



Format-agnostic Approach for Production, Delivery and Rendering of Immersive Media

G. Thallinger, JOANNEUM RESEARCH, Graz, Austria

O. Schreer, Fraunhofer Heinrich-Hertz Institute, Berlin, Germany

G. Thomas, BBC R&D, London, UK

O.A. Niamut, TNO, Delft, The Netherlands

J-F. Macq, Alcatel-Lucent, Antwerp, Belgium

A.Kochale, J-M. Batke, Image Processing Lab, Technicolor, Hannover, Germany

J. Ruiz Hidalgo, UPC, Barcelona, Spain

R. Oldfield, B. Shirley, University of Salford, Manchester, UK

*The research leading to these results has received funding from the
European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 248138*

Today's Dilemma

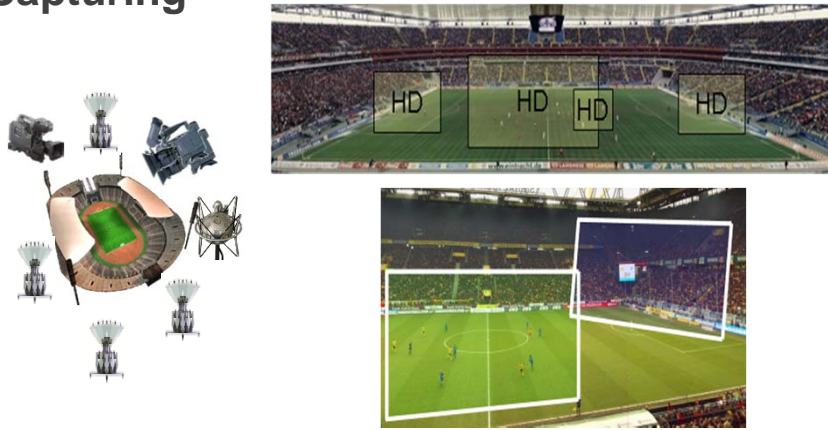
- Multitude of formats & terminals leads to:
 - Complex and expensive A/V capture and production infrastructure
 - Inefficient distribution networks employing multicast
 - Limited or no user interaction, suboptimal AV quality, poor user experience

FascinatE's Approach

- Format agnostic production framework
 - Covering multitude of sensors
 - Employing a layered production framework
 - Enabling rich repurposing
- Shift from image capture to scene capture
- Use of scripting system enabling derivation of multiple simultaneous programs for various terminals
- Development of distributed network infrastructure for intelligent caching, proxying and script processing
- Enabling content-based scalability
- Immersive Experience
 - Interactive navigation in captured scene
 - Immersive consumer experience
 - Highest quality by optimized adaptation to individual terminal capabilities

Immersive Media – Our Research Area's

Layered Panoramic & Omnidirectional A/V Capturing

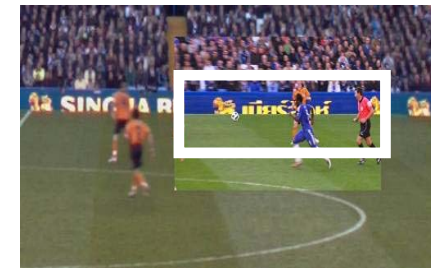


Video Analysis and Automated Editing

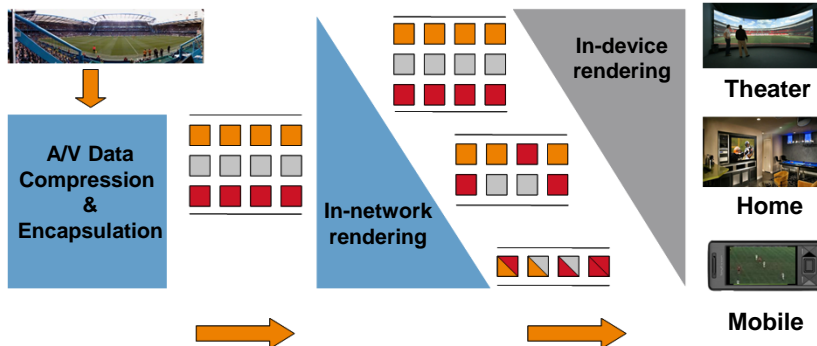


Automated shot framing

Region-of-Interest Detection and Tracking



Scalable Delivery and in-network Adaptation of A/V flows

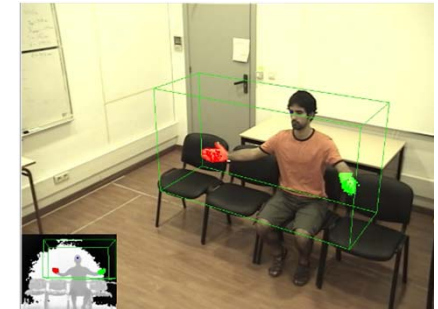


Immersive and Interactive Applications

Flexible and Interactive A/V Rendering



Gesture-based user interfaces



Terminal Categories

**Case 1:
Larger Displays for
Public Viewing**



**Case 2:
More Natural Images
@ Home**



**Case 3:
Mobile Terminals**



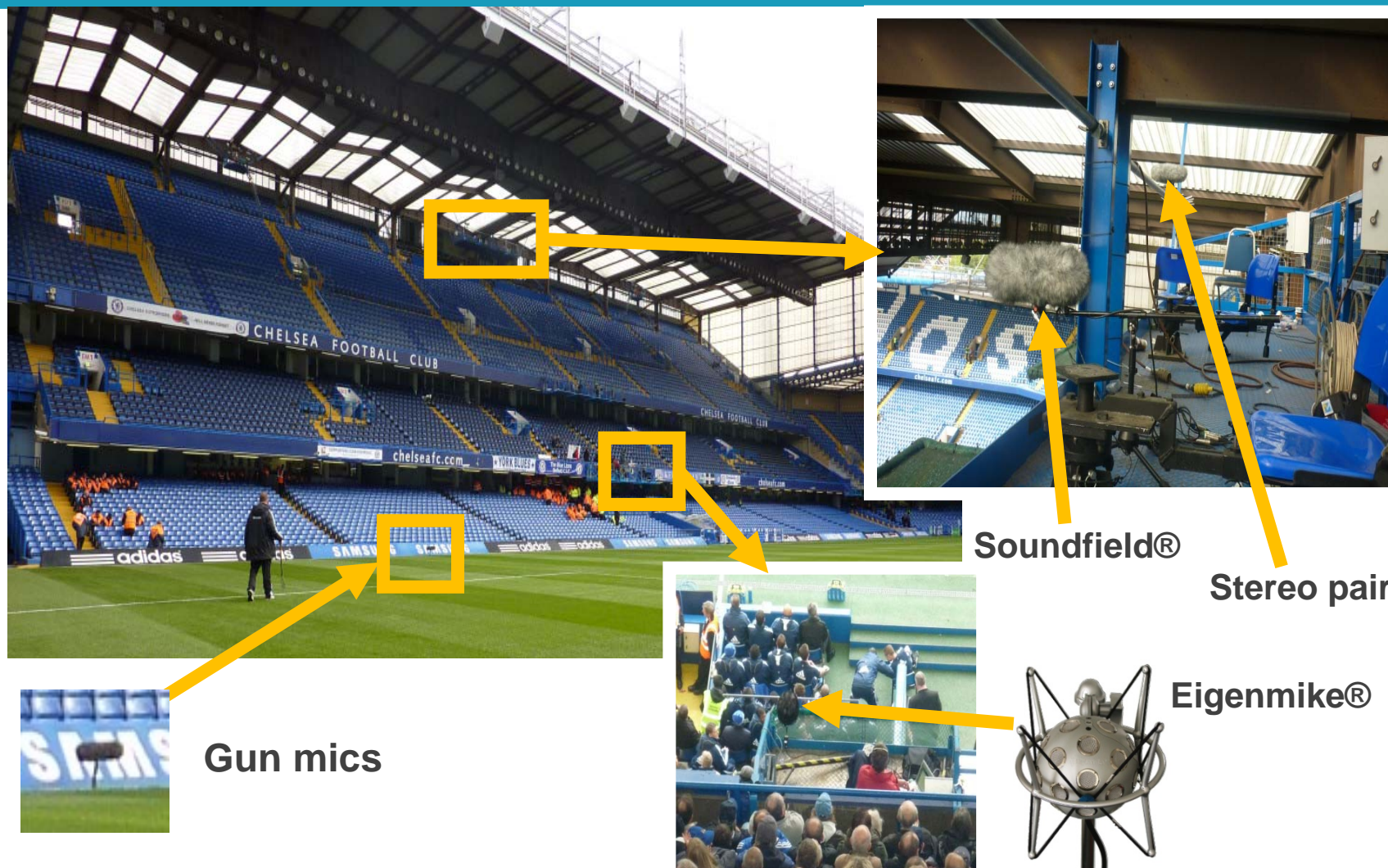
with Interactive Controllers



Video Acquisition



Audio Acquisition



Real-time live panoramic video



Real-time live panoramic video

- First demonstration at IBC 2011
 - of real-time acquisition,
 - warping,
 - stitching,
 - blending and
 - rendering of
 - 6 HD video streams
 - into one panoramic video
 - of 7k x 2k resolution
 - on a single PC exploiting GPU processing

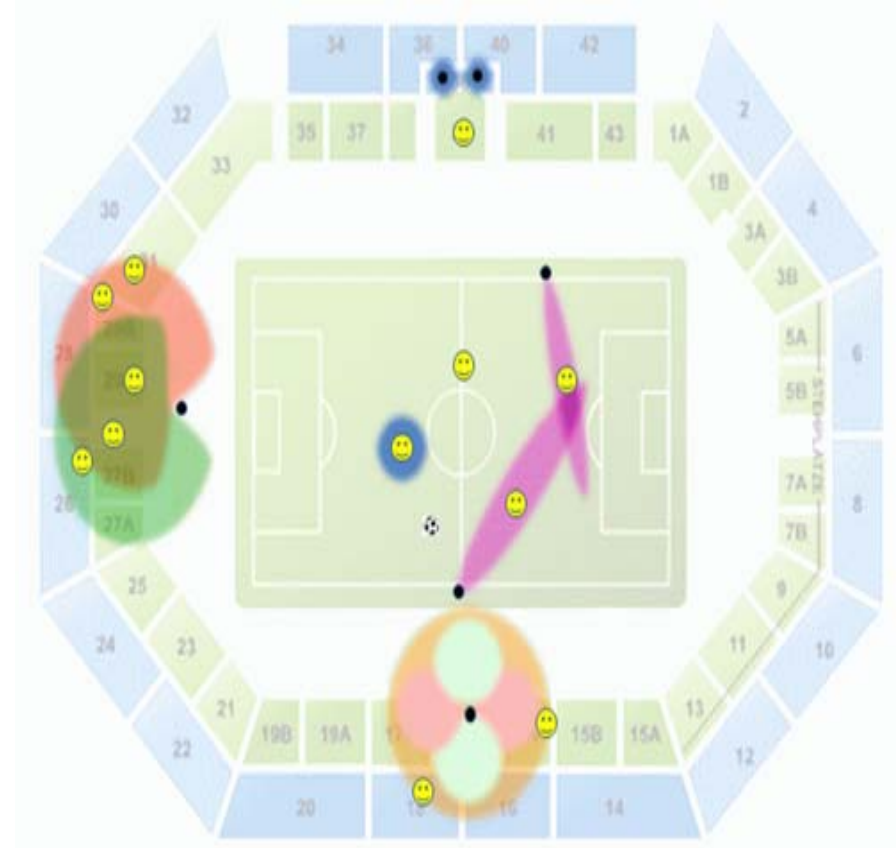


Panoramic & broadcast cameras compared



FascinatE Audio

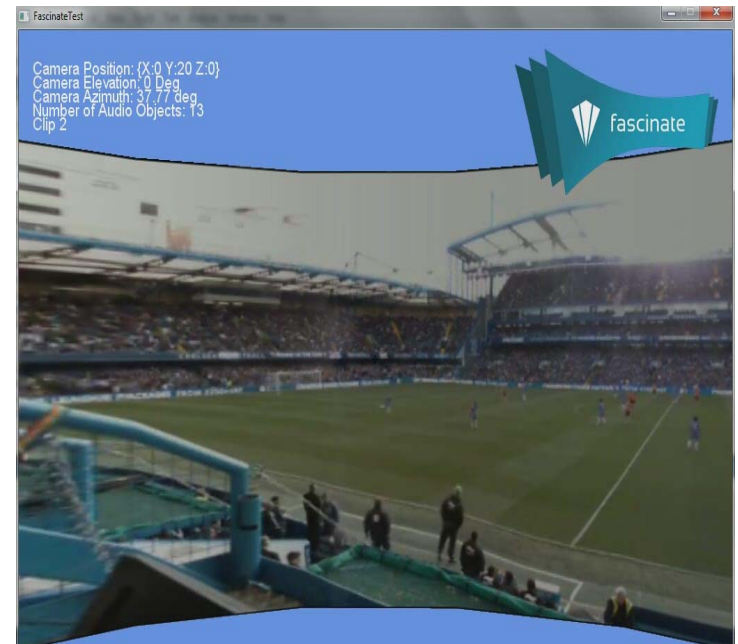
- Object based approach
 - Audio objects
 - Close miked and tracked sources
 - Derived from multiple microphones
 - Ambient sound field component
 - Ambisonics Microphones
 - Eigenmike
 - SoundField microphone
 - Stereo Microphones



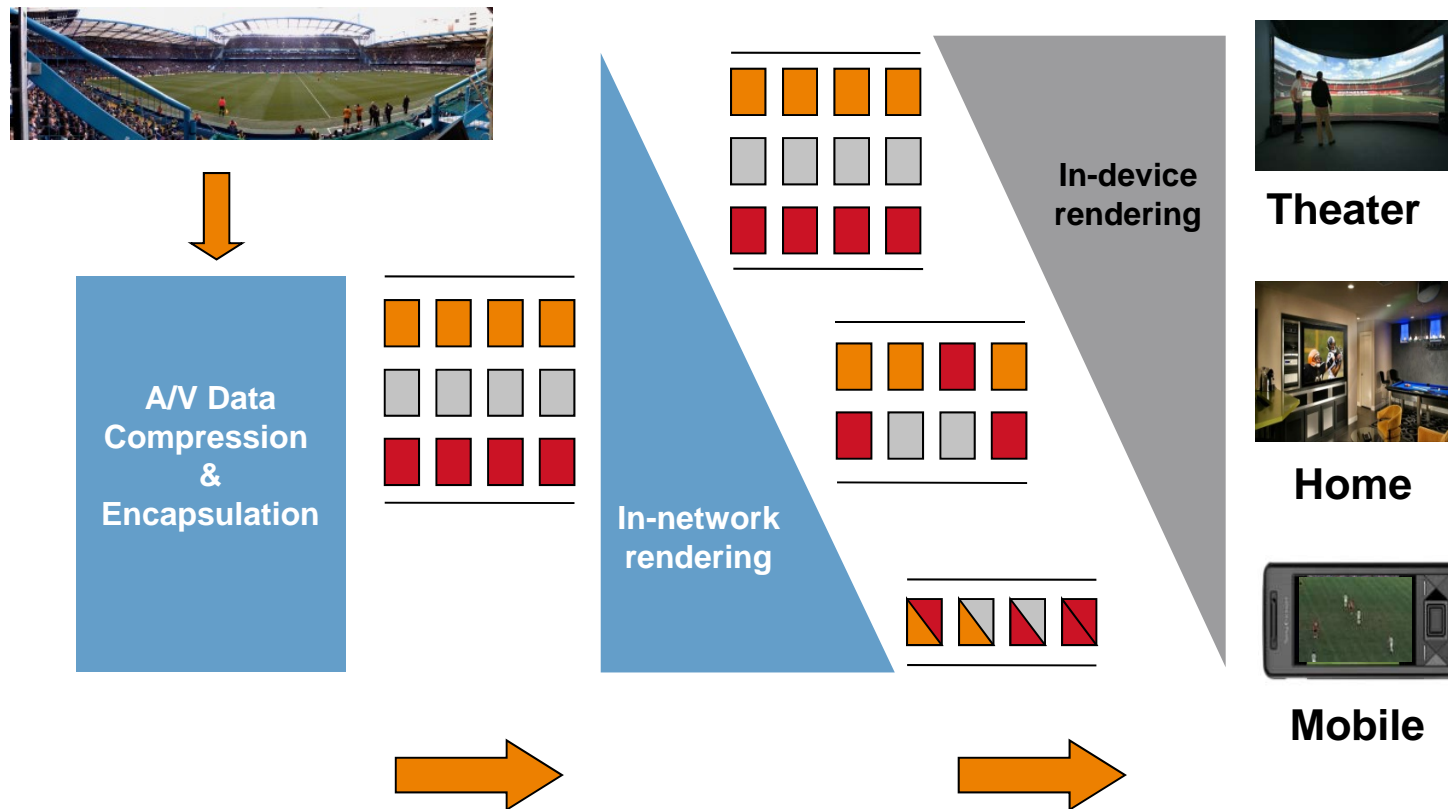
- 😊 Sound Source
- Microphone

Audio Rendering Preferences

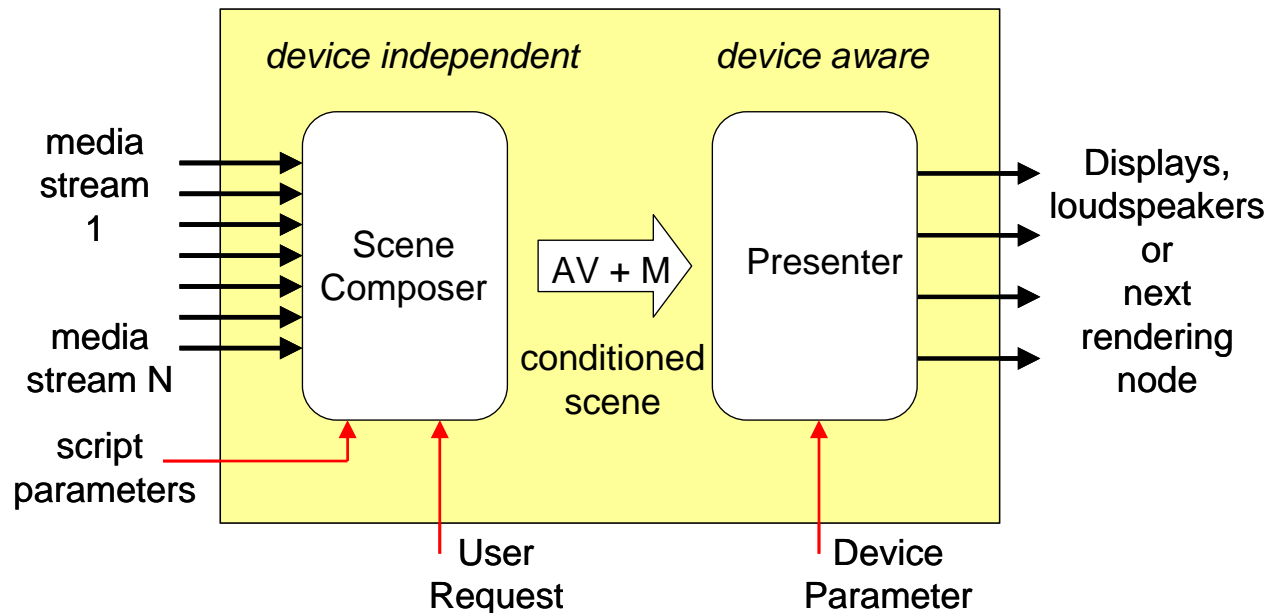
- Audio Rendering
 - Shift in viewer expectation?
 - Passive Viewer vs. Active Participant
- Pilot Test
 - Audio objects placed in 3D model
 - Ambient sound field from Eigenmike®
 - User can pan around scene from fixed position
 - Audio rendering options
 - Static
 - Rotating sound field and audio objects
 - Results
 - Preference towards rotating sound field



Scalable Delivery Networks and Network-based Rendering



Flexible and Interactive Video Rendering

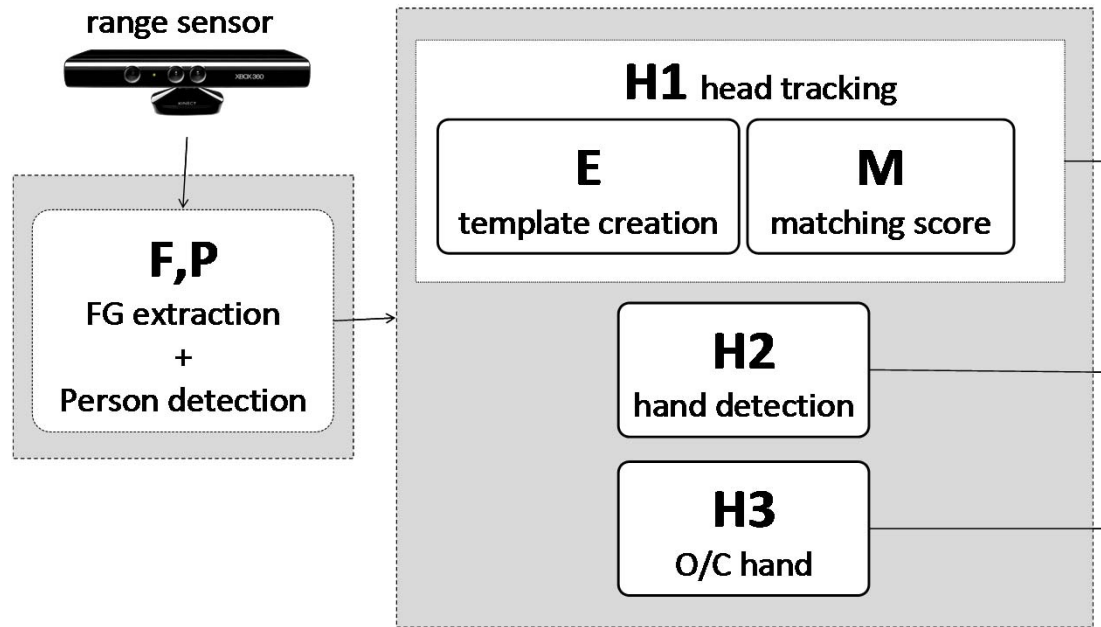


- Content interaction beyond channel switching: Pan/Tilt/Zoom (PTZ)
- Real-time end-device platform for PTZ navigation within a panoramic scene
- Rendering converts format-agnostic A/V into personalized perspectives on end user screens
- XML script controls rendering, easy personalization by selecting script options such as regions of interest

Flexible and Interactive Video Rendering



Gesture-Based User Interaction



- Fast and robust head + hand tracking using depth information from a range sensor, suitable for interactive and immersive applications.
- H1: Head Tracking, Elliptical template resizing and matching score is computed
- H2: Hand detection in 3D hand workspace
- H3: Open/closed hand detection by determining physical hand size

Gesture-Based User Interaction



Conclusion

- The FascinatE project is investigating the practicality of a format-agnostic production system
- Format-agnostic audio-visual acquisition concept has been proven on test shoot at Premier league soccer game at FC Chelsea
- 1st successful demonstration of real-time modules of the FascinatE system at IBC 2011
- FascinatE: provides an interactive experience for immersive media
 - Script-based automation and view selection
 - Gesture-controlled view rendering
 - Scalable network-based tiled streaming and media adaptation
- **Visit our booth in the NEM Summit 2011 exhibition**
- **Visit our web site: www.fascinate-project.eu**