TOWARDS NEXT GENERATION HYBRID BROADCAST BROADBAND, RESULTS FROM FP7 AND HBBTV 2.0

M.O. van Deventer[1], J.J. de Wit[1],
M. Guelbahar[2], B. Cheng[3], F. Gomez Marmol[3], C. Köbel[4],
C. Köhnen[4], G. Rozinaj[5], B. Stockleben[6]

[1]TNO, Netherlands; [2]IRT, Germany; [3]NEC Laboratories Europe, Germany;
[4]THM, Germany; [5]STUBA, Slovakia, [6]RBB, Germany

ABSTRACT

Hybrid Broadcast Broadband Television (HbbTV) is a widely-supported technology that enables broadcasters to run dedicated smart TV applications as a browser overlay on their TV channels. Next generation HbbTV is being developed in the EU-funded HBB-Next project, and being contributed to the standardization process of HbbTV 2.0.

HBB-Next has designed and implemented several next-generation HbbTV frameworks, including multi-user identification, synchronisation between media streams and devices, multi-user recommendations, user-tailored reputation scores for applications and cloud offloading. Integrated proofs-of-concepts are presented in the IBC Future Technologies zone.

The HbbTV standards body has selected a wide set of features for HbbTV 2.0 in four categories: browser, terminal, companion screen and media. About 35 broadcasters, TV manufacturers, technology providers (some from HBB-Next) are actively contributing. Support of second-screen, media synchronisation, and user identity management are high on the HbbTV 2.0 features list.

This paper presents the technology developments by HBB-Next and its contributions to HbbTV 2.0, with attention also to user experience and business considerations.

1. INTRODUCTION

So far, HbbTV applications have mostly linked Internet and broadcast services by offering on-demand content in addition to current linear broadcast services and by replacing old-fashioned Teletext. However, many more services are on the horizon, and the market of hybrid and connected TV attracts attention from all stakeholders in the media value chain. The challenge today is to take hybrid service to a new level where the wide range of expectations of user groups are met.

The European research project FP7 HBB-Next [1] is developing technologies for next-generation HbbTV to enable services that allow more advanced interaction with the HbbTV terminal and with companion screens.

The standards body HbbTV [2] is currently developing the HbbTV 2.0 release, with 35 broadcasters, TV manufacturers and technology providers actively contributing.

This paper presents a selection of next-generation HbbTV technologies developed within HBB-Next, as well as its contributions to HbbTV 2.0 standardisation.

2. MULTI-USER IDENTIFICATION AND - INTERACTION

Accurate user identification constitutes the key to provide personalised services. HBB-Next envisions an HbbTV future in which broadcaster applications can identify and manage multiple users in front of a TV simultaneously in a user-friendly and privacy-conscious way.

2.1 Multi-modal user interface

The television context is “lean back”. This means that the effort for users to identify themselves should be minimal. A multimodal interface, using user inputs from a camera and a microphone, is one of the key features of the HBB-NEXT project, aiming at effortless utilization of the system’s features while assuring that the system only performs user commands which are properly authorized. The multimodal interface focusses on two aspects: user identification and as a means of user input.

User identification of single-user/multi-users requires modified algorithms. Multi-level user identification means that we introduce several levels of security for user authorization. The speaker identification tends to provide basic identification of the possible users located in the system installation area. The basic level would be suitable for not-so-crucial identification tasks, such as loading personal profile. The face detection approach aims to provide more reliable user identification based on users’ faces which contain far more characteristics that can be parameterised in comparison to the voice identification approach. Additionally, the 3D face recognition further extends the possibilities of feature extraction in order to more precisely identify particular persons and can be thus used for the highest level authentication (and authorization) for the most demanding applications (i.e. bank account login, etc.). Iris recognition extends the group of highly secure authentication.

A gesture navigation module has been developed for system control. Static gestures (see Figure 1), as well as dynamic gestures with defined trajectory have been used. Voice command recognition and eye movement recognition are additional possibilities for the system control.

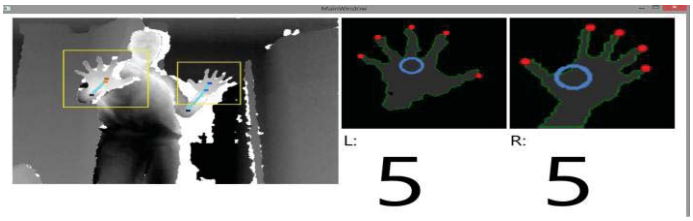

Figure 1 – Static gesture navigation by HBB-Next

In order to allow the system to give feedback within the multimodal interface, speech synthesis can be used. However, voice modality depends heavily on the nationality of the user and requires significant effort especially in multilingual areas.

2.2 Standardisation challenges

HbbTV is currently working on identity management as a novel feature for HbbTV 2.0. The idea is that an HbbTV app can ask the HbbTV terminal which users are present, of course with permission of the users and respecting privacy laws. However, such feature is not without controversy. Whereas identity management is widely supported by broadcasters, equipment vendors oppose for different reasons, such as feasibility of the business model and accountability.

Gesture- and speech navigation are not on the HbbTV 2.0 feature list as such. Yet, HbbTV is working on support of mouse and keyboard devices, and this is one way that a manufacturer could implement those.

3. MEDIA SYNCHRONISATION

Media synchronisation is about synchronising media streams from different broadcast and broadband sources. On a single HbbTV terminal this is called inter-media synchronisation. Synchronisation of media streams between an HbbTV terminal and tablets/smartphones used as companion screens that are used in conjunction is called inter-device synchronisation. Figure 2 shows the involved components in the synchronisation framework, as developed in HBB-Next.

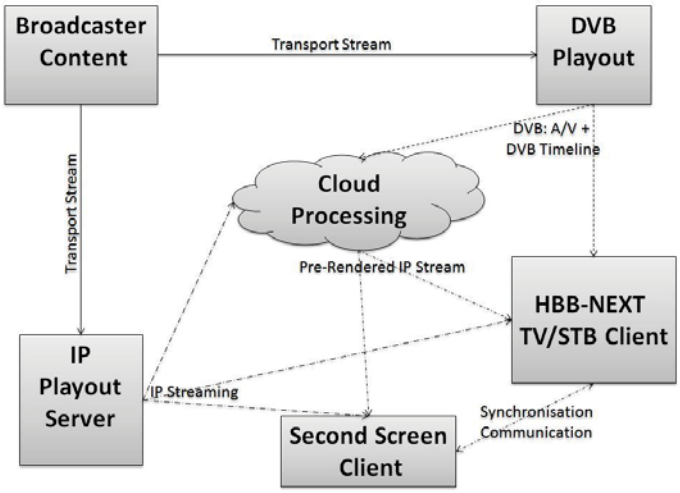

Figure 2 – Generic architecture for inter-device media synchronisation

3.1 Inter-media synchronisation

HBB-Next has developed two approaches to inter-media synchronisation. One uses pre-playout inter-stream synchronised PCR/PTS values; the other utilizes the DVB synchronised auxiliary data timeline [3] to provide event-based absolute timing reference values for synchronisation. The real-world implementation of the DVB timeline in an HbbTV related testbed is a novelty. Instances have been developed for a STB and a timeline generator broadcast server.

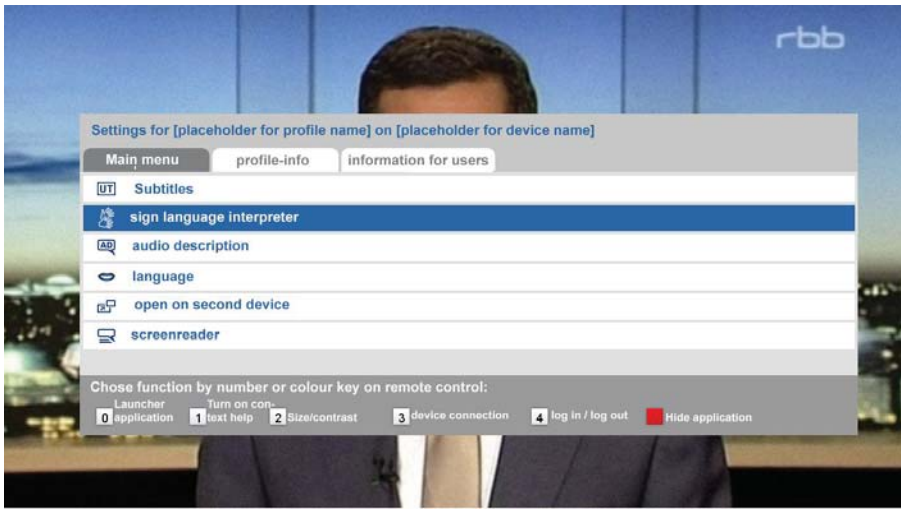

Figure 3 – Settings application providing access to synchronised auxiliary streams

These synchronisation components are controlled by an HbbTV personalised settings application, which allows for customized rendering of media streams such as subtitles in different languages, sign language videos or audio description streams. The settings app has been prototyped, as presented in Figure 3, for all end-devices (STB, tablet) and aims to improve accessibility and to enrich the media experience. All these services require

quasi-lip-sync up to total frame accurate rendering and initial user tests indicate the viability of the approaches chosen by HBB-Next. Also hybrid (mixed DVB/IP) 3D-Half-SBS rendering can be enabled with this solution to provide premium 3D experience to customers. This is prototyped by an enhanced real HbbTV-STB implementation, with novel features like PiP/PaP rendering and DVB Timeline decoding. Furthermore, a GStreamer [4] based synchronisation platform has been developed, which not only allows for frame accurate synchronisation of hybrid media, but also allows for inter-device synchronisation.

3.2 Inter-device media synchronisation

Inter-device media synchronisation is a key enabler for companion screen synchronised multimedia interaction. Frame accurate synchronisation is for example required in case a secondary audio feed or alternative angle camera feed is rendered on the second screen, which requires tight clock synchrony (e.g. below 20 milliseconds for video). In order to perform inter-device synchronisation, clock information is shared between devices. Within the Hbb-Next project, two approaches have been investigated, implemented and analysed:

1. using an absolute common wall clock (provided by e.g. NTP or GPS) and

2. using relative clocks (e.g. communicating the clock differences between devices).

When using NTP, special consideration needs to be taken as to which NTP server is used, since when using multiple (or different) NTP servers, major clock differences can occur between devices [5]. NTP requires full control over the host clock, which does not hold for all target operating systems (e.g. iOS). The clock as provided by GPS, although very accurate, is also not always accessible by the host operation system from a programming level point of view.

As an alternative, devices can exchange clock differences for example using raw TCP sockets or web sockets. Clients communicate to each other their current clock time, and, when playing media, communicate at what time they rendered which frame (e.g. the content time). This way, clock synchronisation can be achieved without the use of a true common wall clock. The “common clock” is inferred from the clock differences between clients.

By resynchronising the clocks in a timely manner, clock skew can be avoided with both clock synchronisation methods. Furthermore, this approach allows clients to communicate whether playout of a certain frame at a certain clock time is feasible (e.g. due to buffer constraints).

Both methods have been implemented in an HBB-Next enabled set top box and a platform based on the open source GStreamer framework in combination with Android and iOS devices. It has been demonstrated that frame accurate inter-device synchronisation can be achieved using both clock synchronisation methodologies, although greater variability (resulting in slight desynchronisation in the order frames) was noticed when using NTP. The accuracy is sufficient for frame accurate second screen experience.

Future research will focus on audio synchronisation (requiring even tighter synchronisation), device discovery (e.g. using uPnP) and Precision Time Protocol [6]. Also, the accuracy of NTP will be investigated in greater detail.

3.3 Standardisation challenges

HbbTV is working on different aspects of media synchronisation for its HbbTV 2.0 specification. Inter-media synchronisation and accurate app-to-media synchronisation are

on the high-priority list, while inter-device synchronisation for companion screen has medium priority. At the time of writing this paper, HbbTV is studying the use cases that need to be supported and the accuracy requirements. Also, an overview was made of to-be-solved issues, which include app-to-TV API, the content time line, impact on buffer model and GOP size, accurate wall-clock sync, media sync messaging and pairing/discovery. Solutions are coordinated with other standardisation bodies, including W3C, OIPF, DVB, ATSC and MPEG.

4. MULTI-USER CONTENT RECOMMENDATIONS

Content recommendations are important in the HbbTV context. Not only do users have access to a plethora of TV channels and carrousel content through broadcast, they also have an overwhelming choice of content that is available through broadband. A recommender system helps by selecting a relevant subset of content items. As TV content is typically consumed by groups of users, multi-user content recommendations should be supported by the recommender.

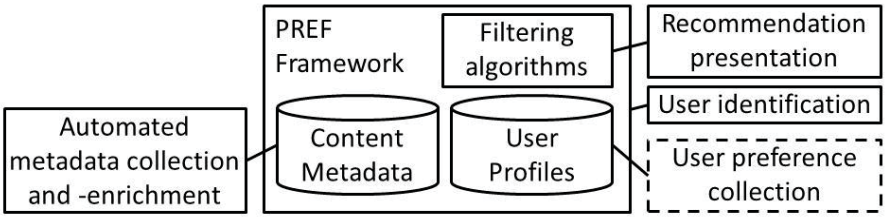

Figure 4 – PREF framework and system components

4.1 Recommender back end

HBB-Next has created a generic Personal Recommendation Engine Framework (PREF), since different recommender systems have so much in common and recommendation algorithm developers like to focus on the algorithms instead of the underlying cogs and gears. PREF provides the developer with i) a generic data model that covers users, items, ratings, preferences, groups, characteristics, and so on (depicted in Fig. 4); ii) a set of internal and external REST API's that enable the creation and storage of this data; iii) fast access to all the persistently stored data; and iv) an internal API and set of readily available algorithms that enable the developer to easily create a group recommender system.

As the TV schedule changes day by day, automated metadata collection is essential for both the calculation of the recommendations (i.e. content-based filtering) and the presentation of recommended items. Fortunately, basic TV metadata is readily available through metadata brokers. In some countries, this is even enforced by law. A system was built that collects metadata and enriches it with additional semantic metadata that is freely available via the internet, e.g. from DBPedia.

Collecting user preferences is a major bottleneck for TV program recommenders. Industry feedback taught us that users are unwilling to provide explicit content ratings in practice. Therefore, an implicit system is needed, based on the watching behaviour of the users. However, providing automated access to the user identities and clickstream runs into major privacy issues and the lack of a viable business-model to broker this information. Furthermore, the collection of implicit ratings in a group context is not straightforward. The current implementation utilizes a fixed user preferences database combined with explicit feedback and leaves the collection issue for further study.

4.2 HbbTV recommender application

The recommendations calculated by the back-end are displayed to the user(s) through an HbbTV application, as shown in Figure 5. The front-end is implemented using CE-HTML and Javascript and runs in the TV browser. A complete content distribution chain was built, including HbbTV AIT signaling, which triggers the application on a channel change. The user inter-face was kept minimalistic and clean. The user pushes the red button on the remote control to activate the recommender. Then, the user is identified and recommendations are provided in a simple grid layout as shown below, which was found the most effective based on user feed-back with various mock-ups. The user interface also shows who is watching, and for whom the recommendations are meant. Once an item is selected, a pop-up page presents further information about the program and offers the user various viewing, recording and sharing options. The playout itself was not implemented, as these are standard TV set and HbbTV functionalities.

Figure 5 – User interface of group recommender

4.3 Business considerations

In the evaluation of the system with media-industry players, conflicting requirements have been identified between user experience and business models. There is no clear place for an “identity provider” role in the current HbbTV ecosystem. Also the collection of implicit feedback runs into both business and privacy objections. Further work is needed on group recommender algorithms, explaining group recommendations to users, shielding the user’s privacy to third parties (including friends and family), and the business model for group recommendations for TV programs.

5. USER-TAILORED REPUTATION SCORES FOR APPS

By introducing HTML applications within Smart TVs, HbbTV is fostering the proliferation of applications’ repositories, stores, markets, etc in this new environment. Yet, as in many other scenarios where these appstores already exist for quite

Figure 6 – HBB-Next app-store with reputation scoring

some time now (like in the case of smartphones, for instance), reputation management is being used as an effective and powerful tool to discern, on the one hand a) which applications are properly developed and found useful by the end-users, and on the other hand b) which applications do not usually meet the end users' expectations or even contains suspicious or malicious code.

Moreover, the reputation computation engines integrated in such appstores today are not flexible and they always provide the same reputation scores to every user [9], regardless their actual preferences on the applications (usability, responsiveness, price, etc). Hence, within the context of HBB-Next, a novel reputation computation engine has been developed (see Fig. 6), addressing the aforementioned shortcomings, i.e., providing user-tailored reputation scores. Such customization of reputation scores is achieved by giving a higher weight to those feedbacks coming from users with similar preferences w.r.t. HbbTV applications.

As future research directions, we plan to develop a novel mechanism to select the most suitable reputation engine at each moment, depending on the current system conditions.

6. CLOUD OFFLOADING OF TERMINAL FUNCTIONALITY

Second devices are playing a key role in enabling interactive media applications and providing seamless media consumption experience in HbbTV. However, due to the resource limitations of mobile devices, lots of media processing has to be offloaded from the terminal side to the back-end server side, such as video/audio transcoding, overlay editing, and image resizing, amongst others.

Based on the OpenStack [7] cloud infrastructure, we designed and implemented a flexible and scalable media cloud platform for handling various media processing pipelines in the elastic cloud. As Figure 7 shows, our media cloud receives media content from broadcasters via different sources and then launches appropriate processing components across multiple virtual machines to generate adaptive streams according to the capabilities of terminal devices. A novel machine learning based resource allocation [8] has been developed to efficiently minimize cloud resource usage without affecting quality of generated streams.

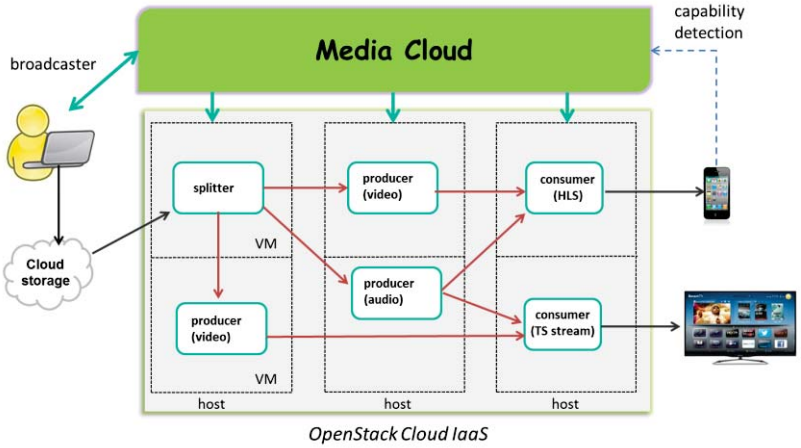

Figure 7: media cloud for media processing offloading

7. BUSINESS CONSIDERATIONS

The SmartTV eco-system as outlined by the HBB-Next demonstrators enables TV broadcasters to enhance the quality of the TV experience by improving content retrieval from both broadcast and broadband sources as well as individual accessibility. This update to the classic hassle-free shared media experience has the potential to strengthen the broadcasters' traditional position as the prime media providers in the living-room. However the devised technical platform as well supports shared interactive experiences on

connected devices beyond mere content browsing. Whether these will develop into a market of its own right or remain added value proposition to TV content remains to be seen.

The HBB-Next eco-system defines new roles in the value chain, such as recommendation-, identity- and rich-media-providers. These roles could each be adopted either by broadcasters as a vertical enhancement of the value chain or by third party players, offering their services for horizontal integration by multiple broadcasters, to independent content providers or even as stand-alone applications in itself. By providing such services in a horizontal manner, they could build on a broader user base and hold more potential for innovation. A broadcaster-centric model on the other hand would have the benefit of a tighter integration between recommendation and content management, and provide less issues with regards to content licensing. HbbTV 2.0 will be open to either approach by defining non-discriminatory open standards.

8. CONCLUSIONS

Future generations of HbbTV are envisioned to support synchronised watching of multiple media streams on multiple screens (main, companion, …), services personalised to groups of users (content recommendations, app reputation, …), multiple user identification methods (PIN, QR, face, voice, …), and media processing offload to the cloud. The FP7 HBB-Next project has conceived, developed and integrated proofs of concepts demonstrating the technical viability of this vision. The HbbTV 2.0 standard, scheduled for end 2013, will already support some of these functionalities.

REFERENCES

[1] FP7 HBB-Next, “Next Generation Media”, 2011-2014, http://www.hbb-next.eu
[2] HbbTV, http://hbbtv.org/
[3] Digital Video Broadcasting (DVB); Specification for the carriage of synchronized auxiliary data in DVB transport streams http://www.etsi.org/deliver/etsi_ts/102800_102899/102823/01.01.01_60/ts_102823v010101p.pdf
[4] GStreamer, http://gstreamer.freedesktop.org/
[5] IETF RFC 59905, “Network Time Protocol Version 4”, http://tools.ietf.org/html/rfc5905
[6] IEEE 1588 Precision Time Protocol (PTP)
[7] OpenStack, http://www.openstack.org
[8] K. Sembiring and A. Beyer, “Dynamic Resource Allocation for Cloud-based Media Processing”, in 23rd ACM Workshop on Network and Operating Systems Support for Digital Audio and Video, Oslo, Norway, 2013
[9] Félix Gómez Mármol, Marcus Q. Kuhnen, Gregorio Martínez Pérez, "Enhancing OpenID through a Reputation Framework", 8th International Conference on Autonomic and Trusted Computing (ATC2011), LNCS 6906, pp. 1-18, ISBN: 978-3-642-23495-8, Banff, Canada, 2-4/09/201

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the EC Seventh Framework Programme (FP7/2007-2013) under Grant Agreement n°287848 (HBB-Next). Also the following people have contributed to this paper: M. Probst (IRT), E. Mikoczy (ST), S. Shumann (ST), R. van Brandenburg (TNO) and A.T. Veenhuizen (TNO).