

Autonomous Production of Sport Video Summaries

by Christophe De Vleeschouwer

Video production cost reduction is an ongoing challenge. The FP7 'APIDIS' project (Autonomous Production of Images based on Distributed and Intelligent Sensing) uses computer vision and artificial intelligence to propose a solution to this problem. Distributed analysis and interpretation of a sporting event are used to determine what to show or not to show from a multi-camera stream. This process involves automatic scene analysis, camera viewpoint selection, and generation of summaries through automatic organization of stories. APIDIS provides practical solutions to a wide range of applications, such as personalized access to local sport events on the web or automatic log in of annotations.

Today, individuals and organizations want to access dedicated contents through a personalized service that is able to provide what they are interested in, at a time convenient to them, and through the distribution channel of their choice. To address such demands, cost-effective and autonomous generation of sports team video contents from multi-sensored data becomes essential, ie to generate on-demand football or basket ball match summaries.

APIDIS is a research consortium developing the automatic extraction of intelligent contents from a network of cameras and microphones distributed around a sports ground. Here, intelligence refers to the identification of salient segments within the audiovisual content, using distributed scene analysis algorithms. In a second step, that knowledge is exploited to automate the production and personalize the summary of video contents.

Specifically, salient segments in the raw video content are identified based on player movement analysis and scoreboard monitoring. Player detection and tracking methods

rely on the fusion of the foreground likelihood information computed in each camera view. This overcomes the traditional hurdles associated with single view analysis, such as occlusions, shadows and changing illumination. Scoreboard monitoring provides valuable additional input which assists in identifying the main focal points of the game.

In order to produce semantically meaningful and perceptually comfortable video summaries, based on the extraction of sub-images from the raw content, the APIDIS framework introduces three fundamental concepts: "completeness", "smoothness" and "fineness". These concepts are defined below. Scene analysis algorithms then select temporal segments and corresponding viewpoints in the edited summary as two independent optimization problems according to individual user preferences (eg in terms of preferred player or video access resolution). To illustrate these techniques, we consider a basket-ball game case study, which incorporates some of the latest research outputs of the FP7 APIDIS research project.

Multi-view player detection, recognition, and tracking

The problem of tracking multiple people in cluttered scenes has been extensively studied, mainly because it is common to numerous applications, ranging from sport event reporting to surveillance in public spaces. A typical problem is that all players in a sports team have a very similar appearance. For this reason, we integrate the information provided by multiple views, and focus on a particular subset of methods that do not use color models or shape cues of individual people, but instead rely on the distinction of foreground from background in each individual camera view to infer the ground plane locations occupied by people.

Figure 1 summarizes our proposed method. Once players and referee have been localized, the system has to decide who's who. To achieve this, histogram analysis is performed on the expected body area of each detected person. Histogram peak extraction enables assignment of a team label to each detected player (see bounding boxes around the red and blue teams). Further segmentation and analysis of the

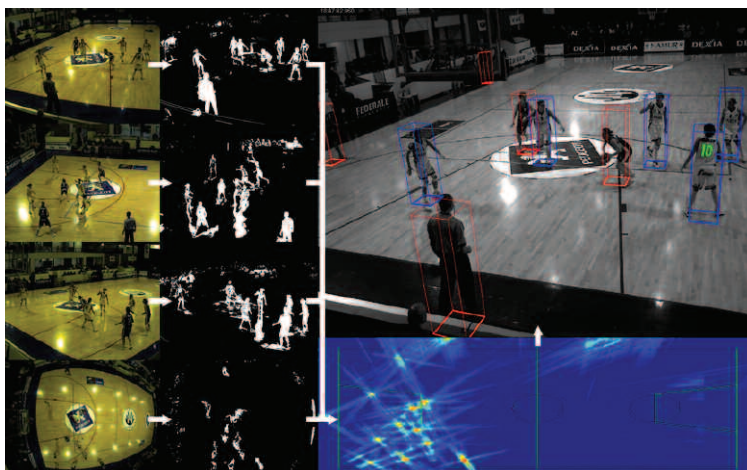


Figure 1: On the left, the foreground likelihoods are extracted from each camera. They are projected to define a ground occupancy map (bottom right in blue) used for player detection and tracking, which in turns supports camera selection.

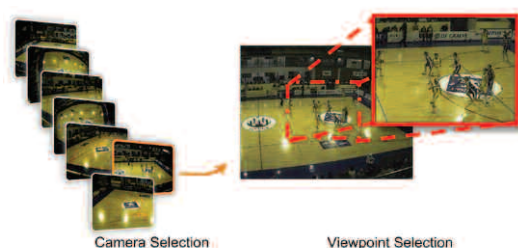


Figure 2: Camera selection and field of view selection.

regions composing the expected body area permits detection and recognition of the digits printed on the players' shirts when they face the camera.

Event recognition

The main events occurring during a basketball game include field goals, violations, fouls, balls out-of-bounds and free-throws. All these events correspond to 'clock-events', ie they cause a stop, start or re-initialization of the 24" clock, and they can occur in periods during which the clock is stopped.

An event tree is built on the basis of the clock and scoreboard information. When needed, this information is completed by visual hints, typically provided as outcomes of the player (and ball) tracking algorithms. For instance, an analysis of the trajectories of the players can assist in decision-making after a start of the 24" clock following a 'rebound after free-throw' or a 'throw-in' event.

Autonomous production of personalized video summaries
To produce condensed video reports of a sporting event, the system selects the temporal segments corresponding to actions that are worth being included in the summary based on three factors:

- Completeness stands for both the integrity of view rendering in camera/viewpoint selection, and that of story-telling in summary.
- Smoothness refers to the graceful displacement of the virtual camera viewpoint, and to the continuous story-telling resulting from the selection of contiguous temporal segments. Preserving smoothness is important to avoid distracting the viewer from the story with abrupt changes of viewpoint.
- Fineness refers to the amount of detail provided about the rendered action. Spatially, it favours close views. Temporally, it implies redundant story-telling, including replays. Increasing the fineness of a video does not only improve the viewing experience, but is also essential in guiding the emotional involvement of viewers through the use of close-up shots.

The ability to personalize the viewing experience through the application of different parameters for each end user was appreciated during the first subjective tests. The tests also revealed that viewers generally prefer the viewpoints selected by the automatic system than those selected by a human producer. This is, no doubt, partly explained by the severe load imposed on the human operator with an increasing number of cameras..

The author thanks the European Commission and the Walloon Region for funding part of this work through the FP7 APIDIS and WIST2 WALCOMO projects, respectively.

Link:

<http://apidis.org/>

Please contact:

Christophe De Vleeschouwer
Université Catholique de Louvain (UCL), Belgium
Tel: +32 1047 2543
E-mail: christophe.devleeschouwer@uclouvain.be

IO LANes: Advancing the Scalability and Performance of I/O Subsystems in Multicore Platforms

by Angelos Bilas

Data storage technology today faces many challenges, including performance inefficiencies, inadequate dependability and integrity guarantees, limited scalability, loss of confidentiality, poor resource sharing, and increased ownership and management costs. Given the importance of both direct-attached and networked storage systems for modern applications, it becomes imperative to address these issues. Multicore CPUs offer the promise of dealing with many of the underlying limitations of today's I/O architectures. However, this requires careful consideration of architectural and systems issues and complex interactions in the I/O stack, all the way from the application to the disk. "IO LANes" (Advancing the Scalability and Performance of I/O Subsystems in Multicore Platforms) is now a EC-funded project led by FORTH-ICS that addresses these issues.

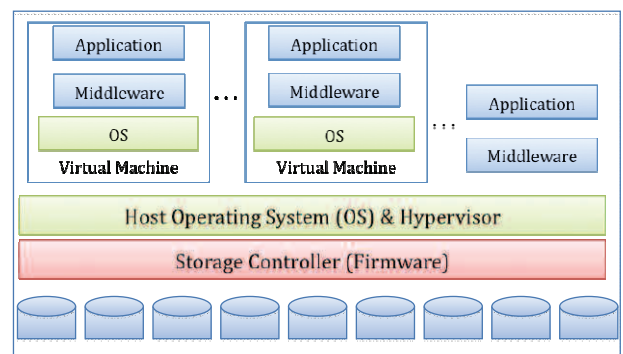


Figure 1: Layers in the I/O path of existing systems.

IO LANes targets three major challenges: (i) dealing with performance and scalability issues of the I/O stack on multicore architectures, (ii) addressing I/O performance and dynamic resource management issues in virtualised, single-host environments, and (iii) examining on-loading and off-loading tradeoffs for advanced functions that are becoming essential in modern storage systems, eg compression, protection, encryption, error correction.

Concept

Our lives are becoming increasingly dependent on electronic information that is processed and stored in computer systems. Individuals, businesses, and organizations cannot survive in today's competitive economy without the use of digital information stored in electronic data storage systems. Data storage is perhaps the most critical and valuable component of today's computing infrastructures. However, crit-