.e., become extremely small) in the re-weighting process.

2.4 Resample

Now we discuss resample of those particles and circle construction. Let in the configuration matrix we have M rows, i.e. total particle number is M. We will resample $4k$ points from M points. $4k \ll M$. We first draw co-ordinate system taking ith frame circle center as origin, x axis of image as x axis and the vector orthogonal to x axis as y axis. Then from each quadrant we resample k points based on likelihood ratio l. If in some quadrant does not have k many points then we resample all points. Then we find out weighted center of mass of all $4k$ resampled particles according to following equation.

$$x_{i+1} = \frac{\sum_{j=1}^{4k} w_j * x_j}{\sum_{j=1}^{4k} w_j}$$
$$y_{i+1} = \frac{\sum_{j=1}^{4k} w_j * y_j}{\sum_{j=1}^{4k} w_j}$$

(x_{i+1}, y_{i+1}) is co-ordinate of the centre of the predicted ball position in $(i+1)$th frame. From this predicted centre we calculate the weighted mean distance of these $4k$ resampled points, which give the radius of the predicted ball.

Thus the predicted circle for frame $(i+1)$ is constructed based on ball position in frame i and frame $(i-1)$.

Algorithm 1 Football tracking algorithm

Given Information(circle): $\{X_i\}_{i=1}^n$ where X_i represents the circle of radius r containing the ball in frame i.

Landmark point selection: for $j = 1...m$, select landmark point on the last circle $x_j = (r\cos\theta_j, r\sin\theta_j)$, $j = 1...m$.

Prediction: $0° \leq \theta_k \leq 360°, T^1_{\theta_k}(x_j) = x^j_{\theta_k} \forall\theta_k$; $k \in \{1, .., n_1\}$, $\forall j \in \{1., m\}$.

Weighting: Total particle set $N = mn_1$. We denote all particle set $\{x^m_{1:n}\}_{m=1}^N$. Compute: $\pi^m_{n+1} = \frac{Kq(\theta_k)l(Y(x\prime^m_{n+1}))}{\frac{\rho(x^m_n)}{\pi}+(1-\rho(x^m_n))N(\theta_k;0,2\sigma^2_\theta)}$

with K such that $\sum_{m=1}^N \pi^m_{n+1} = 1$.

Resample: Total resampled particle $4k$.

Draw axis taking center of X_n as origin. from each quadrant,Draw top k sample point having maximum the discrete probability $\{\pi^k_{n+1}\}_k$ over $\{1, 2, ..N\}$.

Predicted circle: Find center of mass(C.M.) of all $4k$ points. Find distance(d) of all sample points from C.M. Find d_{max}, d_{min}. $r = \frac{(d_{max}+d_{min})}{2}$.

Draw circle with C.M as center and r as radius.

Chapter 3

Experiment and Result

Here, implementation details of the system and experimental results will be presented. We first give a short description of the sports video data used in the experiments. Then, we illustrate ingredients(Likelihood ratio, Dynamics q, Proposal Sampling Function f) used for our football tracking algorithm. Then We evaluate our algorithm with Ground truth and compare with color based mean shift non rigid object tracking algorithm.

Figure 3.1 shows the measurement space creation in 114th frame based on the distance between the centre of already detected ball in $112th and 113th frame$. The resampled paricles and the predicted ball position are also shown in same picture.

Figure 3.2 is the case where inspite of having partial occlusion of the ball in both $152nd and 153rd$ frame, the proposed algorithm is able to capture the ball in $154th$ frame.

3.1 Dataset and Ground truth

One of the motivation of this work is to track football in a live football match. The video data used for our experiment is directly collected from the [18]. This dataset consists of 723 frames each of resolution of (720×1280). In this dataset one team jersey is white. So, it is difficult for any algorithm to track small white ball based only on color information. Camera movement also creates challenge to track football. It makes ball tracking critical by background subtraction method. We also collect another video dataset [19]. In this dataset, only 623 frames contain effective play. Here also each frame has resolution of (720×1280). Here no team has white jersey. We manually label the ground truth of position of the ball in each frame. We compare our algorithm with the representative work in [10] which deals with the problem of non rigid object tracking based on mean shift and color histogram.

3.2 Discussion about some parameters

We have described dynamics of the particle in section 2.2 and discussed about the data model in previous section and the data model is chosen, it remains to devise a proposal

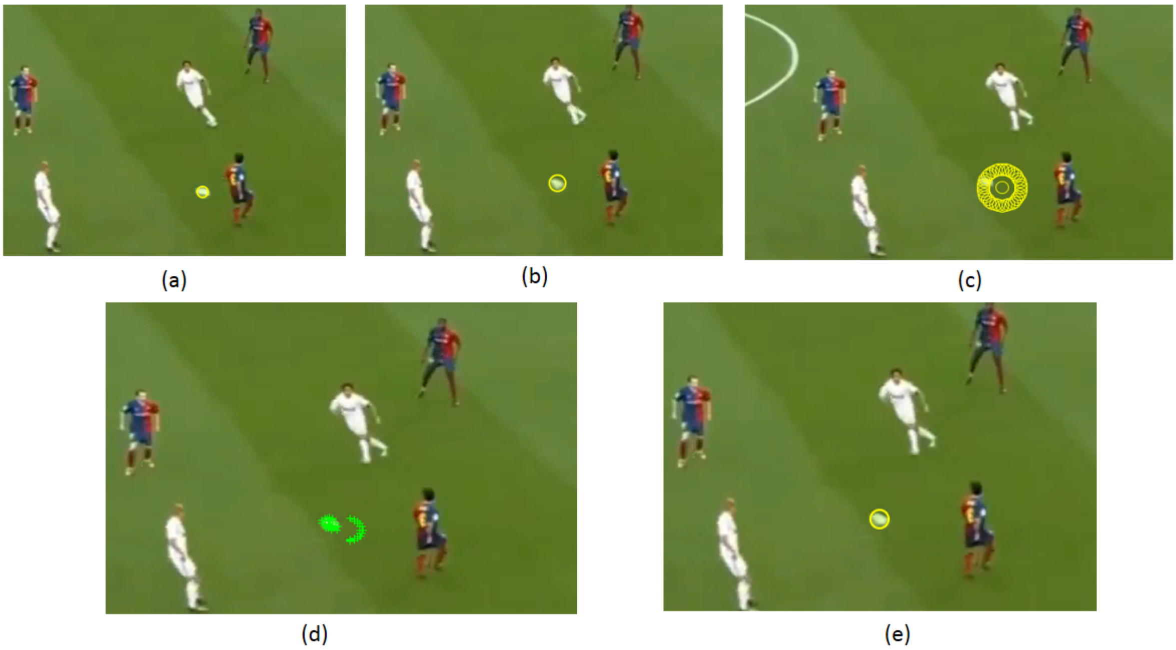

Figure 3.1: (a) Detected ball position in 112th frame. (b) Detected ball position in 113th frame. (d) Measurement Space in 114th frame. (e) Resampled particles in 114th frame. (f) Predicted circle in 114th frame.

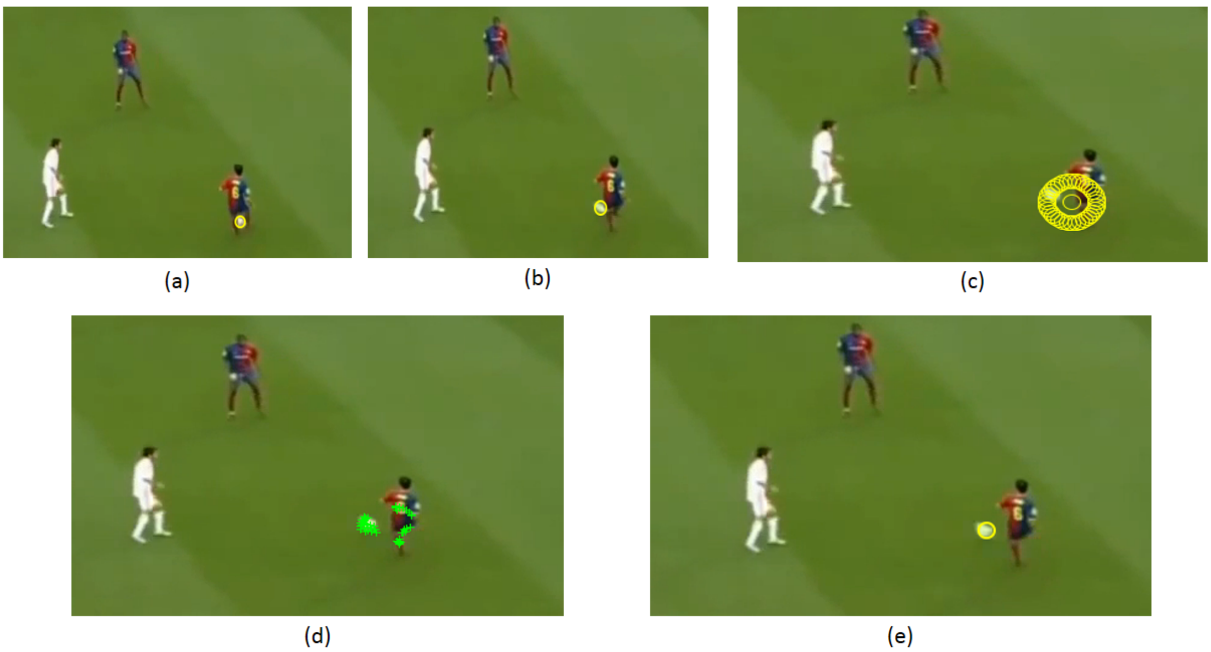

Figure 3.2: (a) Detected ball position in 152th frame. (b) Detected ball position in 153th frame. (d) Measurement Space in 154th frame. (e) Resampled particles in 154th frame. (f) Predicted circle in 154th frame.

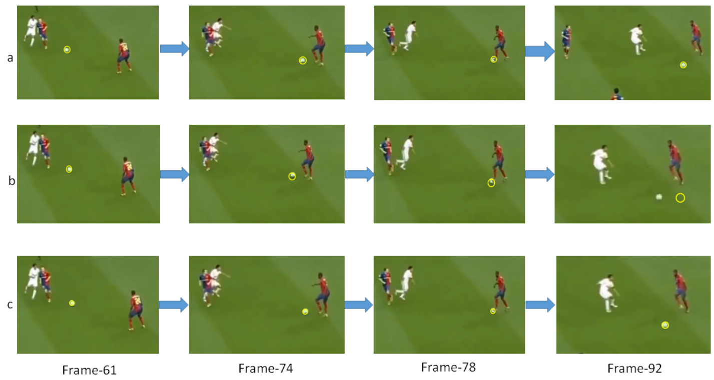

Figure 3.3: Performance evaluation when the ball suddenly changes its direction with high speed, (a) Ground truth result, (b) Result of mean shift algorithm based on color histogram matching, (c) Result of our football tracking algorithm

sampling function f which is as much related as possible to ql under the constraint that it can be sampled from.

Here we are creating 36 new circles(as we have taken $\theta_b = 10°$). So, we got $\frac{360°}{10°} = 36$. We have selected 15 landmark points on the first two frame circles containing the ball. So total number of point on all 36 circles are $15 \times 36 = 540$. We also included the landmark points detected on the last frame circle. So, total points in the measurement space Ω are $540 + 15 = 555$. Here typically $\nu = 0.05$ and $M = 555$.

In resampling step, we are reampling 15 points from each quadrant based on their likelihood. Thus centre of the ball in current frame is predicted from these $15 \times 4 = 60$ resampled points.

3.3 Comparative Results

We first conduct experiment to analyze the performance of our algorithm. The experiment is carried on tracking ball in different critical situation like sudden velocity change of ball, frequent change of ball's direction, frequent camera movement. We also make experimental environment critical by taking jersey color of one team white. Many players wear white shoe. So,shoe also looks like football from long distance. It becomes difficult to track the ball when a player wearing white shoe takes the ball then pass it. We run our algorithm over 721 frames of the video to measure the relevant performance metric.

We take frame 61 to 92 as our experiment frame of dataset [18]. In these frames the football changes its orientation suddenly. In Fig.3.3(a) we are showing our ground truth result where we manually label the ball with yellow circle. In Fig.3.3(b) we run Meanshift algorithm based on color histogram [10] which fails to track the ball when it changes it's direction suddenly. The sudden change in direction is also captured in graphical plot 3.4, which is successfully tracked by our proposed method whereas the competing method fail to track the sudden change in direction. Basically, mean shift is a procedure for locating the

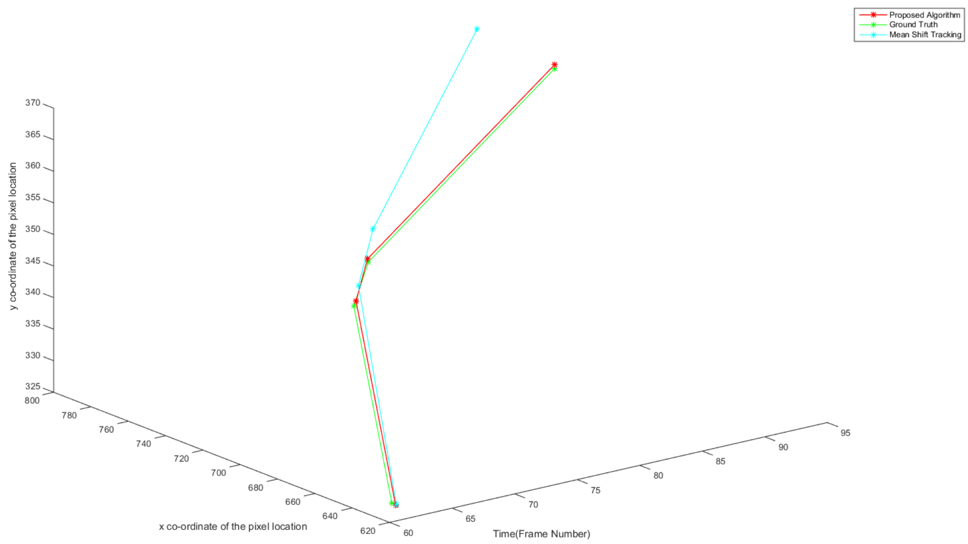

Figure 3.4: Graphical plot indicating sudden direction change of the ball.

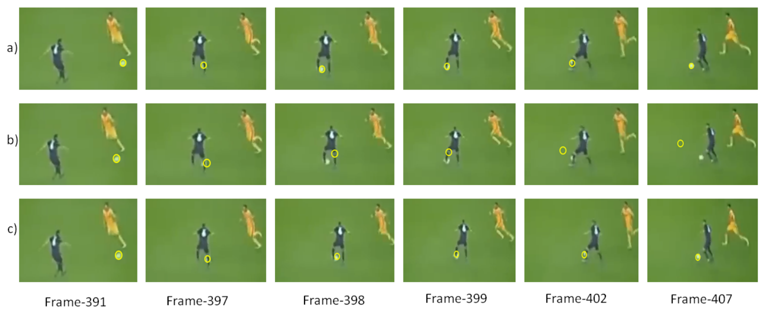

Figure 3.5: Performance evaluation when partial occlusion occurs, (a) Ground truth result, (b) Result of mean shift algorithm based on color histogram matching, (c) Result of our football tracking algorithm

maxima of a density function given discrete data sampled from that function. It is useful for detecting the modes of this density. So when the ball is in touch with a player with white shoe then it is unable to track the ball based on the mode of color based probability distribution function. In Fig.3.3(c) we run our football tracking algorithm which perfectly tracks the ball though it suddenly changes its direction with high speed. It has the ability to track the ball when it changes its direction suddenly as we give a weight on the shape of the ball.

In Fig. 3.5(a) We take frame 391 to 407 as our experiment frame of dataset [19]. In these frames the football gets partial occlusion by player. In Fig. 3.5(b) we are showing our ground truth result where we manually label the ball with yellow circle. In Fig. 3.5(b) we run Meanshift algorithm based on color histogram [10] which fails to track the ball when partial occlusion occurs. This is because for partial occlusion number of pixels on the ball is too low to track the ball based on color component. In Fig. 3.5(c) we run our football tracking algorithm which perfectly tracks the ball though partial occlusion occurs. This is because we give attention to the size and shape of ball and we assign weight based on size

Figure 3.6: Performance evaluation when short time full occlusion occurs, (a) Ground truth result, (b) Result of mean shift algorithm based on color histogram matching, (c) Result of our football tracking algorithm

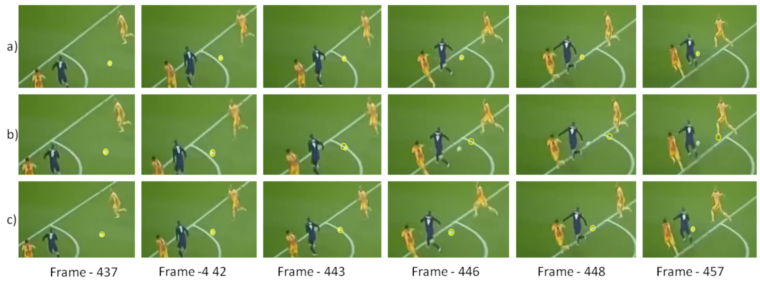

Figure 3.7: Performance evaluation when there is other white objects, (a) Ground truth result, (b) Result of mean shift algorithm based on color histogram matching, (c) Result of our football tracking algorithm

and shape of the ball.

In Fig. 3.6(a) We take frame 420 to 425 as our experiment frame of dataset [19]. In these frames the football gets short time full occlusion by player. In Fig. 3.5(b) we are showing our ground truth result where we manually label the ball with yellow circle. In Fig. 3.6(b) we run Meanshift algorithm based on color histogram [10] which fails to track the ball when partial occlusion occurs. This is because for full occlusion we can't get ball's color pixels. Then it is not possible to keep track the motion of ball. So once color based algorithm fails to track the ball the error occurs in next frames and it can't be rectified. In Fig. 3.5(c) we run our football tracking algorithm which perfectly tracks the ball though full occlusion occurs. This is because our measurement space captures movement of the ball in any direction. In the measurement space suppose the ball gets full occlusion for a small time. When it reappears in the measurement space then our algorithm can capture the balls based on features which we take.

In Fig. 3.7 We take frame 437 to 457 as our experiment frame of dataset [19]. In these frames the football passes through the white line in the ground. In Fig. 3.7(a) we are showing our ground truth result where we manually label the ball with yellow circle. In Fig. 3.7(b) we run Meanshift algorithm based on color histogram [10] which fails to track the ball passes through the white line in the ground. This is because the color of ball matches with the lines in the ground. In Fig. 3.7(c) we run our football tracking algorithm which perfectly tracks the ball though partial occlusion occurs. This is an example where our algorithm tracks the

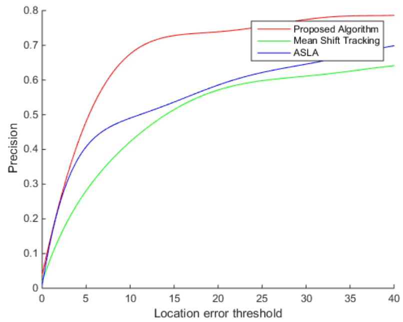

Figure 3.8: Precision Plot.

ball based on shape. Here gradient of edge pixel of ball and white line are almost same due to same background. Based on the shape of the ball we successfully track football when it passes through white line.

3.4 Performance Evaluation

Our main goal is to assess the effectiveness of our proposed algorithm over the state-of-the art algorithms particularly for ball tracking. When an algorithm loses track of the target object, it may resume to track after failure if there exists a re-detection module, or if it locates the target object again as the object reappears at the position where the tracking bounding box is. If we simply average the evaluated values of all frames in an image sequence, the evaluation may not be fair since a tracker may have lost the target in the beginning but could have tracked the target successfully if it were initialized in a object state or frame. Following the above principle, here we use two widely used metric for tracking evaluation.

3.4.1 Precision Plot

One widely used evaluation metric for object tracking is the center location error, which computes the average euclidean distance between the center locations of the tracked targets and the manually labeled ground truth positions of all the frames. When a tracking algorithm loses track of a target object, the output location can be random, and thus, the average error value does not measure the tracking performance correctly. The precision plot addresses this issue by showing the precision, defined as the percentage of frames whose center location error(CLE) are smaller than a threshold, against the CLE threshold. However, the center location error only measures the pixel difference and does not reflect the size and scale of the target object. In Fig 3.8 we have compared our method with color histogram based meanshift algorithm. Our method has shown good performance because we have used color, edge and shape of football respectively as our features. Precision Score at Center Location Error Threshold of 40 also carry significant performance indication. The relatively high CLE, as shown in Fig 3.9, indicates that the tracker doesnot completely lose the track of the ball. In this frame, CLE for our propose algorithm is 38. Thus it can be possible for the

Figure 3.9: Yellow circle is the ground truth at frame 124 of dataset [18]. Blue Circle is the predicted circle by proposed algorithm at that frame yielding CLE of 38.

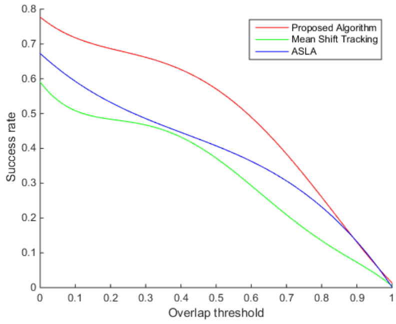

Figure 3.10: Success Plot.

tracker to track precisely at subsequent frames. Whereas at more higher CLE, the tracker completely loses the target, hence start to predict random erroneous location at subsequent frames.

3.4.2 Success Plot

For completeness of the evaluation, the another important metric we use is the Area Under the Curve(AUC) derived from the success plot of tracking algorithms. This is known as **success plot**. Suppose we have a region R_t tracked by our tracking algorithm. We also have ground truth region R_g. By $|R_t|$ we denote the number of pixels in that region. Given R_t and R_g the overlap score is defined as

$$O_{score} = \frac{|R_g \bigcap R_t|}{|R_g \bigcup R_t|}.$$

Then, the success rate of a tracker on a sequence is the percentage of frames whose overlap score O_{score} is larger than a given threshold. By varying the threshold from 0 to 1, one can generate the success plot, and the area under curve(AUC) can be derived afterwards. In Fig.3.10 we have shown our success rate. Our success rate is so high as we have used weight assignment technique of particle filter and we resample coordinates based on the pointwise likelihood ratio. We also take shape and edge of the ball as our features with the color feature which makes our algorithm more accurate than color histogram based mean shift algorithm.

Precision Score (Precision Score at CLE Threshold of 10)						
Video	Number of Frames	Max. Shift between 2 frames	Min. Shift between 2 frames	Proposed Method	Mean Shift Tracker	ASLA
Seq1	54	5.5	0.7071068	0.91	0.7835	0.8532
Seq2	88	9.61	0.3333333	0.8111	0.8272	0.778
Seq3	67	12.77	0.8333333	0.7826	0.696	0.72
Seq4	91	17.160193	0	0.7319	0.7096	0.685
Seq5	85	22.8725435	0.6009252	0.6786	0.602	0.597
Seq6	106	26.44	0.1666667	0.71	0.6111	0.4923
Seq7	100	31.1488363	0.6009252	0.6647	0.602	0.487
Seq8	82	35.0570963	3.0413813	0.6337	0.5715	0.433
Seq9	51	40.5123438	1	0.56	0.4842	0.316

Figure 3.11: Precision Score of different tracker at CLE Threshold of 10.

3.4.3 Accuracy Table

The precision score at Centre Location Error Threshold of 10 of different methods are compared with the proposed algorithm in the table 3.11. The sub-sequences are cropped from the original two dataset [18] and [19]. However the number of frames in each sub-sequences are shown in the second coloumn of the table 3.11.

Chapter 4

Conclusion

The proposed approach uses a probabilistic dynamic feature based approach to track ball in any direction based on likelihood calculation for each sample point. However significant between frame motion and camera movement has also been handled to a certain extend by the construction of dynamic measurement space. Prior weightage on the shape measure of the ball also makes it possible to track the ball successfully upto a extend even in its blurred deshaped position owing to high speed of the ball.

During the course of work, a variety of optical flow based method has been tried to estimate the speed of the ball from the variation of intensity distribution. But all of these methods fail to produce any competitive result due to significant between frame motion. From that point of view, any spatial variation of the features have not been utilized. However it always provide a scope of work to implement it. This will lead to correction of optical flow vector under the constraint of moving camera and thus can create a more robust tracking model for small object tracking.

Bibliography

[1] Moeslund, T.B., Thomas, G., Hilton, A.: Computer Vision in Sports. Springer (2015)

[2] Yilmaz, A., Javed, O., Shah, M.: Object tracking: A survey. Acm computing surveys (CSUR) **38** (2006) 13

[3] Wu, Y., Lim, J., Yang, M.H.: Online object tracking: A benchmark. In: Proceedings of the IEEE conference on computer vision and pattern recognition. (2013) 2411–2418

[4] Yu, X., Sim, C.H., Wang, J.R., Cheong, L.F.: A trajectory-based ball detection and tracking algorithm in broadcast tennis video. In: Image Processing, 2004. ICIP'04. 2004 International Conference on. Volume 2., IEEE (2004) 1049–1052

[5] Huang, Y., Llach, J., Zhang, C.: A method of small object detection and tracking based on particle filters. In: Pattern Recognition, 2008. ICPR 2008. 19th International Conference on, IEEE (2008) 1–4

[6] Xing, J., Ai, H., Liu, L., Lao, S.: Multiple player tracking in sports video: A dual-mode two-way bayesian inference approach with progressive observation modeling. Image Processing, IEEE Transactions on **20** (2011) 1652–1667

[7] Seo, Y., Choi, S., Kim, H., Hong, K.S.: Where are the ball and players? soccer game analysis with color-based tracking and image mosaick. In: International Conference on Image Analysis and Processing, Springer (1997) 196–203

[8] Yamada, A., Shirai, Y., Miura, J.: Tracking players and a ball in video image sequence and estimating camera parameters for 3d interpretation of soccer games. In: Pattern Recognition, 2002. Proceedings. 16th International Conference on. Volume 1., IEEE (2002) 303–306

[9] Ohno, Y., Miurs, J., Shirai, Y.: Tracking players and a ball in soccer games. In: Multisensor Fusion and Integration for Intelligent Systems, 1999. MFI'99. Proceedings. 1999 IEEE/SICE/RSJ International Conference on, IEEE (1999) 147–152

[10] Comaniciu, D., Ramesh, V., Meer, P.: Real-time tracking of non-rigid objects using mean shift. In: Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on. Volume 2., IEEE (2000) 142–149

[11] D'Orazio, T., Guaragnella, C., Leo, M., Distante, A.: A new algorithm for ball recognition using circle hough transform and neural classifier. Pattern Recognition **37** (2004) 393–408

[12] Rathi, Y., Vaswani, N., Tannenbaum, A., Yezzi, A.: Tracking deforming objects using particle filtering for geometric active contours. IEEE transactions on pattern analysis and machine intelligence **29** (2007) 1470–1475

[13] Jia, X., Lu, H., Yang, M.H.: Visual tracking via adaptive structural local sparse appearance model. In: Computer vision and pattern recognition (CVPR), 2012 IEEE Conference on, IEEE (2012) 1822–1829

[14] Dryden, I.L., Mardia, K.V.: Statistical shape analysis. Volume 4. J. Wiley Chichester (1998)

[15] Pérez, P., Blake, A., Gangnet, M.: Jetstream: Probabilistic contour extraction with particles. In: Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. Volume 2., IEEE (2001) 524–531

[16] Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T.: A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. Signal Processing, IEEE Transactions on **50** (2002) 174–188

[17] Doucet, A., Godsill, S., Andrieu, C.: On sequential monte carlo sampling methods for bayesian filtering. Statistics and computing **10** (2000) 197–208

[18] : Dataset. `https://www.youtube.com/watch?v=UBJG3FU4DGM` (2016)

[19] : Dataset1. `https://www.youtube.com/watch?v=Wicq4QTlUCk` (2016)