# Tennis Video 2.0 : A New Framework of Sport Video Applications

Jui-Hsin Lai and Shao-Yi Chien
Media IC and System Lab
Graduate Institute of Electronics Engineering and Department of Electrical Engineering
National Taiwan University
BL-421, 1, Sec. 4, Roosevelt Rd., Taipei 106, Taiwan
larry@media.ee.ntu.edu.tw, sychien@cc.ee.ntu.edu.tw

## ABSTRACT

This video demo presents a new framework of sport video applications called as Tennis Video 2.0. The proposed information extraction scheme retrieves the temporal structure of a video and separates the video foreground and background objects into different layers. With the structure and layer information, the new multimedia is generated. Contrary to the conventional video contents, the proposed new multimedia enables users to generate their own contents and feedback requests to the video players for more interaction. Users even can share their created contents with friends in different transmission bandwidth with considering the semantics.

## Categories and Subject Descriptors

I.4.8 [**Scene Analysis**]: Object recognition, tracking; I.4.9 [**Applications**]; H.5.2 [**User Interfaces**]: Interaction styles

## General Terms

Algorithms, design, experimentation, performance

## Keywords

Tennis video, sport video application, video analysis

## 1. INTRODUCTION

Tennis Video 2.0 is a new framework of sport video applications. It contains an integral and efficient information extraction scheme, and it also includes the presentation of the new multimedia, which are shown as the ÒInformation ExtractionÓ and ÒNew MultimediaÓ parts, respectively, in Figure 1. Information Extraction part is the proposed video analysis scheme to extract structure, score, field, camera, player, and ball information. With structure and score information, the semantic temporal structures of sport video contents are discovered, where a tennis game can be separated into sets, games, and plays. Next, with sprite, which is an image constructed from information of a video object in all frames of a video sequence, the semantic layers of sport video contents are also found, that is, the video objects of field, players, and balls are separated, and the camera information is also extracted [1]. After that, a New Multimedia is generated with these structure and layer information, which is a new structural, interactive, and scalable media. When playing back the proposed new sport video, unlike conventional video contents, the viewers can create their own sport video contents to be presented. For structural media, viewers can browse game structure and watch highlights immediately. For interactive media, viewers use ÒSpotlightÓ function to watch games with multiple players in contiguous motion or use ÒStrategyÓ function to synchronously display players and ball positions on the map for playing strategy discussion. For scalable media, viewers can adaptively transmit video content to others under different bandwidth limitation with considering the semantics. For example, we can share the highlights but not a complete game with friends in low bandwidth.


*MM'07,* September 23–28, 2007, Augsburg, Bavaria, Germany.
ACM 978-1-59593-701-8/07/0009.

## 2. INFORMATION EXTRACTION

In our observation, tennis videos are regular and well structured. They are usually composed of many Basic Unit Shot Sets (BUSSs), as shown in Figure 2. Each BUSS usually presents an event in tennis video. It begins from a key shot, which is the rally shot in tennis video, and ends before the next key shot. To obtain video temporal structure and layer information, the input tennis video is decomposed into BUSSs as the first step. Score information is then extracted from the score box caption in each key shot. After that, sprite mapping generates field and camera information based on warping video frames of key shots into sprite plane, which is the bridge from temporal structured video to layered video [1]. Next, utilizing field and camera information, player and ball information are extracted.

## 3. PRESENTATION OF NEW MULTIMEDIA

After information extraction, these information are used to generate the new multimedia. Figure 3 is the user interface to present the concept of tennis video 2.0.

For **Structural Media** :

**Content** lists all events of the tennis video in the window. Viewers utilize it to browse the game on-demand and watch highlights by double click event items as shown at the upper right part in Figure 3. In this work, the events of volley, drop shot, ace, fault, double fault, and rally would be annotated.

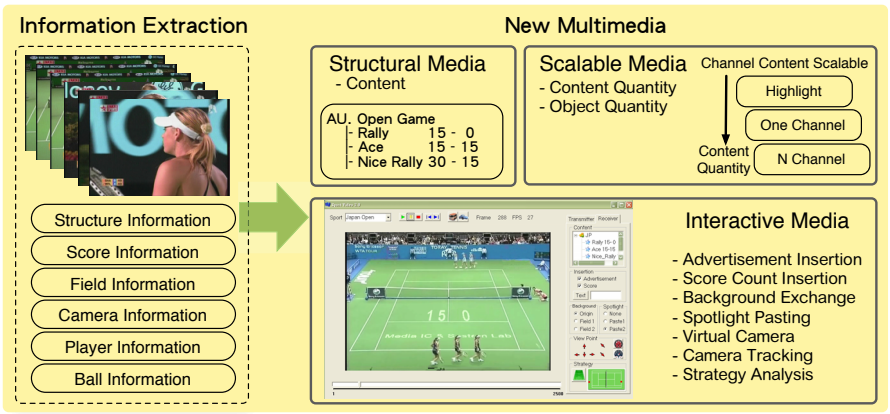


**Figure 1: Tennis Video 2.0 contains Information Extraction and New Multimedia.**

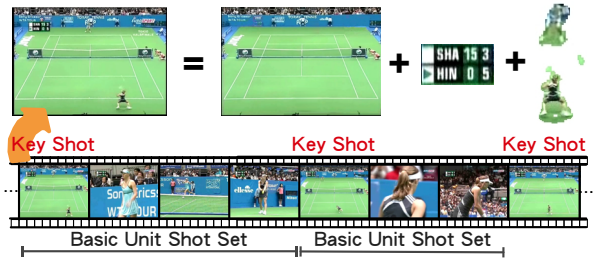


**Figure 2: The concept of Information Extraction.**

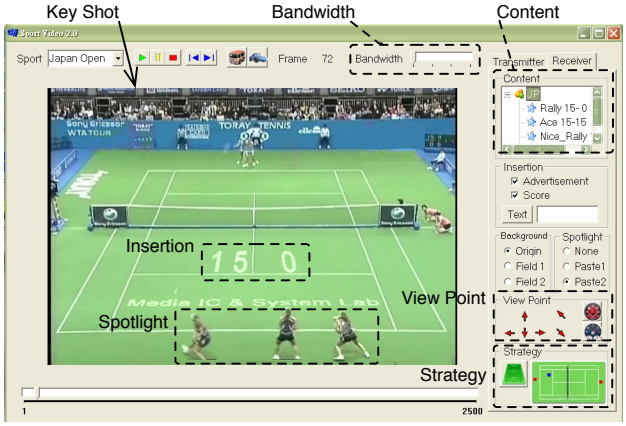


**Figure 3: Tennis Video 2.0 user interface.**

For **Interactive Media** :
**Insertion** enables the content provider or content player to render advertisement or score count on the field without disturbing game proceeding. With score count insertion, viewers can watch full screen game without score box obstructing the window.
**Spotlight** enables viewers to watch games with multiple players in contiguous motion and enriches user experiences, as shown in Figure 3. When watching the highlights, viewers can analyze the postures of players in slow motion.
**View Point** enables viewers to change the camera parameters for virtual camera tracking. Viewers use it to track player's positions and watch the players in close look.
**Strategy** synchronously displays players and ball positions on the map, and it also statistically records the trajectories, as shown at the bottom right of Figure 3. Viewers can discuss the winning strategies of players or analyze trajectories by the small map window.
For **Scalable Media** :
**Bandwidth**, which can be controlled with the bar at the upper part in Figure 3, simulates video content transmission in different bandwidth with considering the semantics. In high bandwidth, all shots of the video will be transmitted, and the receiver gets the most information to know the game. In median bandwidth, only key shot includes audience, player, and ball objects will be transmitted. In low bandwidth, the key shot includes player and ball objects will be transmitted. In very low bandwidth, the text information is transmitted to show the strategy and illustrate what happened in the game.

## 4. CONCLUSIONS

Tennis Video 2.0 is a new framework of sport video applications. Unlike conventional video contents, the viewers of Tennis Video 2.0 can create their own sport video contents to be presented and even share the semantic contents with others. Furthermore, the proposed Information Extraction is an integral and efficient framework to obtain video temporal structure and layer information. With these extraction information, we produce the New Multimedia with more interactivity, functionality, and scalability from signal level to semantic level. Tennis Video 2.0 is implemented and shows the ability to provide various usages and applications to enrich the viewers' experiences.

## 5. REFERENCE

[1] C.-Y. Chen, S.-Y. Chien, Y.-H. Chen, Y.-W. Huang, and L.-G. Chen. Unsupervised object-based sprite coding system for tennis sport. In *Proceedings of the 2003 International Conference on Multimedia and Expo, ICME '03.*, volume 1, pages I–337–340, July 2003.