LIVE, an Integrated Production and Feedback System for Intelligent and Interactive TV Broadcasting

Jianmin Jiang[1], Joachim Köhler[2], Carmen Mac Williams[3], Janez Zaletelj[4], Georg Güntner[5], Heike Horstmann[1], Jinchang Ren[1], Jobst Löffler[2], and Ying Weng[1]

***Abstract*—In this paper, we report recent research activities under the integrated project, Live Staging of Media Events (LIVE), which is funded under European Framework-6 programme, and illustrate how a new LIVE TV broadcasting and content production concept can be introduced to improve the existing TV broadcasting services. In comparison with existing TV content production technologies, we show case that LIVE TV broadcasting format could achieve a range of significant advantages, which can be highlighted as: (i) real-time interaction between TV content production team and viewers to ensure the best possible entertainment experiences and to allow viewers not only to view but also to participate, to influence and to control; (ii) the traditional concept of TV content director can be changed to a TV content conductor, where live TV broadcasting can be conducted on real-time bases and existing content materials can be conducted into multi-streams of TV broadcasting, providing variety of choices to allow the audience to make their own preferences in watching TV programmes; (iii) introduction of significant intelligence in terms of content analysis and behaviour analysis etc. to further improve the quality of services based on the fundamental concept of LIVE project. When tested under field trials for 2008 Olympic Games by ORF, the developed LIVE system has illustrated great potential as well as significant promises to improve the TV broadcasting services not only in terms of the interaction formats, but also in terms of viewing experiences as well as the style of content productions and preparations**

***Index Terms*— Interactive TV, Intelligent TV, and semantics-based video processing.**

Manuscript received Sept. 18, 2009. (Write the date on which you submitted your paper for review.) This work was supported in part by the EU IST FP6 Project "Live Staging of Media Events" (LIVE, IST-4-027312).

J. Köhler, H. Horstmann and J. Löffler are with Fraunhofer Institut, IAIS, Germany (email: {joachim.koehler, heike.horstmann, jobst.loeffler}@iais.fraunhofer.de).

J. Jiang, the corresponding author, J. Ren, and Y.Weng are with Digital Media and Systems Research Institute, Univ. of Bradford, U.K. (e-mail: {j.jiang1, j.ren, y.weng}@bradford.ac.uk, phone: +44-1274-233695; fax: +44-1274-233727).

C. Mac Williams is with Academy of Media Arts Cologne, Germany, (email: carmen@khm.de).

J. Zaletelj is with Dept. of Telecommunications, Univ. of Ljubljana, Slovenia (email: janez.zaletelj@fe.uni-lj.si).

G. Güntner is with Saltzburg Research, Austria (email: georg.guentner@salzburgresearch.at).

Philipp Krebs is with ORF, Austria (email: philipp.krebs@ORF.AT)

I. Introduction

THE EU IST project LIVE investigates novel iTV formats, production methods and tools for staging of live media events in a professional broadcast environment. The vision of LIVE is to change the role of the traditional live TV Director to become a future Video Conductor communicating in real-time with his audience. The concept of "video conducting" can be understood in synergy to a modern day operatic performances on screen. To deliver live staging performances, the Conductor must in real-time direct the actors on stage, the orchestra in the pit, and respond to the audience's mood to ensure a highest possible quality of entertainment.

In LIVE, the Video Conductor is in the real-time environment of a live TV broadcast, who is asked to ensure a quality of drama through the use of parallel video streams to create a live video show. The Video Conductor can synchronize and link the parallel video streams at special points in time to allow the viewers to switch channels corresponding to their interests or mood [1]. The Video Conductor will be able to stage in real-time multiple interlinked streams, in response to the audience's mood and changes of interests yet in the same time a quality of entertainment show is ensured. The benefits for the public will be deeper emotional and intellectual experiences by mood-based navigation through interlinked video streams and tailored services of the audience's interests. The creative innovation behind LIVE is the development of a novel audio-visual language called "video conducting". The Video Conductor not only links the audio-visual frames in a horizontal but also in a vertical manner in real time to create an audio-visual Net-Composition. The vertical connections are not planned in detail beforehand, but will follow the creative spontaneous impulses of the video conductor while listening to his audience. This novel production method will allow the video conductor to improvise and break up the traditional rigid structure of horizontal linear storytelling. The novel LIVE iTV format can be best described as a constantly changing and growing net-composition around mass media events (e.g. Political elections, Olympic Games, natural disasters and war) by linking dynamically (on the fly) live multi-stream media with archived material and to spontaneously respond to the feedback of the audience [2].

Currently, in the time critical production process of live broadcasting, there is little time for the director and the production team to search databases for new materials, and thus archival content is normally pre-selected. This places a constraint on the ability of the production team to respond to the feedback of the audience in real time to satisfy their requests, to respond to unforeseen events and to satisfy creative impulses during a live broadcast. The technical innovation behind LIVE is the capability to analyze, link and recommend content from multiple content sources in a spontaneous and fast moving environment of the live broadcast [3]. The LIVE system during the live broadcast automatically analyses and aligns content from the multiple incoming streams and other available materials pre-selected from archives. Additionally, feedback coming in from the TV viewers (explicit requests, voting and implicit switching behaviour etc.) is also analyzed. The meaningful connections between viewer preferences and analysed video materials are then processed in real-time and fed back to the control room to support the production process.

Such proposed LIVE concept has been tested during 2008 Olympic games in Beijing, and an Olympic field trial has been carried out in Austria by the public Austrian TV Station ORF, which delivered a successful real-time LIVE broadcasting entertainment with the help of LIVE production system in a closed-wall IPTV environment for 500 participating end viewers. The expected impact of the LIVE project is the creation of a new broadcasting concept and introduction of a new market sector in the digital iTV industry: intelligent and interactive Television Services for the public media consumers.

The existing state of the art relevant to the LIVE broadcast and production concept is represented by digital interactive TV formats, such as those being used by BBC via their UGC Hub [3], CNN via their iReport [4], NBC via its Olympic Broadband Portal [5], and Sky News [5] via its live news coverage facilities etc. In comparison with the concept of LIVE, however, all the existing interactive formats are limited to the nature of being "off-line," where audience's involvement is primarily facilitated via web portal. Such an off-line interaction has a range of drawbacks, which can be highlighted as: (i) lack of guidelines and focus, result of which brings enormous unsolicited materials and thus creating serious problems for broadcasters in terms of editing or pre-filtering costs, making it no cheaper than the conventional and traditional journalism; (ii) lack of real-time involvement from the audience for live debate about the running production by giving live feedback to the director who may alter the running news programme; (iii) although the audience is allowed to filter the content according to their interests and change the size of the parallel videos on TV screens with NBC's interaction production, the role of the director is essentially unchanged, due the fact that NBC director is still not able to stage multiple interlinked streams on real-time basis and hence respond to the audience's mood yet, in the same time, ensure the quality of video production. In contrast, the iTV format implemented by the LIVE consortium allows the audience to influence the live show by giving live feedback directly to the director and production team in the control room. The novel multi-perspective iTV format, production method and LIVE tools enable the Video Conductor to create intelligent interactive TV services with the essential advantage that feedback from the audience can be spontaneously reacted to.

On the other hand, advance of digital technology, computer science, and semantics-based multimedia content processing has made it closer to reality for interactive and intelligent TV technologies. Representative work include video-on-demand [12], personalization of broadcasting services [6], interactive broadcasting [7-10, 13, 14], anytime TV [11, 16] and metadata analysis towards semantics-based audiovisual information descriptions [15] etc. In comparison with LIVE, the unique feature provided by LIVE is the coordinated efforts between TV broadcasters and computer scientists towards full exploitation of existing advancement of digital technology in digital media processing for intelligent and interactive TV content productions.

To provide a comprehensive picture about the LIVE broadcasting iTV format and its production concept, we present the rest of the paper in four sections, where Section 2 describes the LIVE tools towards interactive TV production, Section 3 present the LIVE technologies developed under the integrated project towards intelligent TV production, Section 4 reports the results of field trial during the Olympic Games 2008 and finally, Section 5 provides concluding remarks.

II. LIVE Interactive TV Production

The introduction of interactive TV [1-2] opens up new opportunities in the existing world of TV services. The interactive TV applications enable consumers to actively provide feedback on the TV programmes. In this way, the consumers' behaviour including content or channel preferences can be tracked, collected by the application, and hence fed-back to the content producer. The information on the TV audience's preferences will not only be used for the planning of the future TV programmes, but also be exploited during the production process of the live programmes such as sport transmission and news coverage etc. Such a concept would require that the user feedback is collected in real time and that the whole content delivery process is adapted to effectively by using this new information. The collection of user feedback through the ipTV application will enable modeling and tracking of the TV viewers within the TV production, and personalized content production for the target viewer groups.

User modeling and personalized content selection is used in a number of fields such as web documents, e-mail, news, multimedia, TV programmes, music, et al [3]. Personalization within the interactive TV domain is usually done on the level of selecting a suitable programme or channel out of a set of available channels for the individual viewer, and this service is called personalized electronic programme guide (pEPG) [4]. To be able to provide recommendations, the system must derive user's preferences towards content, which is done through analysis of viewer's past interaction with the TV content. The recommender system must thus track what the user is watching and collect viewers' feedback. The content-based recommender then learns the profile of the user's interest based on the features present in the past items that the user has explicitly rated or consumed. The content item is then compared to the user model and thus rated for each individual consumer.

A. The LIVE System

The IST-FP6 project "LIVE: Live Staging of Media Events" [7] is developing a production support system which will provide personalized content recommendations as well as an overview about viewers' feedback for the production team in real-time. The basic difference to the existing iTV recommender systems is that LIVE provides personalization of the programme content itself during the live production [5] along with selection among a set of available programme items.

Fig. 1 shows a conceptual diagram of a LIVE Feedback system, which was employed during the production of the Olympic Games 2008 trials for live broadcasting at the Austrian Broadcasting Company (ORF). The main components are described as follows:

- **Consumer's ipTV application** which is capable of displaying the notifications (questions) sent by the Director to a specific group of the audience, and to return feedback information.
- **Real-time notification and feedback channel,** which transfers the viewers' actions to the TV production team. The ipTV solution of Telekom Austria was modified and integrated with the LIVE production support system to display notifications (messages) from the production to the viewers, to transfer the viewer actions (responses to the voting and channel switch information) and to the TV production studio.
- **Feedback collection and analysis services,** which contains software tools developed by the LIVE team to provide real-time analysis of the audience, including their preferences and responses to the LIVE interaction mechanism.
- **Feedback application (GUI)** was developed as a front-end application which allows setting up the analysis parameters and graphically displays the results to the production team.

Content-level interaction for the viewer is enabled by the feedback channel, through which the viewer can influence the actual live TV content in the following manner: (i) Guided navigation through content ("switch" invitations) is offered to the viewers, where a message is displayed to inform the viewers about the content available on different sub-channel when the topic of the sub-channel changes. The message includes the invitation to instantly switch to proposed sub-channel by simply pressing the OK button. In effect, the production team can create a consistent story composed from several content items on different sub-channels, and guide the audience or structure the audience within several channels; (ii) Content selection by voting, where the Director is specifying several possibilities for the continuation of the live show, and the viewers are able to vote for their preferred content. After the votes are processed, the director can follow a popular opinion and show the preferred content. Through voting, a single viewer is able to influence the decision on the future content of a channel, which means that interaction between the viewers and the content producers is achieved; (iii) Information flashes are set up by the production team to provide additional information for the content to be broadcasted.

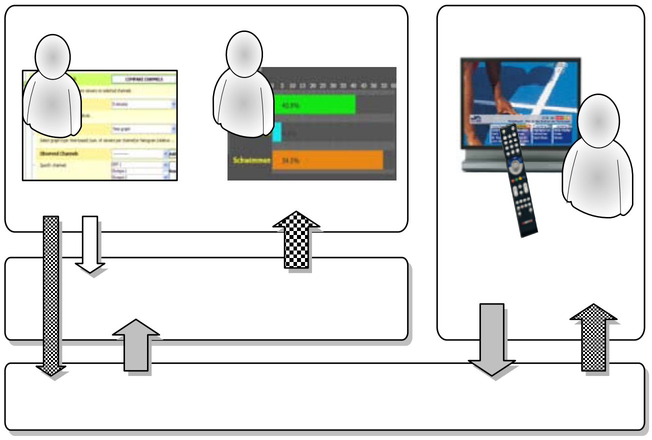

Fig. 1. Illustration of LIVE Feedback System.

The interactive options for the consumer are part of a new Video Conductor's vocabulary, and by using the tools like polls, switch proposals and information flashes, the production team is able to start a productive dialogue between the Director and the Audience, and use the real-time feedback as a guide for the content production. Under the LIVE interaction scheme, a number of modalities are analysed in real time and presented to the Video Conductor, which include: i) Number of viewers per channel, ii) Preferences of the channel audience, iii) Analysis of the user groups on an observed channel, iv) Trends of viewers, and v) Voting statistics.

The analysis is based on the feedback which is collected from every individual viewer, and then integrated over channel audience. Two basic forms of consumers' feedback were applied within LIVE system, which include: (i) Explicit user feedback by voting, where the viewer can choose one from the set of provided answers, and these answers were defined on-the-fly in the live production; and (ii) Implicit user feedback by tracking channel switches, where the switching between channels is collected by the system and the watching times are collected for each sub-channel.

B. Feedback Analysis

The consumer feedback is analyzed by the Feedback Analysis System and presented to the Director, in which the essential components include channel statistics, change rate of viewers, viewer group analysis and average relevance. Details of such analysis are given below.

Channel statistics are estimated to provide an overview of number of viewers on each of the channels. Let $N_i(t)$ denote the number of viewers on the ith channel at a time t. Let $u_{k,i}(t) \in \{0,1\}$ be a binary function which defines if the *k*th viewer is watching the *i*th channel at time t. The number of viewers per channel $N_i(t)$ is calculated for each of the 5 interlinked channels.

The dynamic aspect of the channel audience was also calculated as a *changing rate of viewers* $R_i(t) = \partial(G_\sigma * N_i(t))/\partial t$. A Gaussian kernel was used for smoothing the data. Similarly, $G_j(t)$ – the number of viewers was counted for each of the 32 user groups, which gave an overview on the preferences of the current audience towards different sport disciplines.

The group profile was derived from the individual profiles of the viewers. The consumer profile is represented by a vector $\boldsymbol{P}_k = [p_{k,l}]$, where $p_{k,l} \in \{0,0.5,1\}$ gives the *preference* of the user *k* to the attribute (sport discipline) *l.* The $r_{k,l} \in \{0,1\}$ represents a relevance of the attribute *l* to the user *k.* The group profile is calculated so that the *average preference* for the attribute *l* over all consumers who have provided preference value is calculated as

$$\overline{p_l} = \left(\sum_{i=1}^{N} p_{i,l} \cdot r_{i,l}\right) \Big/ \sum_{i=1}^{N} r_{i,l} \qquad (1)$$

Average relevance of a sport discipline to the user group is calculated as a relative number of users who have rated this sport discipline and denoted as $\overline{r_l}$. Finally, relative importance of the attribute (sport discipline) to the whole user group is calculated as the product $\overline{p_l} \cdot \overline{r_l}$. The final vector which represents a group profile in terms of attribute preferences is $\boldsymbol{G}_l = [\overline{p_l} \cdot \overline{r_l}]$.

Fig. 2 is an example of the weighted group profile $\boldsymbol{G}_l$ for the audience group “Men,” indicating that the football is the most popular sport. Similar group profiles could be calculated for the audience of a target channel at any time, giving the idea about preferences of the viewers toward different types of content. Feedback Application is an end-user application for professional users, which analyses and displays various modalities of feedback from the entire TV audience. Feedback Application was designed in such a way that the Video conductor is able to set up feedback analysis parameters and to review the real-time aggregated statistical data on his audience, including preferences of the channel viewers, how they rate his show, the results of voting and channel statistics etc.

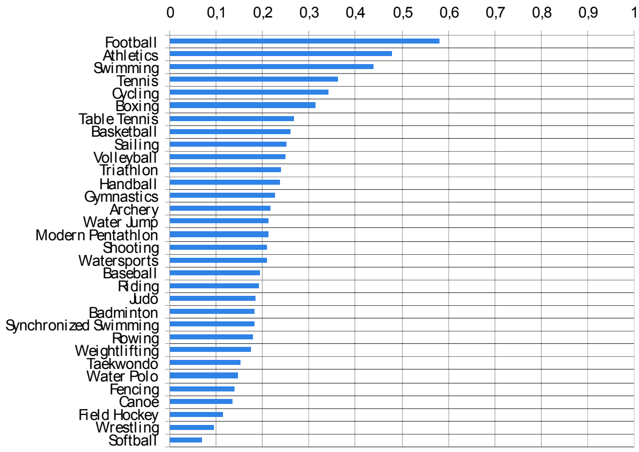

Fig. 2. Audience profiles for “Men”

Fig. 3 illustrates the Feedback Application window, which consists of two panes. While the left one contains a tree-structure for setting the graph parameters, and the right one contains the graphs created using the settings from the left pane.

- **Voting and rating (explicit feedback)**. The viewers are able to express their opinion whether they like the current program or not (rating) and are able to answer for various voting polls. The results are shown in a graphical format as a rating chart (histogram or pie chart), and voting result chart (**histogram** or **pie chart).**
- **Channel comparison graph** based on implicit consumer feedback. The watching times of Consumers are collected for each channel and time-based channel viewing statistics is computed. This feedback is even more powerful in combination with information on consumer profiles. There are two types of graphs available for displaying channel viewing statistics: *time-based graph* and *histogram*.
- **Observation of the target groups.** The Director is able to track the distribution of several pre-defined user groups on the target channel, which provides information about demographic or interest distribution of viewers on a target channel.

Dynamic behaviour of the Audience

As shown in Fig. 4, histogram charts display trend markers to make it easier for the user to have a feeling about whether the number of viewers is increasing or not. The Feedback GUI remembers the past values of viewers on different channels and compares them to the present values. It then draws line indicators where previous values used to be. The lines are coloured red and green, depending on whether the number of consumers decreased or increased.

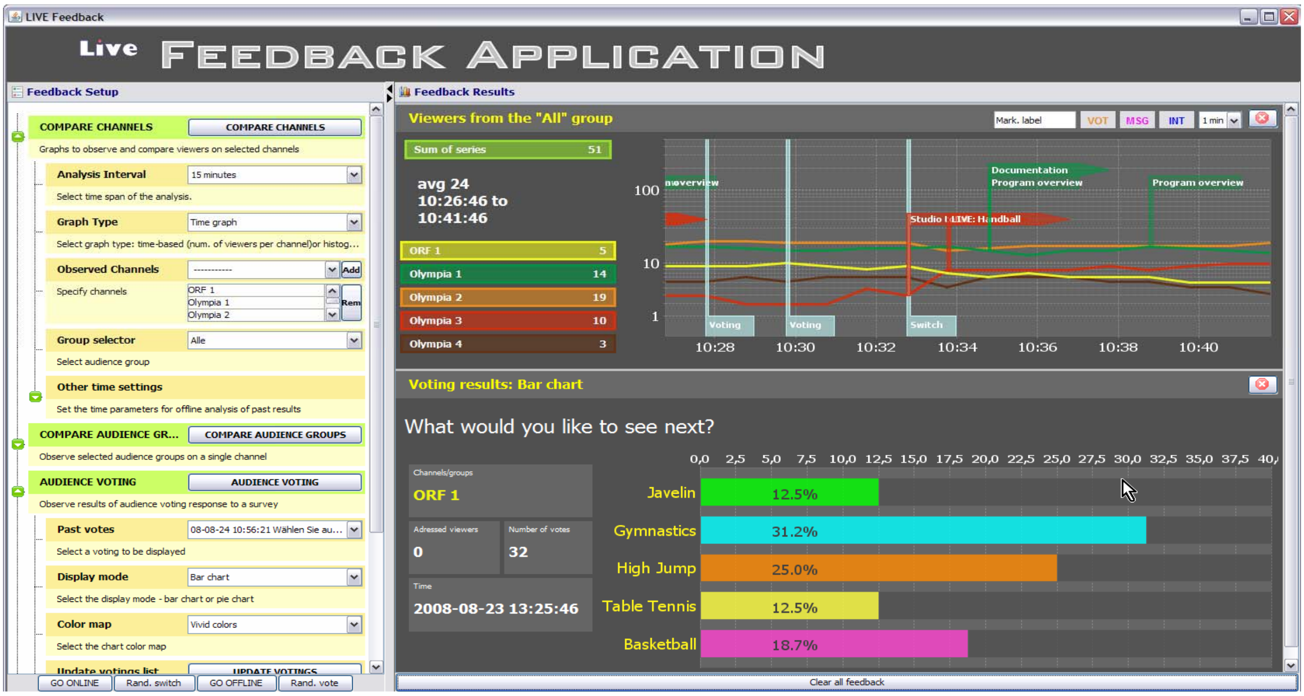

Fig. 3. Feedback Application window with histogram and time-based comparison of target channels.

Notifications and alerts for the production team

Time-based chart provides some additional options and functionalities. To the right of the graph title there is a set of buttons, which are used for displaying graph markers. Markers denote a moment in time when something significant happened (the director initiated a voting, sent a message etc.). There are three types of markers: for voting, for messages and for time-intervals. The markers allow the team to easily observe the effect of some actions (such as initiation of voting) on the audience over time.

Besides the markers, the time-graph provides an option of alarming the professional about some important trends (positive or negative) in the audience. An example of a negative trend would be a sudden drop in number of viewers on a channel or no viewers on the other etc. When an event occurs, the system alarms the user by painting a colour frame around the chart on which the event occurred. The frame is red for negative events and green for positive.

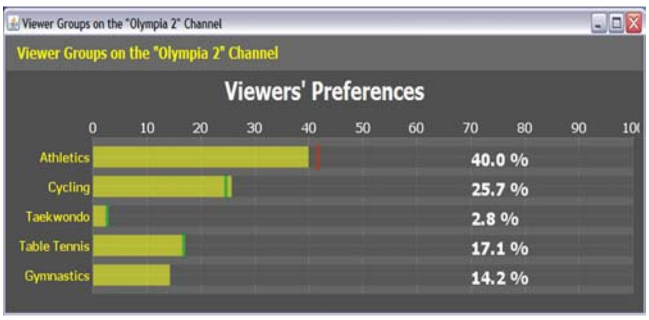

Fig. 4. A typical histogram with trend markers. Green markers denote the increase in the number of viewers, wile, red markers denote decrease.

III. LIVE Intelligent TV Production

At present, management and operation of the video streams is done by operators, members of the production team. The planning of the contents of the TV show is supported by common PC-based software tools such as spreadsheets and is not tightly integrated with the production system [17-19]. Today's TV production team is working with audio-visual content directly, and is not supported by dedicated and integrated content management and processing systems. The production workflow is also not explicitly modeled and supported by any integrated information system, which means that such live production requires highly trained and experienced production teams.

The introduction of multi-channel live TV broadcasting has brought new challenges and increased demands for live TV production teams. The volume of the input sources for large live events, such as Olympic Games Beijing 2008, is huge with 12 parallel multi-channels. As a result, while the existing multi-channel production is carried out by multi-production teams on one-team-one-channel basis, the technologies invented and the concept introduced by LIVE makes it possible to complete the multi-channel production by one team. Additional benefits brought by LIVE also include the concept of the lively interlinked output channels, which are synchronized to create consistent stories across channels. Consequently, the content production team by LIVE system is capable of not only dealing with the enormous volume increase of content, but also tracking and synchronizing individual content items on output channels in parallel.

In an approach to manage the complexity of the concept of interactivity in a live multi-channel production environment and to provide intelligent decision support, the LIVE consortium has developed a software system (subsequently called the "LIVE system") supporting the broadcaster's production team during the preparation and the staging of a media event. The LIVE system allows capturing, managing, filtering and retrieving situational statements about the context of audio-visual content used to create the interactive TV show. From the architectural point of view, the LIVE system consists of five major sub-systems:

- The Metadata Generation System is responsible for the detection, extraction and annotation of live content. It contains automatic annotation tools as well as the human annotation tool, a graphical user interface to be used by a human annotator.
- The Intelligent Media Asset Information System is a component of the Intelligent Media Framework which is responsible for the management of the controlled vocabulary, event information and intelligent media assets. The system also implements services to manage and query this data.
- The Video Conducting System covers all components necessary for the LIVE staging process. This includes the user interfaces to be used by the editor and video conductor as well as interfaces and adapters for the broadcasting environment.
- The Recommender System is responsible for the calculation of content recommendations for professional users as well as for consumers. In addition, the recommender system collects and analyses feedback provided by the consumer.
- The Consumer System supports the functionality needed at the user's end and provides personalisation and feedback options, which are associated with the non-linear multi-stream video show.

A. *The LIVE Staging Domain*

The LIVE system provides support, on one hand, during the preparation phase of an interactive TV show. On the other hand, it is operational during the production phase. In a field trial in August 2008, when the system was productively tested and successfully evaluated during the Olympics Games in Beijing, the LIVE system had to face the following challenges: Up to twelve parallel clean feeds of audio-visual live material were provided by the International Olympic Committee. The production team could monitor all the incoming streams and identify audio-visual content meeting the requirements of the current staging concept. Additionally the production team could produce five parallel output streams in order to create a compelling and consistent story across channels. Furthermore the successful production of the show was highly depending on the availability of archival material on demand. The technical platform of the LIVE system therefore comprised the concept of a "production archive" in which several hundred archival clips were kept ready for immediate on demand playback. A description of the clips (metadata) of the production archive was developed in the preparation phase based on the records of the broadcaster's archive.

In such a scenario, the production team is confronted with the complex task to maintain the overview about twelve incoming streams, to produce five thematically interlinked parallel output channels, to keep track of available archival material in the production archive, to insert interactive elements (e.g. switches and pollings) and to monitor the corresponding audience feedback. The overall concept to support and manage the production process included the design of a knowledge based middleware to migrate and integrate information about the main aspects of the staging domain. As a first step to introduce intelligent decision support for the staging domain and to design the intelligent content model for LIVE, the domain was analyzed and decomposed into four main informational aspects:

- **Event**: Information about the event to be staged is crucial for planning the show. During the event, a production support system has to collect and analyze information in real-time to support the LIVE staging process. Typically event related information is provided by dedicated information systems.
- **Content**: The system has to handle essences (i.e. the raw video streams) as well as their metadata. Live essences are created by filming the event (the live feeds) and activities around the event (e.g. interviews, analysis of experts, etc.). Also metadata about this kind of content is created under (near) real-time conditions. Content may also come from the production archive. The production archive is filled in the preparation with content from the long term archive. The provisional user responsible for this preparation is the archivist. Metadata of prepared content is available prior to the production phase and typically does not change during the staging process. However, the usage of prepared content may result in additional metadata and annotations.
- **Staging**: Staging comprises all knowledge about how an event is presented to the consumer. This includes: (1) content formats, which consist of general stream types, different types of switching between streams as well as screen layouts like split screen or picture in picture; (2) stream profiles that define hot topics for a specific stream and the target audience of each specific stream; (3) stream interrelations defining situations, in which linkages between two streams are possible; and (4) audience profile modeling for the targeted audience of the show.
- **User Roles** and **User Spaces**: The LIVE staging process defines different user roles and three user spaces. The "annotation space" covers all human and software agents assigned to annotate the incoming live feeds. The "production space" covers all the professional users working in the control room (the conducting team). This includes the video conductor, editors, archivists and some assistants. This team performs the actual staging of the show. Finally, the "user space" covers the consumption of the show by the consumer and the interaction of the consumer with the LIVE system and the video conducting team by providing implicit and explicit feedback.

Among these informational aspects, a series of relations was considered to derive requirements for the design of the intelligent content model. For example, the fact that some live content is produced in the course of a specific event creates an immediate relation between that content and the event. Additionally, the editor and the video conductor can use archival content to present special aspects of the event to consumers, thus relating archival content to the current state of the event and the show. In the preparation phase this relation is important to search and find archived content. In the production phase this relation is important to search and find available content suitable for the current state of the event.

Apart from the domain specific aspects the following core knowledge requirements were identified as the foundation of the design of the intelligent content model for the LIVE domain:

- Make information machine-processable: Knowledge must be encoded using a formal language and interpretable by machines to enable the system implementation supporting knowledge workers in handling content objects.
- Support cross-domain interoperability: Some aspects of the knowledge about the content are independent of the application domain, e.g. media coding information, rights management information, etc. Established standards of these domains must be respected and aligned to the concept of knowledge content.
- Allow multiple interpretations: This aspect is typically neglected by traditional content management systems, yet this is important in production environments in which employees with different job roles have different perspectives on the same content objects.
- Unleash hidden information: This covers the idea to link content with knowledge that cannot be directly derived from the content: Media content only contributes to a larger narrative. Even though only a small part of the story may be supported by some content it must still be possible to state that the whole story is related to the content. Looking at it the other way round, this requirement explains also the significant weakness of systems which only rely on feature recognition to explore the meaning that is hidden inside media objects.
- Make knowledge independent of content: In traditional media asset management systems, knowledge is typically dependent on the content object. As the analysis of the domain showed, this is not a valid assumption for most of the interrelations between knowledge and content. Particularly, planning activities of knowledge workers first create a description of the topic of interest, and later search for suitable content objects, which support the conceived story. In general, descriptions of the subject of the content seem to be more dependent on the topic and the user than on the content itself. Knowledge descriptions (especially annotations) therefore must be autonomous from the content objects.

B. *Towards An Intelligent Content Model for LIVE*

Based on the domain analysis and the core knowledge requirements, several content models have been assessed for the LIVE staging domain (see [22] for an overview and comparison), e.g. MIPS (Multimedia Information Presentation System) developed a distributed, knowledge based media annotation and presentation system ([26]). With the rise of semantic technologies, content models with ontological grounding were developed, e.g. the ZYX model ([21]), the Enhanced Multimedia Objects (EMMOs, see [22]), and the Knowledge Content Objects (KCOs, [20]). These models incorporated other multimedia content description schemes, such as the MPEG-7 and MPEG-21 standards, to describe facets of the content models.

The LIVE staging domain was finally mapped to a knowledge structure which forms the basis of the intelligence of the LIVE system. The KCO model was used as a starting point for LIVE's intelligent content model because the KCO concept already addressed several aspects of the core knowledge requirements of the LIVE staging domain. However, even this model had to be extended to cover the real-time aspects and the capability aspect to contextualize content depending on user preferences and to allow efficient and comprehensive storage and retrieval ([25]). The resulting knowledge model consisted of three sub-models which will be described in the following sections.

1) *The term model*

The Term Model defines the foundation of the knowledge structure. It is based on functionalities defined in the Simple Knowledge Organization System (SKOS) and supports the definition of vocabularies, terms, term metadata and semantic relations in between terms. In addition, the Term Model defines a base typology for terms (the "term type hierarchy") which on the one hand aims to ease the definition of controlled vocabularies for application domains by providing a simple hierarchy, and on the other hand provides additional semantics exploited by the other knowledge models of the LIVE staging domain. All terms of the controlled vocabularies are aligned to the term type hierarchy.

The semantics for the categories are based on the DOLCE foundational ontology ([27]) and one of its extensions, the Descriptions and Situations (DnS) ontology ([23]).

2) *The event domain model*

The Event Domain Model provides a knowledge structure to describe events. It supports the definition of formal statements describing the current state (annotating) or a future state (planning) of events. Such statements are used to describe the staged event as well as the staging process itself. They are instantiations of concepts defined in the Event Domain Model (e.g. Agent, Activities, Events, Settings). Basically this model can be used to describe any kind of event but in the scope of LIVE it is mainly used to formally describe planning information about sport events. During the course of the event, it is used to store knowledge about specific activities and the athletes performing those activities. With respect to staging, this model is used to define the setting of a show and the single streams, anticipated parts of the show as well as specific activities performed by the video conductor or the editor as part of the production of the show. The Event Domain Model uses controlled vocabularies based on the Term Model as part of its knowledge structure.

The foundation of the Event Domain Model is provided by two design patterns of the DOLCE foundational ontology ([27]). The first one is the participation pattern describing that individuals and objects can participate in activities. The second one is the agent-activity pattern, which provides a specialisation of the DnS pattern used for the term type hierarchy. This pattern, as shown in Fig. 5, describes that individuals with a specific role participate in actions corresponding to specific tasks.

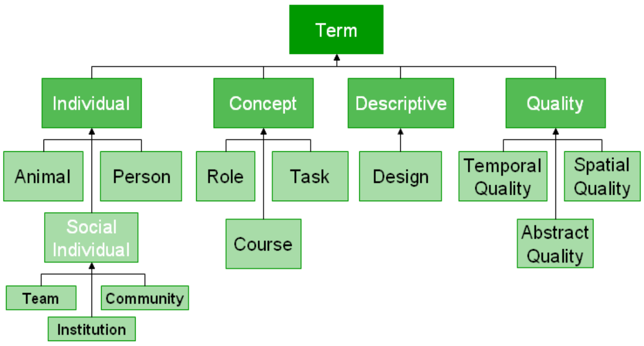

Fig. 5. LIVE term type hierarchy.

C. The Intelligent Media Framework

Once the knowledge model for the LIVE staging domain is defined, a software framework can be introduced to make all this knowledge accessible and operational: The role of the "Intelligent Media Framework" (IMF, shown in Fig. 6) in the production process is to accept and handle partial information about media items, to add semantic information to the items and to infer and attach contextual knowledge to the items by relating them to the events. It furthermore provides knowledge services that offer controlled vocabularies related to the current context of a stream to guarantee the unambiguousness of the terms used. The IMF provides the knowledge infrastructure for the LIVE system and is based on a multi-tier service-oriented architecture in which all interfaces to external components are modeled as Web services (see [24] and shown in Fig. 7).

- Knowledge Service: Provides access to the vocabularies and terms stored in the LIVE knowledge base. The knowledge base holds the terms used in the LIVE system. As an example, the Human Annotation Tool uses the Knowledge Service to retrieve relevant information about all participants of an event, as well as initialising its user interface with annotation terms based on the vocabulary of the discipline of the selected event.
- Event Information Service: Provides access to information that is specific to a staged event. During preparation, this includes the date and location of the event as well as the participants. While staging an event, related information to this event, such as activities, are collected and stored.
- IMA Service: Provides access to knowledge based content annotations – the Intelligent Media Assets (IMAs). This service can be used to access information about prepared content as well as content annotations that have been created during the live staging about the incoming live feeds.
- Semantic Context Service: This service uses a semantic index to provide queries and access to the semantically enhanced context of any resource managed by the IMF. This includes terms, event settings, events, activities, agents as well as intelligent media assets. The IMF supports three levels of semantic contexts based on the topic classification type.
- Action Message Queue: The IMF includes a messaging system to support the real-time aspect of the LIVE staging environment. The Action Message Queue represents this subsystem as a notification service about the current state of the event, the staging and available content to the other components of the LIVE Production Support System.
- User Service: Provides authentication to applications.

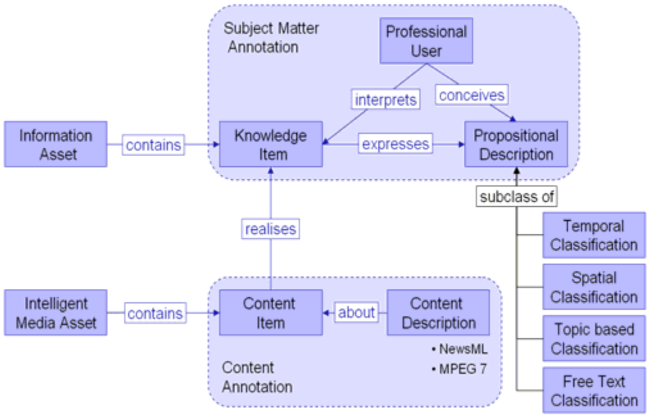

Fig. 6. Architecture of the intelligent media framework.

D. Putting the Intelligence into "Action"

In the previous sections, we pointed out the knowledge structure of the LIVE system and introduced the Intelligent Media Framework to make this knowledge structure operational. This section gives three examples of how the intelligence of the LIVE system is actually applied in areas of real-time annotation, to information visualisation and to the recommendation of audio-visual materials.

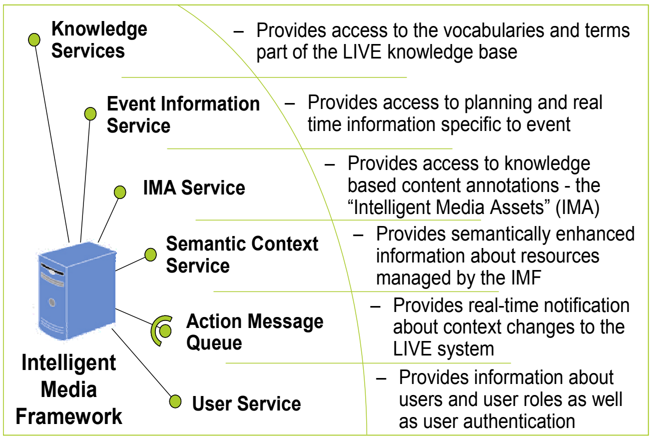

Fig. 7. Services of the intelligent media framework.

Conceptually we distinguish three types of applications (1) Tools and components used for the population of the knowledge model in the preparation phase (e.g. the Event Ingest Tool to retrieve and edit event related information). (2) Tools for real-time annotation and for the update of information in the production phase (e.g. the Human Annotation Tool). (3) Tools and components utilising the acquired knowledge as well as the semantic context service to provide intelligent decision support in the production phase (e.g. the Staging Console and the Content Recommender). Common to all three types of applications is the fact, which the underlying complex knowledge model had to be "flattened" and "hidden" on the user interface level to allow an easy-to-use access to the functions even under the pressure of the real-time requirements of a live production environment. At the same time, the references to the knowledge model has to be maintained and meaningful, and the contextualized information has to be delivered to (or retrieved from) the Intelligent Media Framework in the background.

To give some figures from the field trial which was executed on three weekends during the Olympic Games in August 2008 with an average production time of five hours per day: The LIVE knowledge model was populated with 18.000 terms in 65 vocabularies. These terms comprised 23000 athletes, 14000 officials, 6500 events and 1100 teams. During the production phase, a total of 110000 action messages were sent, 100000 activities were detected and stored and 18000 knowledge items were created.

1) Knowledge extraction and real-time annotation

Although content-based retrieval of digital media is found more effective in querying multimedia data, it suffers one basic problem to fill the gap between low-level features and high-level semantics [28-29]. Accordingly, extraction of semantics for automatic content annotation is emphasized, especially for content-rich video sources [30-31]. The preliminary purpose for automatic annotation is to release the huge labour burden from manual annotation and also improve the flexibility and accuracy. This can be achieved through automatic extraction of semantics in the video sources, which include semantic events and concepts as well as classification of its associated disciplines [32-34]. Accordingly, the work will help to fast categorising and retrieving of relevant videos towards effective management for post-production applications.

Real-time annotations are used to create event situation statements and set content in context of an event. They are also used to inform the LIVE system about updates on the state of an event. This is accomplished in three ways: (1) By automatic detection components monitoring the incoming live streams and sending annotations to the Intelligent Media Framework. (2) By human annotators producing manual annotations of monitored audio-visual streams using a graphical user interface. (3) By messages from an external event information system (called Sport Information System) which filtered real-time information from the official Olympic information system and transferred them to the Intelligent Media Framework.

The automatic annotation system can be combined with other supporting technologies and form a complete system for content-based video annotation, retrieval and delivery. To illustrate how the system works, a diagram is given in Fig. 8 which contains two main parts, i.e. extracting part and application part. The former focuses on extraction of several events and semantic concepts for content indexing and automatic annotation; while the latter uses these indexed semantics for video retrieval and delivery. The retrieved results can be summarised for efficient delivery in accordance to the requirement of the application, i.e. a quality index of the video determined by the network for data transmission. As seen, this helps to provide a useful strategy for content adaptation in retrieval and network transmission applications.

In order to achieve automatic annotation, effective extraction of semantic concepts and events from video sequences is required. Consequently, some technical requirements for the system can be summarized as follows.

- Compressed-domain processing for efficiency, where access to DC-images and motion vectors is essential for several functional components in terms feature extraction and classification;
- Machine-learning techniques such as support vector machine (SVM) and Bayesian statistical analysis are used for decision making;
- Data organisation and indexing techniques for efficient data management.

For efficiency, online processing is implemented in the compressed domain to avoid time-consuming inverse discrete cosine transforms (IDCT) [35-36]. Several novel features can be extracted including statistical analysis of conventional DC-images [37-42], the inter-frame difference and motion prediction errors. These indicators are then used for shot boundary detection (SBD) [37-38], camera motion detection [39, 43] and extraction of human objects and several semantic concepts [42]. Please note that the SBD techniques had been successfully applied in TRECVID'07 in which the performance on cut detection and an overall evaluation are ranked the first and the third among all participated teams all over the world [37].

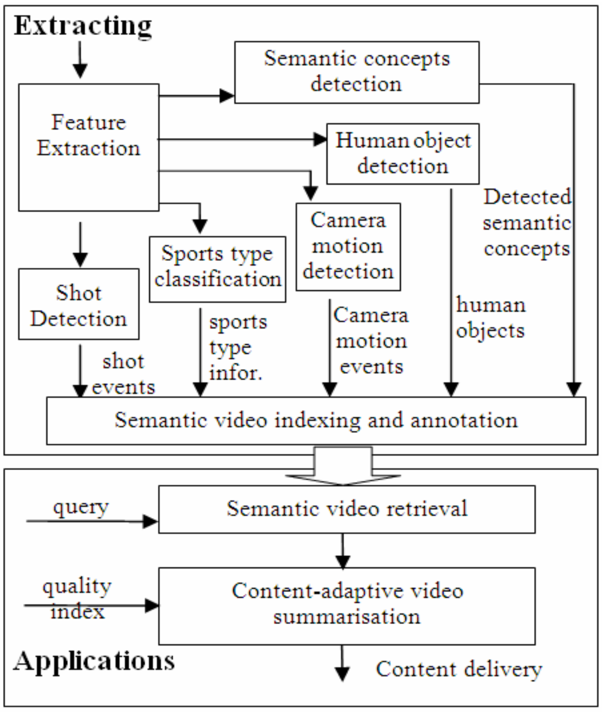

Fig. 8. Diagram of knowledge extraction and automatic annotation system.

Camera motion here refers to zooming, panning and tilting, where zooming in of human objects is particularly desired as close-up events which is often encountered in sports videos [39, 43, 56]. In addition, human objects are detected via statistical modeling of skin pixels plus spatial-temporal filtering [44-51]. The statistic approach is found robust to the change of pose and size where template-based techniques such as Haar-like features may fail.

Semantic concepts including outdoor, indoor, building, sky and plant together with hierarchical classification of sport disciplines are implemented [42, 57]. In fact, the support vector machine (SVM) is utilised for this purpose, using low level features including colour, edge and texture measurements. For plant and the sky, dominant colours of green and blue are respectively exploited. Meanwhile, edge information can be used to distinguish natural scenes and artificial objects [52]. The former contains edges of arbitrary directions and lengths and the latter has long and straight lines, especially for building and indoor scenes. In addition, there are no apparent edges in the sky object. Furthermore, coarse and fine texture details can be found respectively in outdoor vs. indoor and building vs. sky scenarios. These are then used for supervised learning in detecting the semantics concepts for retrieval and annotation applications.

With extracted semantics, videos can be retrieved using their associated semantic contents. To cope with their large sizes for efficiency in delivery, content adaptive video summarization is proposed which aims to represent the video in a short version while keeping the most important contents [53-55]. In principle, the essential strategy is to choose the most meaningful parts of videos to form the summary while ignoring the less important ones, which are often referred to as content of interests (COI). In our system, these COIs are defined using our extracted semantics hence the summarised videos are capable of keeping these semantics for consistency in video retrieval. Again, the developed summarisation techniques have also been applied to BBC Rush Videos in TRECVID'08 with very promising results being achieved [53], in which effective shot detection plays crucial roles in dividing input video streams into content-consistent sections [58-60]. According to the network performance, a quality index is then determined to decide the size of summarised video for transmission. As seen, this helps to provide a useful strategy for content adaptation in retrieval and network delivery applications.

In the annotation system, most of the low-level processing is embedded for simplicity [57]. The user only needs to specify a sequence as input, then the system will automatically analyze the video and generate knowledge as output to denote the extracted semantics. Meanwhile, the user can control the play of the video and also browse the annotations. Finally, the annotated outputs can be stored for re-usage in the live stage production process.

The Human Annotation Tool for example used the Event Domain Model to retrieve information about the event, such as its participants, the types of activities related to the discipline of event (e.g. "overtaking" for a bicycle race), etc. During the field trial a team of annotators simultaneously annotated two Olympic live streams each based on a controlled vocabulary (obtained from the Term Model), but also free text annotations were allowed. Internally annotations are encoded as MPEG-7 annotations and transferred to the IMF message queues for further processing. This mechanism is also used for automatically detected annotation (e.g. black screen detection, no motion detection, face recognition, etc.).

Real-time event updates are accomplished by connecting the LIVE system to an external event information system called "Sport Information System," which provides real-time updates on the state of an event (e.g. time of an athlete after the first lap, the result list of a race, a detailed status about the participation of an athlete, location information, weather information, etc.) Most of this information is retrieved in a push type mode and aligned to the LIVE knowledge structure by the messaging and subscription mechanisms provided by the Intelligent Media Framework.

2) Staging support

The general purpose of the LIVE Staging Console (see in Fig. 9) is to provide useful information to professional users in the production room. The Staging Console is a central decision support application filtering and highlighting annotations and messages in the context of the show. It makes full use of the underlying knowledge model (e.g. to semantically enriched annotations), yet presents this information in a user friendly manner. Several panels are offered by the application: the event schedule is displayed in the style of a Gantt chart and as a list. Moreover, the incoming live streams are displayed with additional information (e.g. type of sport, nationality of participants) about the actual context of the streams. This information is retrieved from the knowledge model by combining the content and event information with the notifications from the automatic and manual annotation components (e.g. black screen detection, no motion detection) in real-time. Three panels of the Staging Console show filtered, prioritized and highlighted event situation statements created by the human annotators or by the Sport Information System. The statements of the human annotators usually contain valuable information for the production team and the audience.

3) Content Recommender

The main goal of the Content Recommender is to enable usage of archived content within the live production scenario. It provides a user interface as shown in Fig. 10 for the professional user to search and retrieve archived clips. On the other hand, it provides automatic methods, which combine preferences of the target audience (the TV viewers) with the predefined topics of the TV program, in order to find the relevant content in the production archive. In this way it supports the production of personalized and targeted TV channels, directed to specific TV audience groups.

To enable the generation of live context-dependent and personalised recommendations from the production archive, the Content Recommender receives action messages produced by automatic or human annotation systems. These messages carry information about the actual activities (actions) within the event (and the video streams). This information enables the Content Recommender to build a real-time working context for the content search. An example of such a real-time event action context is given below:

```
Event_Action_Context  {
   Agent = { Participant Lane 2 }
   Role = { Winner }
 }
```

The action message updates the working context of the Recommender, so that the Athlete Name value of the Athlete Interview Template is set to “Pellegrini,” who won the actual race. The Recommender can be programmed to react to Role Winner (to the action message when somebody wins) by triggering automatic search. The search query is based on the current template combined with the real-time action event message. In this way, the interview or portrait of the winning athlete is retrieved from the archive.

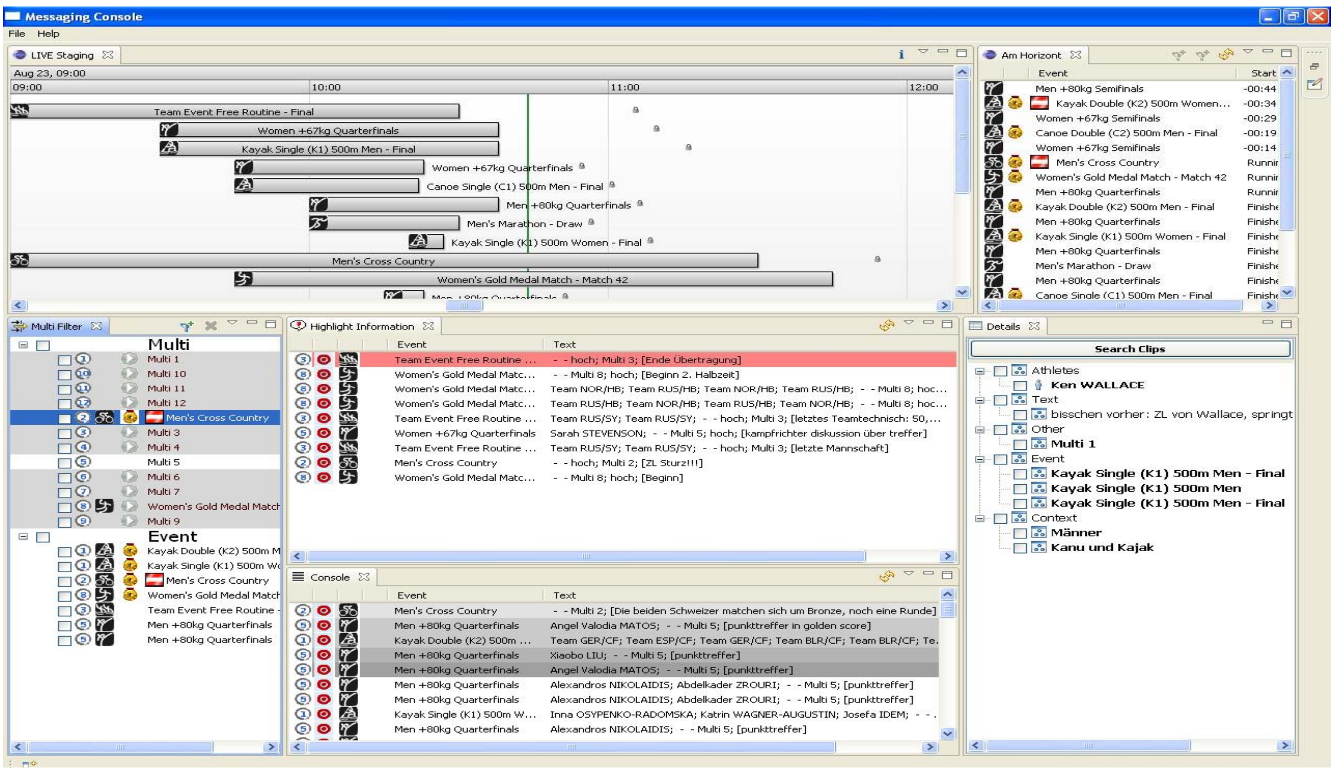

Fig. 9. LIVE Staging Console.

Fig. 10. LIVE content recommender.

IV. Evaluation of LIVE System via Olympic Field Trial

The LIVE Olympic Trial is carried out by six days of broadcasting multi-channel Olympic programmes, which totaled to 138 hours of video materials that was broadcasted to 489 registered users. During the trial, the new tools for Director – Audience communication and feedback were continually used, and the production team made a sequence of experiments on how to use the tools in a most efficient manner. Assessments were made on how such new interactivity options are accepted by the test viewers.

During six days of the trial, a total of 37 information flashes were sent to the viewers, the majority of them (59%) were announcing the next programme topic, for example “Stay tuned! Coming next: 100-meter Sprints – Men”. Latest sport news, apologies and voting results were also presented by the information flashes.

The polling (voting) was also used in the field trial of the LIVE system, and total of 39 polls were made. The polls are divided into 5 categories: Program planning (59%), Viewers’ opinion (18%), Users’ acceptance of the content or format (15%) and entertainment (8%). Questions were mostly sent to all channels, only some were sent to individual channel.

The “switch” option allowed the production team to control and guide audience over the channels. The team could propose to the viewers to switch to another channel, by announcing the contents of the target channel. The option was extensively used (total 144 times) during the trial.

The interview with the Director proved that the new real-time communication and feedback options became very important for him during the trial, and that he considered this as an indispensable tool which he wanted to use in his daily work. He considered the polling function as a kind of decision support for his work, and through the results he gained validation and confirmation of his decisions from his viewers. Through the trial, he gained positive experiences in the developed tools, and his final verdict was positive, as the feedback system became for him a personal highlight of the trial production.

The users’ acceptance of the LIVE program format and interactive elements was assessed by the market research done by GfK Austria. Results show that the overall user satisfaction was very high, and 94% respondents stated that they enjoyed the trial. The satisfaction was the highest among users who are generally interest in interactive TV. In terms of interactive elements, the switch function is well received, with which 87% users were satisfied. The next satisfactory level is with the voting system, with which 75 % of the audience is positive, and many viewers appreciated that by using this function they were able to influence the program content. As a matter of fact, the viewers stated that through polls they gained the feeling that the broadcaster responded to their interests. Additionally, 21 users liked the possibility to express their own opinions.

To prove and evaluate the results achieved in more than two years of research, the LIVE project performed a field trial during the Summer Olympic 2008 in Beijing with invited IPTV customers in Austria. LIVE's partner ORF and a team of researchers produced an interactive, multi-channel live TV programme for six days to provide consumers with an enhanced viewing experience. By dynamically linking the programme streams, providing always up-to-date programme information and with the help of a moderation team, the audience was given orientation and navigation support within the multi-channel programme bouquet. Further, the test audience directly influenced the flow and the content of the broadcast by switching channels and giving feedback. The following paragraphs give more detail on the setting and preparation of the field trial, including detailed description and discussion of its execution and results.

A. Setting up

The Olympic field trial was primarily planned by ORF like any other live sports production. A specific production plan dedicated to the LIVE consortium was created by specifying the schedule, the personnel and the technical resources. Since the field trial productions had to be performed in parallel to ORF's main coverage of the Olympics, the limited resources had to be shared in most the production processes. Thus the field trial production also benefitted from the proceeding preparation on ORF's main channel ORF1. Tasks, such as booking of live feeds from Beijing, creating a dedicated on air screen, design or producing special documentaries on China, were performed once and were used for both productions. A full-fledged sports control room, a studio, a commentator spot (off-tube), a room for the technical equipment and last but not least a dedicated production team was provided by ORF. The team consisted of:

- a director, performing the role of the video conductor;
- an editor, responsible for content selection and preparation;
- two video and two audio engineers, dealing with two outgoing signals each
- an ITV assistant, timely creating interactive elements
- two production archive assistants, taking care of the archived video content
- three human annotators, annotating two incoming feeds in parallel
- two moderators, a couple of commentators and additional technical staff

Twelve live feeds from Beijing, a number of additional clips stored on the production archive, studio cameras and a mobile camera for "behind-the-scenes" footages were available for the production of six live shows with four interlinked channels. To support the production team in the task of staging an interactive multi-channel programme, the second prototype version of LIVE's technical system was set up at the production site. A technical project team was available before and during the productions to assure the accurate operation of the system.

The test audience comprised nearly 500 invited and registered Austrian households with an installed IPTV platform. A dedicated test platform provided by Telekom Austria was used to distribute the bouquet, and its interactive elements in parallel to the regular TV programmes were made available only to the test audience.

B. Preparation

The preparation of the Olympic field trial comprised production related aspects like planning of resources, content related tasks including selection of archive clips as well as technical setups and tests. Since the production plan and other details of the setting were already introduced, the following paragraphs will concentrate on the preparation of content and other technical aspects.

1) Content Preparation

The preparation of content has already been identified to be most crucial and time consuming at an early project phase. Audio-visual material from the long-term archive had to be specified and selected, tailored where necessary and ingested as distinct clips to the production archive. The related metadata was imported, enhanced and stored in LIVE's Intelligent Media Framework. The ingested clips were also used for training the automatic detection, extraction and annotation algorithms.

Apart from the audio-visual material, the Intelligent Media Framework itself had to be prepared. Terms, topics, and categories of the Summer Olympics and their relations had to be modeled and represented in LIVE's central knowledge base. Moreover, the set of event information provided by the central information system of the Olympic Games had to be imported and on-demand updated during the production.

2) Technical tests and setups

Systematic testing of the LIVE System was started at an earlier stage of development. The system components were installed on a dedicated hardware platform and iteratively tested in relation to distinct usage scenarios. The testing was completed by final system tests in concentrating on aspects of performance, robustness and scalability. Afterwards, the system was deployed as part of the broadcasting infrastructure at the premises of ORF.

The deployment of the system was planned and prepared in close cooperation with ORF's technical department well in advance. Inspections of the premises took place and detailed cabling and setup sketches were created to guarantee a fast and successful integration of the system. Most of the cabling and dedicated network connections were prepared in conjunction with the required adaptations and reconstructions of the control room RP12. Since this control room normally is equipped for producing one outgoing signal at a time, additional equipment and work spaces had to be integrated to allow for four outgoing signals in total. Besides a second video and audio engineering work space, places for the ITV and the two production archive assistants and their related work stations had to be arranged. The three human annotation spaces were set up outside the control room in a near by room where the technical team was monitoring the operation of the LIVE System

C. Execution

In cooperation with the LIVE research team and supported by the LIVE System, the production team produced six live shows with four interlinked channels over three weekends during the Olympic Summer Games 2008. Each production took about six hours and ran in parallel to ORFs main Olympic coverage broadcasted on ORF1. 12 live feeds from Beijing and archived material were therefore used, including live camera feeds from the studio. The main channel of the interactive bouquet Olympia 1 comprised interviews with studio guests and "Behind-the-scenes" reports shot by the mobile camera team. In addition, a number of documentaries about China, the athletes and highlights of the Games were presented on this channel. The three remaining channels Olympia 2 to 4 concentrated on the live coverage of the competitions and on archived highlights of the past few days.

In most of the times, the production team followed a detailed plan (run-downs) for the current production day based on a common staging concept. Prior to each production, a short briefing of the team took place, where the editor presented the current plan and the director instructed the moderators and the mobile camera crew. In case of a programme change on one of the four channels, the team was instructed by the director in accordance to the plan. He also instructed the commentators sitting either in Beijing or in the local off-tube compartment next to the studio. To share the commentators abort but also to coordinate the programmes, the director kept in contact with the team producing in parallel on ORF1. The goal was not to show the same pictures as ORF1 since the channel was also included in the interactive bouquet for the test audience.

Invited by the moderators and informed by textual inserts, the audience could get in contact with the production team and thereby influence the programme. Besides various poll offerings, the consumers could also ask questions and comment on the programme via email or chat. Consumers not actively responding to the participation offerings can still influence the programme through their viewing behaviour. On a central screen, the video conductor and his team can observe the current audience charts and reacted accordingly (See Figures 3 and 4).

To inform the audience about the current events, new programme items within the bouquet or breaking news, the production team created textual and picture-in-picture inserts included in the broadcast as well as information and interactive elements distributed via the IPTV platform. By setting switches the audience not only was informed about a programme offering. Those interactive elements allowed also for linking to the related channel. Confirming the option was sufficient to directly switch to the programme item.

D. Findings and Results

The field trial demonstrated impressively the advantages of the LIVE concept for both consumers and producers. This was proofed by the audience response during and by consumer market research that followed the field trial. The results of the latter were already presented at the beginning of the section. Therefore the following concentrates on the production and system related findings of the field trial.

1) Production related

One of the main tasks of the field trial was to evaluate the LIVE concept and its idea of a novel interactive multi-channel TV format. Besides creating a good and interesting programme, the goal was to establish a dialogue between the audience and the production team. This goal was achieved with the help of the LIVE System and the IPTV platform, allowing the interactive elements to be presented to the consumers and their behaviour and responses be reported back into the control room, especially the poll results that influenced the programme flow more often in an instant operation manner. It proves effective and interesting by asking the consumers for their views on a topic or favourite programmes, which initiates spontaneous polls in accordance with their current topics.

Conducting his team and staging the Olympics was well-versed performed by the director of the productions already on the second production day. Whereas the first day was used to get experience and confidence in concept and technical system, the team produced the interactive TV programme of four interlinked channels in a high quality on the following production days. Only when programme items ended at nearly the same time and therefore more than one "What comes next?"-decision had to be taken the team was a bit more tensed.

Hence the field trial is essentially used to try out direct ways of communicating with the audience. A video telephony tool was used not only to talk or chat with consumers, but also to bring them directly into the show on the screen. Since this idea was innovative, the audience was initially interested and curious. Such attitude changed during the field trial and, in the market research study, it was revealed that this way of communication was predominantly not acceptable. Nonetheless, the overall LIVE concept and especially its interactive elements were positively rated by the consumers throughout the trial, as indicated by both the market research and email reactions.

2) Technical aspects

The LIVE System supports the production team with its research and development results integrated as software tools and services in ORF's production environment. On average, 12000 annotations per hour were created, respectively received and processed by the LIVE System. Despite this enormous load, the system performed well during the productions, which is illustrated by the fact that 254 interactive elements were launched on the IPTV platform, 60 percent of them switches, 25 percents polls and 15 percent information boxes. A total of 11800 consumer feedback messages were returned to the LIVE System, which were then analysed and presented to the video conducting team.

During the field trial, possible improvements of the developed LIVE tools and systems are also closely monitored and analyzed, which also reveals that (i) modular design of the LIVE system helps the consortium to easily identify the strength and weakness of the system; (ii) the field trial arranged not only provides opportunities for the consortium to get a deeper understanding of the TV production process, but also helps to bridge the gap between the media production and technical developments; and (iii) experiences gained from the field trial provide a range of possibilities for further research and improvements on the LIVE concept

V. Conclusions

At present, cutting-edge media organisations support the end users to create a personalised experience, add pre-filtered A/V Content, write comments to a blog and to participate in their own community. No existing example was found in the iTV market yet where media professionals and organisations ask their audience to give feedback during a live mass media production in order to open a live public debate and change the running show accordingly to the audience feedback.

In the field trial of LIVE during the Olympic Games 2008, the public TV Station ORF allowed the audience to get in contact with the live sport director in the control room during the live Olympic Games Broadcast in order to influence the show. The LIVE team realised the world novel concept of "Video Conducting" and the results of the field trial are documented in this journal paper.

The iTV technology enables new interaction possibilities between the TV production and the TV consumers. Through the real-time feedback channel and analysis of feedback, the TV viewer is able to actively influence the content of the live TV program such as transmission of the sport event. This new concept of interactive TV production was tested within EU project "LIVE" during the Olympic Games 2008. The evaluation of the concept included 489 households in Austria which were able to receive multi-channel live coverage of the Olympic Games, produced by ORF and supported by the presented system. The results of the evaluation have shown the great promise of this concept and the acceptance of the presented tools by the professional users inside the chain of TV productions.

The LIVE System is an important advancement on the road to fully interactive and personalized TV. On one hand it brings into the production office the new real-time information on audience behaviour and preferences, which can also be employed for future content planning, audience acceptance level analyzing, and particular media content identifying. On the other hand, the response from the consumers who watch the LIVE channels shows that they really appreciate new interactivity options and that they feel, that through the voting feedback they can really influence the content of the show, and hence they can become more engaged with the TV programmes.

Acknowledgement: the authors wish to acknowledge the funding support as an integrated project, LIVE, under EU IST FP-6 Research Programme (Contract No. IST-4-027312).

References

[1] S. Grünvogel, C. Mac Williams, G. Trogemann, and R. Wages, "Video composer and live video conductor: future professions for the interactive digital broadcasting industry," Proc. 4th European Conference on Interactive Television (EuroITV), Athens, Greece. pp. 32-38, 2006.

[2] R. Wages, S. Grünvogel, J. Zaletelj, C. Mac Williams, and G. Trogemann, "Future live iTV production: challenges and opportunities," Proc. 2nd Int. Conf. Automated Production of Cross Media Content for Multi-Channel Distribution (AXMEDIS), Leeds, UK, pp. 325-328, 2006.

[3] Citizen journalists challenge BBC, by Matthew Eltringham, BBC NewsWatch, 11 April 2006, http://news.bbc.co.uk/newswatch/ifs/hi/newsid_4900000/newsid_4900400/4900444.stm.

[4] http://edition.cnn.com/2009/US/05/01/week.in.irpt/index.html.

[5] http://www.nbcolympics.com/tv_and_online_listings/online.html.

[6] http://news.sky.com/skynews/UK-News.

[7] D. Melendi, R. Garcia, X.G. Paneda, and V. Garcia, "Multivariate distributions for workload generation in video on demand systems," IEEE Communications Letters, vol. 13, no. 5, pp. 348-350, 2009.

[8] H. Shirazi, J. Cosmas, T.J. Owens, Y.H. Song, and A. Centonza, "A QoS monitoring system in a heterogeneous multi-domain DVB-H platform," IEEE Trans. Broadcasting, vol. 15, no. 1, pp. 124-131, Mar. 2009.

[9] M. Lopez-Nores, Y. Blanco-Fernandez et al. "Receiver-side semantic reasoning for digital TV personalization in the absence of return channels," Multimedia Tools and Applications, vol. 41, no. 3, pp. pp 407-436, 2009.

[10] R. Kaiser, M. Hausenblas, and M. Umgeher, "Metadata-driven interactive web video assembly," Multimedia Tools and Applications, vol. 41, no. 3, pp 237-467, 2009.

[11] I. Mas, V. Berggren, R. Jana, J. Murray, and C.W. Rice, "IMS-TV: an IMS-based architecture for interactive, personalized IPTV," IEEE Communications Magazine, vol. 46, no. 11, pp. 156-163, Nov. 2008;

[12] H. Jabbar, T. Jeong, J. Hwang, and G. Park, "Viewer identification and authentication in IPTV using RFID technique," IEEE Trans. Consumer Electronics, vol. 54, no. 1, pp. 105-109, Feb. 2008.

[13] T. Yoshihisa, M. Tsukamoto, and S. Nishio, "A broadcasting scheme considering units to play continuous media data," IEEE Trans. Broadcasting, vol. 53, no. 3, pp. 628-636, Sep. 2007.

[14] Y.P. Huang, L.W. Hsu, and F.E. Sandnes, "An intelligent subtitle detection model for locating television commercials," IEEE Trans. Systems, Man and Cybernetics, Part B-Cybernetics. vol. 37, no. 2, pp. 485-492, April. 2007.

[15] S.J. Lee, S. Lee, K.W. Kim, and J.S. Seo, "Personal and mobile satellite DMB services in Korea," IEEE Trans. Broadcasting, vol. 53, no. 1, pp. 179-187, Part2, Mar. 2007.

[16] J. Kim, H. Kim, and K. Park, "Towards optimal navigation through video content on interactive TV," Interacting with Computers, vol. 18, no. 4, pp. 723-746, Jul. 2006.

[17] R. Goularte, M.D.G.C. Pimentel, and E.D.S. Moreira, "Context-aware support in structured documents for interactive-TV," Multimedia Systems, vol. 11, no. 4, pp. 367-382, Apr. 2006.

[18] W.F. Poon, K.T. Lo, and J. Feng, "Provision of continuous VCR functions in interactive broadcast VoD systems," IEEE Trans. Broadcasting, vol. 51, No. 4, pp. 460-472, Dec. 2005.

[19] B. Nikolaus, J. Ott, C. Bormann, and U. Bormann, "Generalized greedy broadcasting for efficient media-on-demand transmissions," IEEE Trans. Broadcasting, vol. 51, no. 3, pp. 354-359, Sep. 2005.

[20] W. Behrendt, N. Arora, T. Bürger, and R. Westenthaler, "A management system for distributed knowledge and content objects," Proc. 2nd Int. Conf. Automated Production of Cross Media Content for Multi-channel Distribution, December 12-15, 2006, Leeds, UK., 2006.

[21] S. Boll and W. Klas, "Z Y X - a semantic model for multimedia documents and presentations," Proc. 8th IFIP Conf. Data Semantics (DS-8): 'Semantic Issues in Multimedia Systems.' Rotorua, New Zealand, 1999. http://citeseer.ist.psu.edu/boll99semantic.html

[22] T. Bürger, and R. Westenthaler, "Mind the gap – requirements for the combination of content and knowledge," Proc. 1st Int. Conf. Semantic and Digital Media Technologies (SAMT), Athens, Greece, 2006.

[23] A. Gangemi and P. Mika, "Understanding the semantic Web through descriptions and situations". In: On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE, pp. 689-706, 2003

[24] G. Güntner, T. Bürger, D. Glachs, and R. Westenthaler, "Bringing intelligence to iTV: The intelligent media framework," Proc. 6th European Conf. EuroITV, Salzburg, Austria. 2008.

[25] G. Güntner, D. Glachs, and R. Westenthaler, "Utilizing semantics in the production of iTV shows," Proc. European Semantic Web Conf. (ESWC), Heraklion, Greece, pp. 841-845, 2009.

[26] C.A. Macnee, W. Behrendt, et al., "Presenting Dynamically Expandable Hypermedia," Information and Software Technology, vol. 37, no. 7, pp. 339-350, 1995.

[27] C. Masolo et al. "The WonderWeb library of foundational ontologies preliminary report," (WonderWeb Deliverable D17), 2002. Available online: http://wonderweb.semanticweb.org/ deliverables/documents/ D17.pdf (last visited: 31.05.2009)

[28] L.-Y. Duan, M. Xu, Q. Tian, C.-S., Xu, and J.S. Jin, "A unified framework for semantic shot classification in sports video," IEEE Trans. Multimedia, vol. 7, no. 6, pp. 1066-1083, 2005.

[29] S.-F. Chang and A. Vetro, "Video adaptation: concepts, technologies, and open issues," Proc. of IEEE, vol. 93, no. 1, pp. 148-158, 2005.

[30] A. Hanjalic and L.-Q. Xu, "Affective video content representation and modelling," IEEE Trans. Multimedia, vol. 7, no. 1, pp. 143-154, 2005.

[31] N. Dimitrova, "Context and memory in multimedia content analysis," IEEE Multimedia, vol. 11, no. 3, pp. 7-11, 2004.

[32] C. Cotsaces, N. Nikolaidis, and I. Pitas, "Video shot detection and condensed representation: a review," IEEE Signal Proc. Mag., vol. 23, no. 2, pp. 28-37, 2006.

[33] Z. Rasheed and M. Shah, "Detection and representation of scenes in videos," IEEE Trans. Multimedia, vol. 7, no. 6, pp. 1097-1105, 2005.

[34] C. Kim and J. N. Hwang, "Object-based video abstraction for video surveillance systems," IEEE Trans. Circuits Syst. Video Technol., vol. 12, no. 12, pp. 1128-1138, 2002.

[35] H.J. Zhang, S.Y. Low, and S.W. Smoliar, "Video parsing and browsing using compressed data," Multimedia Tools and Apps., vol. 1, no. 1, pp. 89-111, 1995.

[36] V. Kobla, D. Doermann, and K.-I. Lin, "Archiving, indexing, and retrieval of video in the compressed domain," In: Proc. SPIE, vol. 2916, pp. 78-89, 1996.

[37] J. Ren, J. Jiang, and J. Chen, "Shot boundary detection in MPEG videos using local and global indicators," IEEE Trans. Circuits and System for Video Technology, vol. 19, no. 8, pp. 1234-1238, 2009.

[38] H. Fang, J. Jiang, and Y. Feng, "A fuzzy logic approach for detection of video shot boundaries', Pattern Recognitions, vol. 39, no. 11, pp. 2092-2100, 2006.

[39] Y. Weng and J. Jiang, "Real-time and automatic close up retrieval from compressed videos," Int. J. Automation and Computing, vol. 5, no. 2, pp. 198-201, 2008.

[40] J. Jiang, Y. Weng, and P. Li, "Dominant Colour Extraction in DCT Domain," Image and Vision Computing, vol. 24, pp. 1269-1277, 2006.

[41] J. Jiang and G. Feng G, "Robustness analysis of facial image description in DCT domain," Electronic Letters, vol. 43, no. 24, pp. 1354-1355, 2007.

[42] J. Ren, J. Jiang, J. Chen, and S. S. Ipson, "Extracting objects and events from MPEG sequences for video highlights indexing and retrieval," To appear in Journal of Multimedia (JMM), Academy, vol. 5, no. 2, 2010.

[43] Y.P. Tan, D.D. Saur, S.R. Kulkarni, and P.J. Ramadge, "Rapid estimation of camera motion from compressed video with application to video annotation," IEEE Trans. Circuits Syst. Video Technol., vol. 10, no. 1, pp. 133-146, 2000.

[44] S. L. Phung, A. Bouzerdoum, and D. Chai, "Skin segmentation using colour pixel classification: analysis and comparison," IEEE Trans. Pattern Anal. Machine Intell., vol. 27, no. 1, pp.148-154, 2005.

[45] L. Sigal, S. Sclaroff, and V. Athitsos, "Skin colour-based video segmentation under time-varying illumination," IEEE Trans. Pattern Anal. Machine Intell., vol. 26, no. 7, pp. 862-877, 2004.

[46] M.J. Jones and J.M. Rehg, "Statistical colour models with application to skin detection," Int. J. Computer Vision (IJCV), vol. 46, no. 1, pp. 81-96, 2002.

[47] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-colour modelling and detection methods," Pattern Recognition, vol. 40, no. 3, pp. 1106-1122, 2007.

[48] N. Habili, C. C. Lim, and A. Moini, "Segmentation of the face and hands in sign language video sequences using colour and motion cues," IEEE Trans. Circuits Syst. Video Technol., vol. 14, no. 8, pp. 1086-1097, 2004.

[49] B. Leibe, K. Schindler, N. Cornelis, and L. Van Gool, "Coupled object detection and tracking from static cameras and moving vehicles," IEEE Trans. Pattern Anal. Machine Intell., vol. 30, no. 10, pp. 1683-1698, 2008.

[50] P. Viola, M. J. Jones, and D. Snow, "Detecting pedestrians using patters of motion and appearance," Int. J. Comput. Vision (IJCV), vol. 63, no. 2, pp. 153-161, 2005.

[51] Y.-T. Chen and C.-S. Chen, "Fact human detection using a novel boosted cascading structure with meta stages," IEEE Trans. Image Proc., vol. 17, no. 8, pp. 1452-1464, 2008.

[52] J. Jiang, K. Qiu and G. Xiao, "An edge block content descriptor for MPEG compressed videos," IEEE Trans. Circuits, Systems for Video Tech. vol. 18, no. 7, pp 994-998, 2008.

[53] J. Ren and J. Jiang, "Hierarchical modelling and adaptive clustering for real-time summarization of rush Videos," IEEE Trans. Multimedia, vol. 11, no. 5, pp. 906-917, 2009.

[54] C. W. Ngo, Y. F. Ma, and H.-J. Zhang, "Video summarization and scene detection by graph modelling," IEEE Trans. Circuits Syst. Video Technol., vol. 15, no. 2, pp. 296-305, 2005.

[55] X. Zhu, J. Fan, A. K. Elmagarmid and X. Wu, "Hierarchical video content description and summarization using unified semantic and visual similarity," Multimedia Systems, vol. 9, no. 1, pp. 31-53, 2003.

[56] Y. Weng and J. Jiang, "Fast camera motion estimation in MPEG compressed domain," to appear in IEEE Transactions on Circuits and Systems for Video Techn., 2009.

[57] J. Ren and J. Jiang, "Extracting semantics and content adaptive summarisation for effective video retrieval," to appear in Proc. NEM Summit, 2009.

[58] J. Yuan, H. Wang, L. Xiao, W. Zheng, J. Li, F. Lin, and B. Zhang, "A formal study of shot boundary detection," IEEE Trans. Circuits Syst. Video Technol., vol. 17, no. 2, pp. 168-186, 2007.

[59] G. Boccignone, A. Chianese, V. Moscato, and A. Picariello, "Foveated shot detection for video segmentation," IEEE Trans. Circuits Syst. Video Technol., vol. 15, no. 3, pp. 365-377, 2005.

[60] J. Bescos, G. Cisneros, J.M. Martinez, J.M. Menendez, and J.Z. Cabrera, "A unified model for techniques on video-shot transition detection," IEEE Trans. Multimedia, vol. 7, no. 2, pp. 293-307, 2005.