# TRACKING THE TRANSLATIONAL AND ROTATIONAL MOVEMENT OF THE BALL USING HIGH-SPEED CAMERA MOVIES

*Hubert Shum and Taku Komura*

City University of Hong Kong
Tat Chee Avenue, Kowloon, Hong Kong SAR
*hubert@cityu.edu.hk, itaku@cityu.edu.hk*

**ABSTRACT**

Skills to spin the ball are important for athletes in many sports such as baseball, soccer and tennis. Recently, researchers in aerodynamics and sports science started to analyze the correlation of the rotation and the trajectory of the ball; high speed cameras are used to analyze the rotation of the ball. However, the analysis had to be done either manually, or under special lighting condition. In this paper, we propose a new algorithm to track the translation and rotation of the ball shot by high speed cameras under outdoor ambient light condition. Using our system, the athletes can analyze their skills by shooting their motion and the resultant rotation of the ball. It can also be used as a tool to enhance sports TV programs by providing further scientific information to the audience.

## 1. INTRODUCTION

Computer vision techniques are often used today for sports analysis. For sports that use balls, the analysis of the movement of the ball is important for athletes to learn how they can improve their skills. A lot of research has been done to track the ball in the field using TV quality movies [1] [2] [3]. Yu et al.[1] successfully tracked the soccer ball in the field using the information of the whole trajectory. The results were stable comparing to methods based on frame-to-frame tracking. Guèziec [2] tracked the baseball in TV program videos using Kalman Filter and reconstructed the 3D trajectory of the ball in virtual space. Shum et al.[3] did a similar work based on single-view camera information; aerodynamics were used to estimate the depth of the ball. The detection of the translation of the ball was the main problem to be solved in these researches. The athletes will be able to analyze their skills even further if they can know about the rotation of the ball. For example, in baseball, the pitchers control their arms skillfully to change the spinning pattern of the ball, inducing it to curve to different directions. The curvature of the trajectory of the ball and its spinning pattern has a strong correlation, and therefore, whether they can pitch good curve balls depend on how skillful they are in spinning the ball. With the help of high quality equipments, researchers have started to analyze the rotation of the ball. Alaways [4] analyzed the angular velocity of a baseball using a high speed camera and precisely explained the physics of baseball. Theobalt et al. [5] used multi-exposure images of high quality camera to track the motion of the baseball under a special lighting condition. However, for analysis of pitchers, it is preferable that the method can be used outdoor under ambient light condition, which will broaden the field of analysis. In this paper, we propose a new method to track and analyze the movement of flying baseballs using video movies shot by high speed cameras. We assume the ball is shot under an outdoor, ambient light environment. Using our system, the athletes can analyze their skills by shooting their motion and the resultant rotation of the ball. Our system can also be used as a tool to enhance sports TV programs by providing further scientific information to the audience.

## 2. METHODOLOGY

### 2.1. Overview

The system is composed of two modules: the ball tracking module and the rotational analysis module. The former uses a searching window method to track the correct position of the baseball, while the latter uses the extracted baseball texture to calculate the angles of rotation.

### 2.2. Window Based Tracking

In order to analyze the ball, it is necessary to extract it from the scene. This can be done by subtracting the background from the image. The background image has to be either prepared in advance, or has to be calculated from the video. This image is then matched with the scene and subtracted from the image. In case the background image has to be created from the video, it is generated based on the statistics of the pixel color [6]. Since the foreground may contain

**Fig. 1**. The background image calculated from the video (a), and the extracted foreground (b).

noise according to video quality, a morphological closing filter is applied to reduce the noise. In Figure 1, an example of calculating the foreground image is shown.

The ball can now be cut out from the scene by fitting a rectangular window into the foreground image. The position of the ball is manually specified in the initial frame. In the following frames, it can be tracked throughout the movie by searching the region the ball appeared in the previous frame. As we assume perspective camera view, the size of the ball in the scene gradually changes. Therefore, the size of the window has to be changed dynamically. The ball found by foreground extraction may miss some of the ball pixels that contain color similar to the background. Therefore, we need to rectify the searching window before the final segmentation process. In most of the cases, some area of the ball is relatively easier to be distinguished from the background. Generally, the face of the ball reflecting the light appears more clearly, and therefore the boundary can be extracted more accurately. As the ball appears brighter on the side the light comes from, the direction of the light can be estimated by evaluating the luminance over the ball region. An example of an extracted ball is shown in Figure 2. The arrow indicates the direction of the light. The boundary of the ball closer to the light source can be extracted more precisely; in case of Figure 2, the upper, right boundary of the searching window is more reliable. The searching window will be placed on the reliable side. The size of the ball will be calculated based the distance between the reliable pair of parallel edges of the searching window. In case of Figure 2, the size of the ball is the distance between left and right side.

However, the size and position of the ball may jitter around the correct value when processing the whole movie, due to the noise at the foreground extraction stage. Therefore, we used Kalman filter [2] to stabilize the results. The general form of the Kalman time-update equation is:

$$\hat{x}_k^- = Ax_{k-1} + Bu_k \tag{1}$$

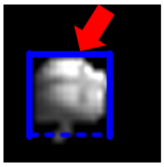

**Fig. 2**. Finding reliable sides in searching window

where $x_{k-1}$ is the state of the system in time $k-1$, $\hat{x}_k^-$ is the priori estimate of the state at time $k$, and $u_k$ is the control input. The Kalman measurement-update equation can be written by the following form:

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \tag{2}$$

where $\hat{x}_k$ is the state vector used for the result, $z_k$ is the measured data of the state, $K_k$ is the Kalman gain, and $H$ is a constant that correlates the state $x_k$ to the measured value $z_k$. In our system, we use a $R^3$ space for the position $(x, y)$ and the size $(s)$ of the ball, i.e. $x \in R^3$. Since we do not have any control input, $u_k$ is considered zero. Using the Kalman filter, the effect of noise in the searching window can be minimized.

### 2.3. Rotation Parameters Estimation

The purpose of this research is to track the translation and rotation of the ball. The rotation is estimated by matching the texture pattern of the ball in consecutive frames. Before matching the texture pattern, the image must be preprocessed and the effect of lighting has to be cancelled. This is to avoid the lighting condition to affect the results. To subtract the lighting effect from the ball, we have to know the lighting condition of the environment. However, this information will not be provided in most cases, and it is difficult to estimate it only from the luminance of the ball. One simple and practical method to remove the lighting effect is to subtract the average luminance of each pixel throughout the movie. This process is shown on Figure 3. The segmented image of the ball from the video is shown on the left, the average luminance of each pixel in the ball region throughout the video is shown on the middle, and result of subtracting the average luminance from the original image is shown on the right. The right-most image is used for tracking the rotation of the ball.

Next, the image texture is parallelly mapped onto a sphere to estimate the axis and angle of rotation. The radius of the sphere is set same as that of the segmented circle. The rotation is represented by three Euler angles, which represents consecutive rotations around the x, y and z axis. Therefore, the problem here is to find out the value of the Euler angles. Each coordinate on the surface of the sphere will be transferred to another point on the sphere by applying the

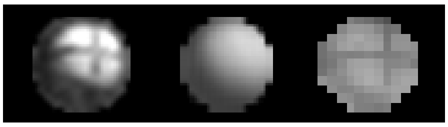

**Fig. 3**. The segmented image of the ball from the video (left), the average intensity of the ball region throughout the video (middle), and result of subtracting the average intensity from the original image (right)

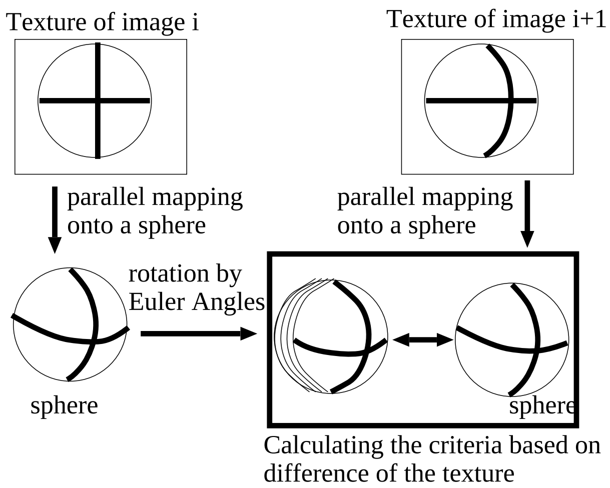


**Fig. 4**. The process of estimating the rotation of the ball

following transformation:

$$p' = R_x R_y R_z p \quad (3)$$

where $R_x$, $R_y$ and $R_z$ represent 3D rotational transformation around the $x$, $y$ and $z$ axis, and $p$ and $p'$ represent the point on the sphere before and after the rotation, respectively. The Euler angles can be calculated by minimizing the following criteria:

$$\sum_{x,y \in S_i} |T_{i+1}(x', y') - T_i(x, y)| / a(S_i) \quad (4)$$

where $S_i$ is the surface area of the ball in frame $i$ and $a(S_i)$ is its area projected onto a plane vertical to the z axis, $(x, y)$ and $(x', y')$ are the x and y coordinates of $p$ and $p'$. The flow chart of this process is shown in Figure 4. Note that because we only have the texture information from one side of the ball, some part of the ball which was originally hidden in frame $i$ will appear in frame $i + 1$, and some area visible in the frame $i$ will be hidden in frame $i + 1$. Such area are excluded from $S_i$ when calculating the criteria.

Since a high speed camera can capture 1000 frames per second, the rotation between two consecutive frames is very small. As a result, the effect of the noise becomes significant. Therefore, instead of estimating the rotation parameters from only two consecutive frames, we assume the axis

| | **soccer ball** | **baseball** | **golf ball** |
|---|---|---|---|
| **S.D.** | 1.0053 | 1.6680 | 1.0049 |

**Table 1**. The average standard deviation of the Euler angles of the balls in the artifically generated movies.

**Fig. 5**. The ball textures used for evaluation of the system

is constant for $N$ frames, and estimate them by minimizing the following criteria:

$$\sum_{i=M}^{M+N} \sum_{x_i, y_i \in S_i} |T_{i+1}(x'_{i+1}, y'_{i+1}) - T_i(x_i, y_i)| / a(S_i) \quad (5)$$

where $M$ is the current frame, and $(x'_{i+1}, y'_{i+1})$ and $(x_i, y_i)$ are the x and y coordinates of $p'_{i+1}$ and $p_i$, respectively. It is known that, in most of the cases, the axis of rotation of the ball stays almost constant from the moment of release until it reaches the catcher. Therefore, this approach is acceptable if $N$ is much smaller than 1000, which is the number of frames until the ball reaches the catcher. The Euler angles were searched by exhaustive search by changing their values from -5 to 5 degrees. We have only checked the integer values here to reduce the computational cost.

## 3. EXPERIMENT

In order to evaluate the accuracy of the system, we first applied the proposed method to artificial generated movies of various balls with known rotation. Different textures were mapped onto spheres to simulate the scenario of different ball games. A 3D modeling software was used to map different textures onto a sphere, and movies of flying balls were generated using key frame animation. The translation and rotation of the balls were automatically calculated using our system. The rotation parameters were correctly calculated; however, there were rotation error around the correct value. The average standard deviations of the obtained Euler angles for the three balls are shown in Table 1. It can be observed that the result of using the baseball texture is much worse than those by the soccer ball and golf ball texture. This is because the texture of the baseball is narrow and sparse. The other two balls have thick and clear texture comparing to the baseball, which enables the system to obtain correct results in a stable manner.

Next, the proposed algorithm was used to calculate the translation and rotation of a baseball pitched by an amateur baseball player. A high speed video camera system *Phan-*

**Fig. 6**. The image sequence of a pitcher captured using the high speed camera system

| | **x** | **y** | **z** |
|---|---|---|---|
| $N$=**5** | 1.2155 | 1.2045 | 0.8690 |
| $N$=**30** | 0.6072 | 0.6570 | 0.4416 |

**Table 2**. The standard derivation of the Euler angles around the x, y and z axis when N=5 and N=30

*tom,* by Vision Research Inc. was used to shoot the trajectory of the ball at the rate of 1000 frames per second. Figure 6 shows some sample frames of the input video. Two sorts of pitches including the straight ball and the curve ball were analyzed. A cross pattern texture was drawn on the ball, as shown in Figure 5, to assist the pattern matching algorithm. Different $N$ values were tested to calculate the rotational parameters. The standard deviations of the Euler angles when $N$=5 and $N$=30 are listed in Table 2. When $N$ is increased, the results of the optimization become more stable.

Figure 7 shows the reconstructed 3D trajectory of the straight (a) and curve (b) ball from the catcher's side, together with the visualized information of rotation. It is easy to find out the trajectory of the straight ball is rectilinear, while that of the curve is parabolic. The average angular velocity of each pitch is shown in Table 3, where y-axis represents vertical axis, the z-axis is parallel to the depth vector from the pitcher toward the catcher, and the x-axis is the vector from the left to the right (see the left side of Figure 7). The rotational axis of the straight ball is rather perpendicular, where that of the curve ball is diagonal to the horizontal axis. These results match well with the theoretical data [4].

## 4. CONCLUSTION

In this paper, a new method to track the translational and rotational motion of a ball was proposed. The movie is assumed to be shot by a high speed camera in an outdoor environment. The effects of the ambient light were subtracted from the segmented area, and the rotational parameters were calculated using a spatiotemporal approach. The calculated results matched well with theoretical values and the system has also shown good performance when applying the method to computer generated movies. A sports analysis system based on the proposed method, which is to support baseball players, is now being developed. Such system will help the baseball players to find out their problems to spin the ball skillfully; they can watch the high speed camera videos together with the output of this system to get ideas how to improve their pitching skills. The system is also applicable to different sports such as table tennis, tennis, and soccer. The system can also be used as a tool to enhance sports TV programs by providing further scientific information to the audience.

| | **x** | **y** | **z** |
|---|---|---|---|
| **straight** | -71.897 | -25.481 | 28.363 |
| **curve** | 22.154 | 17.811 | 12.129 |

**Table 3**. The angular velocity (rad/s) of a straight and curve ball

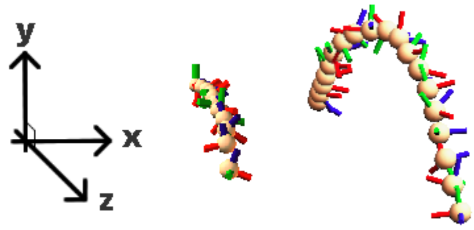


**Fig. 7**. Reconstructed trajectory of the pitched baseball

## 5. REFERENCES


[1] X.Yu, C.Xu, H.W.Leong, Q.Tian, Q.Tang, K.W.Wan, "Trajectory-Based Ball Detection and Tracking with Applications to Semantic Analysis of Broadcast Soccer Video", *ACM Multimedia 2003*

[2] A. Guèziec, "Tracking a Baseball for Broadcast Television", *ACM Siggrapph 2003 Courses*

[3] H.Shum, T Komura, "A Spatiotemporal Approach to Extract the 3D Trajectory of the Baseball From a Single View Video Clip", *ICME 2004*

[4] L.W. Alaways, "Aerodynamics of Curve-Ball - An Investigation of the Effects of Angular Velocity on Baseball Trajectories", *PhD Thesis. University of California*

[5] C.Theobalt, I.Albrecht, J.Haber, M.Magnor, H-P Seidel, "Pitching a Baseball - Tracking High-Speed Motion with Multi-Exposure Images", *SIGGRAPH 2004*

[6] D.Farin, P.N. de With, W.Effelsberg, "Video-Object Segmentation using Multi-Sprite Background Subtraction", *ICME 2004*