Useful Information Extraction and Providing System from Video of Tennis Match

Shun Kitahara
Graduate School of Engineering, Tokai University
4-1-1 Kitakaname, Hiratsuka, Kanagawa, 259-1292 Japan
2bdrm008@mail.tokai-u.jp

Osamu Uchida
Dept. Human and Information Science, School of Information Science and Technology,
Tokai University
4-1-1 Kitakaname, Hiratsuka, Kanagawa, 259-1292 Japan
o-uhida@tokai.ac.jp

Abstract: - In recent years, in sports such as football, baseball, volleyball, and tennis, quantifying various data including the movement of the ball and players is utilized to improve the players' skills and tactics of the match. Equipment that can record various types of sports data automatically exists; however, the equipment is expensive in general and not everyone can use them with ease. In this paper, we focus on tennis and propose a method to analyze the content of the match from videos of the match taken by commonly-marketed digital video camera automatically, and to provide players with useful information comprehensively.

Key-Words: - Tennis, Video image analysis, Sports data analysis, User interface

1 Introduction
In recent years, in sports such as baseball, football, volleyball and tennis, quantifying various data including the movement of the ball and players is utilized to improve the players' skills and tactics of the match. However, in many cases, aggregation of these data is still manual. Therefore, much time and effort is needed to record the details of a sports event. There exist equipment that record various sports data automatically; however, they are expensive in general and not everyone can use them with ease. In this study, we focus on tennis. The purpose of this study is to record various data from tennis match video, to provide the player with useful information extracted from recorded data comprehensively, and to implement the system inexpensively to be available for anyone.

In the previous study, we proposed a system to extract the players' coordinate, the bounce points of the ball, and the stroke points from tennis match videos automatically [1]. However, it is required that both the start and the end time of each point (the sub division of the match between the time when one of the players serves the ball and the time when one of the players gets a point) are inputted into the system manually. It is a burden on users, then in this paper we propose a method to recognize the scene of serving the ball automatically in order that the start time of each point will be inputted into the system automatically. Moreover, we propose a user interface that provides the player with useful information by visualizing various data extracted from recorded data comprehensively.

2 Existing Study
There have been several studies on sports data analysis using video image processing in the past [2]-[7]. For example, D. Connaghan proposed a method to detect the fame and the change of ends of a tennis match by using nine network cameras [2]. Nishita et al. proposed a system to improve tactics for doubles tennis matches [3]. The purpose of this study is to consider a method to record more detailed information from a video of tennis match and provide the player with analyzing results.

3 Proposed Method

3.1 Overview
In the proposed method, two tennis videos taken from an angle such as Fig. 1 for each player by using commonly-marketed digital video cameras are inputted to the system. Then, the ball and players' positions are extracted and record the time of serve action, bounce and stroke positions, and players' movement for each point. The data collected by the
above scheme is provided to users through an application which consists of two Us.

### 3.2 Extraction of the ball

The coordinates of the tennis ball is extracted by using a similar method of Ref. [1]. In the proposed method, however, to accerate the processing time, to determine the most probable region of the ball from candidate regions, the following way is used instead of a template matching used in Ref. [1]. To determine the most probable region of the ball, we calculate the Euclidean distance between the candidate regions of the ball and the reference value determined beforehand. The region the distance value of which is smallest is determined as the ball region and the center of it is decided as the ball position. In the verification experiments described later, the reference value were set to $R = 230$, $G = 230$, and $B = 150$, respectively.

### 3.3 Estimation of the coordinates of bounces and strokes

By using the method proposed by Kitahara and Uchida [1], the coordinates of bounces and strokes are estimated from the coordinates of the ball obtained by the method stated in Section 3.2. To find positions of bounces or strokes, we calculate the angle formed by two vectors, the one is the vector which connects the positions of the ball on frames $T$ and $T - 1$, and the other is the vector which connects the positions of the ball on frames $T$ and $T + 1$. If the angle is smaller than the threshold value, the position can be considered as the rapidly changing point, and the position is considered as a candidate for a position of a bounce or a stroke.

### 3.4 Decision of the coordinates of the player's foot

By using the method proposed by Kitahara and Uchida [1], the coordinate of player’s foot is decided. By using a particle filter, the region of the player is extracted and tracked. Next, the skin color regions are extracted from the region of the player. Then, the foot coordinates of the player is decided in the way described in Ref [1].

### 3.5 Extraction of the scene of serving the ball

In order that the start time of each point will be inputted into the system automatically, the scene of serving the ball is extracted. Fig. 2 shows an example of the movements to serve the ball. A serve is an action that the player shot the ball into the opponent's court firstly in each point. Moreover, a toss is an action of throwing the ball straight up to serve the ball. In this study, we extract the scene of serving the ball by combining the extraction of the toss action and the detection of the rapidly ball changing point.

Extraction of the scene of serving the ball is done based on the angle of moving vectors of the ball. We define the vector which connects the positions of the ball on frames $T$ and $T + 1$ as the moving vector at the frame $T$. The angle of the moving vector at the $T$ frame is computed by

\[
\theta = \begin{cases} 
360 - A & \text{if } y_1 - y_0 > 0, \\
A & \text{otherwise}, 
\end{cases}
\]

where

\[
A = \cos^{-1} \frac{x_1 - x_0}{\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}},
\]

and $(x_0, y_0)$ and $(x_1, y_1)$ denote the coordinates of the ball at the $T$ and the $T + 1$ frames, respectively. If the number of moving vectors the angle of which is between 45 and 135 degree is greater than or equal to 8 at the frame between $T$ and $T - 10$, the time of the
frame $T$ is decided that the time when a toss action occurs. If there exists a rapidly changing point less than or equal to 40 frames from the time when a toss is detected, the point and its time are decided as the serve hitting.

### 3.6 Extraction of characteristics of the play
Useful information for the player is extracted by using the coordinates of the player, bounce, and stroke estimated by the method stated above.

(a) **Projective transformation**
The coordinates of the player and the ball are mapped to on a two-dimensional coat image viewed from directly above (Fig. 3). The blue circles and the red circle mean the coordinates of the players and the coordinate of the stroke, respectively. To derive the homography matrix of the projective transformation, four coordinates of the tennis court are inputted manually.

(b) **Movement distance of the players**
The movement distance of the player is derived by using the data of coordinates of the player. Movement distance of $i$-th point denoted by $D_i$ is calculated by

$$D_i = \sum_{j=1}^{N_i-1} \sqrt{(x_j^i - x_{j+1}^i)^2 + (y_j^i - y_{j+1}^i)^2}$$  \hspace{1cm} (2)

where $(x_j^i, y_j^i)$ denotes the coordinate of the player on $j$-th frame of $i$-th point, and $N_i$ denotes the total number of frames on $i$-th point.

(c) **Number of the strokes in the rally**
A rally is a sequence of shots within a point, and the number of the strokes in $i$-th rally denoted by $R_i$ is calculated by

$$R_i = \left\lceil \frac{h_i}{2} \right\rceil$$  \hspace{1cm} (3)

where $h_i$ is the total number of the strokes extracted in the $i$-th point.

(d) **Numbers of the forehand and the backhand strokes**
The numbers of the forehand and the backhand strokes for each player are obtained. The system can judge whether the stroke is forehand or backhand by using the positions of the ball and the player (the player’s dominant arm is manually inputted to the system in advance).

(e) **Average speed of the shots**
The average speed of the shots for each player is obtained. Average speed of the shots in $i$-th point denoted by $V_i$ is calculated by

$$V_i = \frac{1}{M_i} \sum_{k=1}^{M_i} \frac{d_k^i}{f_k^i}$$  \hspace{1cm} (4)

where $M_i$ denotes the total stroke number of the player in $i$-th point, $d_k^i$ denotes the distance from the position of $k$-th stroke and the following bounce position in $i$-th point, and $f_k^i$ denotes the number of frames from the position of $k$-th stroke and the following bounce position in $i$-th point.

### 3.7 Application to provide useful information
We implemented an application to provide useful information. This application consists of two UIs, one is the interface to look at feature trends (Fig. 4), and the other is the interface to browse the movie (Fig. 5). Users can easily look back the match and obtain the hint to improve their skills by using the application. These UIs are implemented by .Net Framework 4.5 and C#.

#### 3.7.1 Interface to look at feature trends
When users to use this interface, users first chose two features (see Table 1). Then, dots corresponded to each point are plotted to the court image based on the features of each point. Fig. 4 is an example in the case where the vertical axis and the horizontal axis are the number of the strokes in the rally and the average speed of the shots of player 1. Only by clicking the dot on the court image, the practical movie correspond to it can be browsed.
3.7.2 Interface to browse the movie
On the screen of the interface, users can browse the practical moves of each player and the duplicated movie of the match viewed from directly above. The duplicated movie is created by the projective transformation which is mentioned in Section 3.6 (it is assumed that the ball flies linearly). The color of the arrowed line used on the duplicated movie to display the trajectory of the ball is changing depending on the speed of the shot. The duplicated movie viewed from directly above is expected to help users to grasp the whole picture of the match easily.

4 Verification Experiment
To evaluate the usefulness of the proposed method, we conducted an experiment to obtain the success rate of the extraction of serve action. We recorded three practice tennis matches of a tennis club at Tokai University (the players of which were experts) using two commercially available video cameras (Panasonic HC-V620) from an angle such as Fig. 1 for each player. We applied the proposed method to randomly selected 50 scenes of serving the ball. The extraction of the serve by the proposed method is identified as successful if the following two conditions are satisfied; (i) the number of frames between the frame judged such that the serve stroke is occurred by the proposed method and the frame labeled as correct manually in advance is less than or equal to 2, (ii) the Euclidean distance between the correct position of the serve stroke and the coordinates that is extracted as the serve stroke

<table>
<thead>
<tr>
<th>Table 1 Selectable features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Movement distance of player 1 or player 2</td>
</tr>
<tr>
<td>Average speed of the shots of player 1 or player 2</td>
</tr>
<tr>
<td>Number of forehand or backhand strokes of player 1 or player 2</td>
</tr>
<tr>
<td>Number of the strokes in the rally</td>
</tr>
<tr>
<td>Order of points</td>
</tr>
</tbody>
</table>
position by the proposed method is less than or equal to 10. The success rate obtained by the experiment was 0.72 (= 31/50). Fig. 6 and 7 show successful and failure examples of detecting serve action, respectively.

5 Conclusion
In this paper we have proposed a method to recognize the scene of serving the ball automatically and a user interface that provides the player with useful information by visualizing various data extracted from recorded data comprehensively. The result of the verification experiment shows that the success rate of the extraction of serve action is sufficient for practical use. In the future, we will try to implement the function of providing more tangible advice.

References: