Non-linear Video

A cross-platform interactive video experience

Robert Seeliger
Future Applications and Media
Fraunhofer Institute FOKUS
Berlin, Germany
robert.seeliger@fokus.fraunhofer.de

Christian Räck
Future Applications and Media
Fraunhofer Institute FOKUS
Berlin, Germany
christian.raeck@fokus.fraunhofer.de

Dr. Stefan Arbanowski
Future Applications and Media
Fraunhofer Institute FOKUS
Berlin, Germany
stefan.arbanowski@fokus.fraunhofer.de

***Abstract* - Non-linear video is an approach, which makes video content an interactive experience. Non-linear video gives the viewer the opportunity to interact with objects that are part of the video and access supplemental information. On demand, multimedia content is linked with related information. Interactive, time independent navigation opens a new ways to experience video content. This paper shows how such an system could be built upon IPTV and web technology in a cross platform manner.**

Interactive video; non-linear content; user interaction; object description; personalization; IPTV

I. INTRODUCTION

Internet Protocol (IP) based media services, such as Internet Protocol Television (IPTV) [1], are different from conventional TV technology. Via IPTV, television content will be viewed and delivered through technologies used for computer networks. This opens to a wide range of new media services and asks for innovative advertisements and content patterns to satisfy the ever increasing demand of the advertisement industry as well as content production for predefined interfaces to place product and content related information and advertisements in close vicinity to particular objects within the video.

Advertisements in the form of pre-rolls, post-rolls, and overlays are becoming increasingly ineffective as more and more consumers decide to skip these commercials due to the unimportance of the offered information and low individual involvement and interest in the offered product. This fact is boosted by the technological and economic evolutions in digital Television (TV) and media business areas. There is a strong need for new marketing strategies, innovative advertisements and finally a need for new kinds of interactive video content to fit these requirements [2].

Non-linear video concerns the delivery of personalized interactive value added information and related videos to end-users. Such an interactive item could be a video clip that can be paused at any time. Thus the introduced technology enables customers to decide when and which related information, interactive items and advertisements are displayed. In those environments customers can pause each video content at any time.

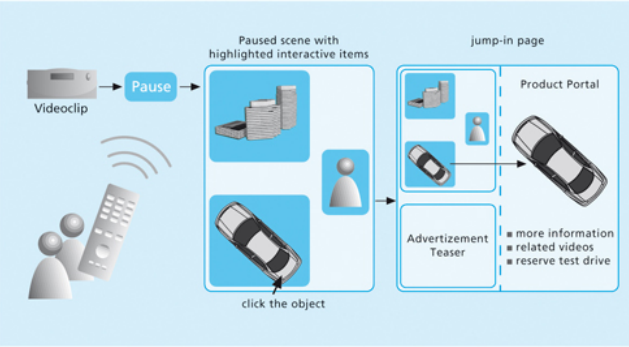

Figure 1. Non-linear video content interaction

Once paused, several objects in that current scene are automatically highlighted. Each highlighted object displays additional information, such as a detailed description and links to other objects or related content, when clicked. The displayed objects, descriptions and links are personalized based on previously learnt end-user profiles. The Non-linear video system identifies which objects are relevant to a particular customer based on his or her current situation and then only displays these objects for a personalized and interactive experience. Figure 1 shows, how different elements, such as a piece of content, a description of the object shown there, and content recommendations are linked to create a new interactive content experience.

II. OVERVIEW

Non-linear video offers the platform for realizing interactive and personalized multimedia content. Videos are linked to related information thus making time independent navigation possible. The linear character of traditional moving pictures and video formats are enhanced with non-linear video towards multiple ways of interaction. The user can navigate at any time through objects, which are contained in the video content such as TV programs and films. As soon as the viewer clicks on an object that interests them, supplemental interactive information or video is displayed. This can be any multimedia content (image, text, animation, video, pdf), websites, or alternative communication methods such as telephone, chat, email as

 ISBN: 978-1-61208-110-6

well as Web 2.0, community and social media services, which can be shown on a time basis or in subject to the user. Non-linear video supplements the previous available technology of pre and post rolls, commercial breaks, product placement, and overlays. Videos become clickable using links to content and additional information which are brought up parallel to the existing moving pictures. Metadata describe the objects in the video as well as possible ways to interact with it.

III. RELATED WORK

As the history of the last ten years and recent developments on interactive multimedia content and related TV or Web-based solutions have descriptive shown, interactive video content pines for easer solutions. Whether previous systems [3] seem to provide all technical issues and apparently fulfill all requirements for those interactive multimedia experiences, in fact the bulk of them disappear due to lack of usability. This includes both, the user side to experience value added services and real interactivity, and the service providers modality to serve such interactive content offers. Those solutions were mostly based on large and complex metadata descriptions as MPEG7 [4], related interactive TV technologies as MHP [5] or more theoretical approaches on top of MPEG4 [6] video scene and object descriptions. Unfortunately these approaches do not affect any products and services on the media market till now. In contrast to the depicted history of interactive video, the envisaged Non-linear video solution tackles a more applicable and practicable route to enable interactive content based on a media platform to serve videos, and the related metadata to provide object descriptions and the associated information as well.

The realization of the above-mentioned vision in section I require the exploitation of a number of ideas, which have been developed in different technology domains. Ideas taken from interactive and non-linear video utilization build the basis for the definition of the new interactive content features as described above. Results taken from previous projects and developments, which dealt with the realization of a recommendation system [7], are used to provide the required personalization features. A central component of the described platform is the Interactive Video Player that utilizes results from different technology domains, such as recommendations, next-generation video platforms [8] as well as standardized IPTV infrastructures [9] to provide the content including the interactive items to end-users.

IV. USE CASE

Following, a usage scenario for interactive content-based on the Non-linear video technology will be introduced. The 'Berlin Tourism' usage scenario can be categorized as a scenario for an interactive tourism information system. In this scenario a local tourism centre wants to promote their special offer for a weekend trip to Berlin. The campaign includes a series of short video clips about the most famous sights of Berlin, which are delivered through our envisioned

Figure 2. Browser based player for interactive content

new interactive non-linear video platform. The tourism centre links detailed tourist information to the sights and identifies a number of additional objects in the videos. Affiliates of the tourism centre are allowed to place their own information and related content on the previously identified objects.

Affiliates of the tourism centre select the content that is provided by the tourism centre to feature their own products and services on the tourism centre website. The *traveller* is an end-user, who visits the tourism centre website looking for a place to go on his vacation. He will take advantage of the interactive service in terms of being able to plan his trip to Berlin and get informed in detail by watching and navigating through the interactive videos using the non-linear video technology on his mobile phone, Laptop or TV at home.

V. HOW CONTENT BECOMES INTERACTIVE

Traditionally, videos and TV content have been created to be consumed passively. Using non-linear video, TV and multimedia content is made to be experienced interactively. This technology creates a seamless transition between additional information, valued added services and video content. The resulting interactive content functions just like a website. Individual sections can be annotated and linked to continuous content such as text, images, video and links. This information is represented by XML based metadata and is made available using an interactive video player. Web-Standards are employed as the access control technology. The FOKUS Tagging-Tool enables the annotation of the raw video data. This intelligent software supports the editor by identifying relevant objects and sections of a video-scene, placing information and by highlighting. The tagging tool is designed to be used as a web-based solution and can be conveniently utilized within the browser. In addition to the time and spatial data, you can also add keywords to describe objects such as the type, category or kind of interaction.

VI. CROSS PLATFORM INTERACTIVE VIDEO UTILIZATION

Non-linear video works with many of the user devices available today. The technology of Fraunhofer FOKUS enables the convergent use of interactive video content on TV, the web as well as mobile phone. This unique

 ISBN: 978-1-61208-110-6

environment makes interactive media available regardless of

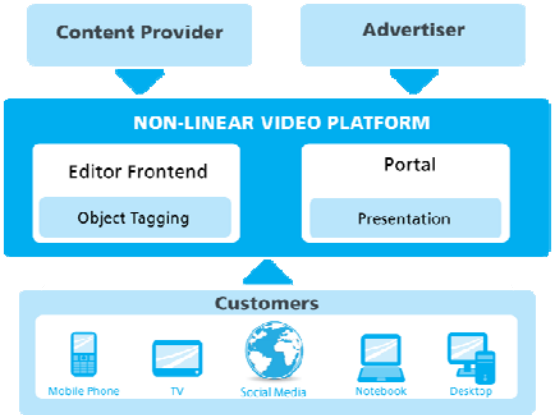

Figure 3. Cross platform approach

the end device and access platform.
The different ways for the viewer to interact, the graphic presentation of the supplemental information as well as the content itself all come in the best possible format for any type of user device. Non-linear video thus offers diverse variations of both the visual aspects of and interaction with the content. As a result, it can offer interactive content, relevant additional information and communication channels custom-made for both the user and the device.

VII. ARCHITECTURE

This section provides a brief overview on the overall architecture of the envisioned interactive media platform and its downstream marketplace. The envisioned architecture is a three-tier architecture that consists of a rich content "media player component" and a video tagging and annotation toolkit both in the presentation tier, a set of backend components in the application tier and one or more servers in the data tier.

A. *Building blocks*

The high-level architecture of the platform consists of six different building blocks including sub modules:

- Interactive Media Player (client side)
- Object Tracking and Linking Tool (client-side)
- Media Server (server-side)
- Recommender Server (server-side)
- Marketplace (server-side)
- Advertising Server (server-side

Fig.4 shows the high-level architecture and the above-mentioned building blocks and their functionality.

B. *Workflow description*

The depicted technology for user-initiated content interaction based on our Non-linear video technology refers to the following set of features:

- The raw video is enriched by metadata.
- Users can initiate a content interaction session at any time.
- Prepared objects will be highlighted for interaction.
- Consumers can select several objects to get further information (dive into the content).

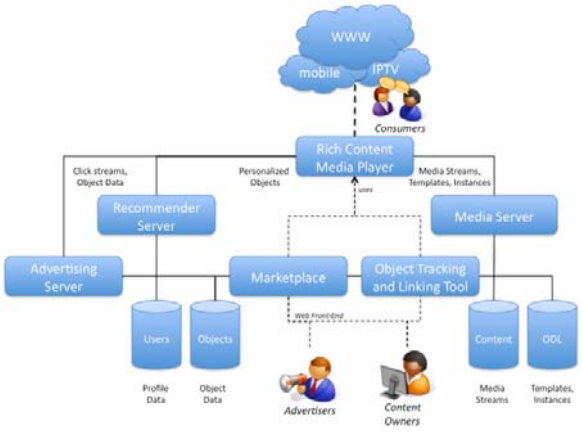

Figure 4. Overall system architecture

The additional information, which is displayed when a highlighted object is selected, provides an added value to the user. In this way, he is capable of getting a description of the highlighted object as well as links to related content of any type. Navigating the hyperlinked content spawns a hierarchical object tree, which represents the user's particular interests at a given moment in time. In contrast to traditional content, which has been produced for linear and non-interactive TV, the user is free to choose what content he wants to consume now and what will be next – the user goes interactive. In a first step, the raw video-content has to be analyzed and annotated. This results in an identification and description of the objects in the scenes (e.g., a car, watch, jacket, sight). The outcome of this procedure is metadata information describing the video content. Advertisers and content producers can place product related information and interactivity by annotating the predefined objects. After completion of the annotation and aggregation step, the raw video content is enriched with interactive items for user interaction. If the user initiates a pause while watching the video, the rich media player highlights all objects within the current scene, which have been pre-annotated for user interaction. The user is now skilled to interact directly with the interactive items. Based on the given metadata information, various possibilities are conceivable, e.g., link to product portals, direct shopping, product related add-on information and nonlinear video scene navigation.

 ISBN: 978-1-61208-110-6

VIII. REFERENCE IMPLEMENTATION

A. Interactive Media Player

The Interactive Video Player component displays interactive videos, which are delivered in the format that has been defined as Non-linear video. The media player highlights all objects, which have been previously identified and which are described in the corresponding metadata. When an end-user clicks on an object a new user interaction session is started and all actions that are linked to the selected object are triggered. To ensure a high attention rate among the target audience objects can be filtered or differently coloured based on their predicted relevance. For that, a recommender system is used to calculate the relevance of all objects in this scene for a particular end-user. Fig. 5 shows the implementation of these interactive media player component in terms of our mobile solution running on Apples iPhone. Other implementations for Web-Browser and Hybrid TV sets are also available. The system provides multiple interaction layers, which enables the user to leave the current video by watching the next clip about an object in a hierarchical way, he can navigate throw the media in a non-linear manner.

Figure 5. Non-linear video implementation for smartphones

B. Object identification, tracking and linkage tool

The Object Tracking and Linking Tool is used to identify all objects in a video. This is a three-step process: At first an object template has to be defined for each object that should be tracked in a video. Next a new instance has to be created for each scene that contains this particular object. At last the object has to be tracked in each scene. The object template contains the definition of the tracking shape, a link to an HTML page that provides further information about the object and a list of related Web links. Each object instance contains information about a particular scene in which the object is visible. Single frames numbers and the position of the tracking shape in the video frames are recorded.

The output of the tracking and linking process is the object template definition and the list of object instances for all objects in a video. This metadata is described in an XML-based *Object Definition Language (ODL)*, which has been adopted on the experience of XML based metadata as TV Anytime [10] [11]. The information, which is described in the ODL, is used by the Interactive Video Player to enable the end-user interaction as described in the previous paragraph.

In the following example, implementations of metadata separate from the video data are given, i.e., metadata and video data exist as separate files and database entries which are combined in the video data player. This has the advantage that the video data can remain unchanged. The video data player has to load the video as shown in "example.obj" in the sample code of the ODL given below. Further, the video player has to load the description of the identified objects in the video data and the respective linking.

C. Object Definition Language

The following sample XML-code describes the object identification. For each object that a user can interact with, a

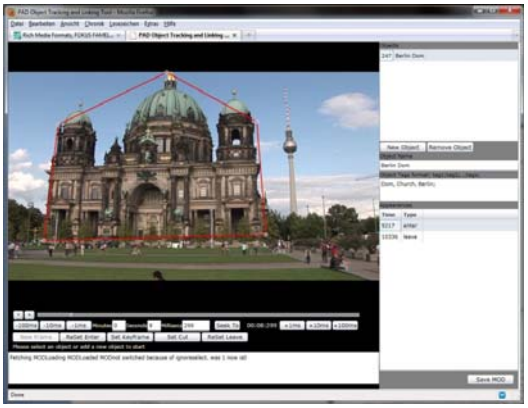

Figure 6. Web-based user interface of tagging toolkit

set of bounding boxes is defined. Each bounding box is

```
<?xml version="1.0" encoding="utf-16"?>
<MediaObjectDescription xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance" xmlns:xsd="http://www.w3.org/2001/XMLSchema">
 <ObjectAppearances>
  <ObjectFrame>
   <ShapeTimeFrames>
    <ShapeInTime>
     <TimeFrame>3769</TimeFrame>
     <Points>
      <RelPoint>
       <X>0.48125</X>
       <Y>0.26379310344827589</Y>
      </RelPoint>
     </Points>
   </ShapeTimeFrames>
   <Object>
    <ObjectID>94</ObjectID>
    <Name>Berlin Cathedral</Name>
    <Tags>Berlin;Point of Interest;Cathedral;</Tags>
    <PredictedRelevance>0</PredictedRelevance>
    <LinkedAdvertisement>
     <AdvertisementID>38</AdvertisementID>
     <Name>Berlin Cathedral</Name>
<Text>The Berlin Cathedral…</Text>
<PictureURI>http://sampleURL/samplepic.jpg</PictureURI>
      <URI>http://samplevideo.com;feature=related</URI>
    </LinkedAdvertisement>
   </Object>
  </ObjectFrame>
 </ObjectAppearances>
 <MediaID>15</MediaID>
 <MediaURI>http://sampleuri/samplevideo</MediaURI>
 <Owner>
  <CoID>5</CoID>
  <Name>tester</Name>
  <CoTags />
 </Owner>
</MediaObjectDescription>
```

 ISBN: 978-1-61208-110-6

specified with its type (e.g., “a rectangular box”) and size. Then the set of frames where a particular object is shown is identified. Last, but not least transitions of the bounding boxes between individual frames are defined.

IX. CONCLUSION

As opposed to traditional video, which limits interaction and use due to its linear nature, interactive video opens up new and diverse possibilities. FOKUS non-linear video supports multiple levels of interaction in the content itself as well as between the interactive objects and the supplemental information available. The viewer can access interactive video content using three basic levels of interaction:

- Moving picture content (video material)
- Interactive objects in video (text, audio, video, image, web)
- Communication channels (telephone, chat, email, web 2.0, social networking)

Whereas the level of moving picture content sparks someone’s interest in a topic or product, the newly created levels enable customization of content and make detailed information about the objects in the video available. In combination the interactive possibilities, that are carried out to produce a non-linear story line that, can be freely chosen and interacted with by the viewer.

The linking of information using non-linear video can be applied ideally for multistep information access and business models: from free use, which grants access to general information to registration sites as well as premium content access through a pay system. Since interaction in today’s video platforms is limited to commenting on posts and placing advertising banners, the non-linear video technology offers new possibilities for interactive video use. Multimedia data, video objects and additional information are linked to interactive content and enriched by customized object interaction. Related information, communication channels and content are all customized for the user and his end device. Supplemental information enables advanced applications such as interactive advertising, video portals and edutainment.

Fostered by the innovative technology and simultaneously upcoming opportunities, manufacturers and CE-Industry will be empowered to open up new markets. Interactive media services become reality and will boost the user experience to a higher level. The current solution provides interactive video content via rich internet application technology as MS Silverlight [12]. The integration to our ETSI TISPAN [12] and Open IPTV Forum standards compliant Open IPTV Ecosystem [14] has already been done. It shows how interactive content may be used in future IP based TV, and Hybrid TV environments [15]. Next steps will introduce more differentiated interaction models for identified interactive object items within the videos. This will imply enhancements on the object description language, the tagging toolkit and the associated interactive media players for the depicted various target platforms.

References

[1] O. Friedrich, R. Seeliger, and A. Al-Hezmi, „Prototyping interactive and personalized IPTV-services on top of open IMS infrastructures” 6. EuroiTV2008, Salzburg Austria

[2] T. Haaker, N. Ito, S. Smit, J. Vester, M.G lbahar, K. Shepherd, and J. Zoric, Business Models for Networked Media Services, 7th European Interactive TV Conference (EuroITV 2009), Leuven, Belgium, June 2009

[3] Hang Nguyen, Julien Royer, Durgesh O Mishra, Olivier Martinot, and Françoise Preteux, “Atomatic generation of explicitly embedded advertisement for interactive TV: concept and system architecture”

[4] MPEG-7 Overview (v.10), ISO/IEC JTC1/SC29/WG11 N6828 Palma de Mallorca, October 2004

[5] ETSI ES 201 812 V1.1.2 (2006-08) Digital Video Broadcasting (DVB);Multimedia Home Platform (MHP) Specification 1.0.3

[6] MPEG-4 BIFS tutorial http://gpac.sourceforge.net/tutorial/bifs_intro.htm

[7] Christian Räck, Stefan Arbanowski, and Stephan Steglich, “A Generic Multipurpose Recommender System for Contextual Recommendations”, 8th Intern. Symposium on Autonomous Decentralized Systems (ISADS 2007), Sedona, Arizona, March 2007

[8] Draft ETSI RTS 182 027 V3.1.0 (2008-11). Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); IPTV Architecture: IPTV functions supported by the IMS Subsystem

[9] Open IPTV Forum Release 1 Specification: Volume1 Overview V1.0 available at http://www.oipf.tv

[10] TV-Anytime Forum, http://www.tv-anytime.org/, 2010-06-10

[11] ETSI TV Anytime XML Schema: ETSI TS 102 822

[12] MS Silverlight Rich Internet Application Technology http://silverlight.net/

[13] ETSI ES 282 001: "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); NGN Functional Architecture Release 2

[14] Fraunhofer FOKUS Open IPTV Ecosystem, http://www.fokus.fraunhofer.de/en/fame/_pdf/FOKUS-IPTV-Ecosystem.pdf, 2010-06-09

[15] Philips Net TV, press information http://www.digitalnewsroom.philips.com/press/nettv-pr.doc, 2010-06-11

 ISBN: 978-1-61208-110-6