

THE BROADBAND PLATFORM: STREAMED 3D INTERACTIVE MIXED MEDIA IN THE HOME

M. PRICE, B. WEIR

*BBC Research & Development,
Kingswood Warren,
Tadworth,
Surrey KT20 6NP
marc.price@rd.bbc.co.uk*

Abstract: The future ‘broadband platform’ could be any device that is attached to the broadband-enabled home network. In this paper, we focus on the use of PCs and games consoles to provide a shared experience of streamed 3D interactive mixed media. We describe two demonstration applications which are being developed to provide a better understanding of what can be achieved with streamed 3D interactive mixed media using suitable computer game technology.

Keywords: broadband, interactivity, streaming, television, radio, computer games.

1 INTRODUCTION

The broadband-enabled, networked ‘home platform’ (hence ‘broadband platform’) offers a huge range of new possibilities in mixed media. The ability to interconnect a TV set-top box, a PC, a games console, and on a wider scale, home networks of other households, allows a new shared experience of streamed 3D interactive mixed media.

In this paper, we focus on the idea that the ‘broadband platform’ is either a PC or a games console. However in general, it could be any device that is connected to the home network. It is difficult to predict exactly how the features of the future ‘broadband platform’ will be successfully exploited in broadcast applications. One thing we can say for sure though, is that **it will not be television**. Like the spoken word of radio has not replaced books, and indeed, television has not replaced radio, any attempts to make the ‘broadband platform’ replace television will almost certainly fail.

There are a huge number of exciting ideas for future content with the interactivity of a computer game, and the controlled narrative of a professional broadcast. Hence, the broadcaster would be negligent to ignore the ‘broadband platform’. If the evolution of the platform needs to be driven according to the needs of the user - in the BBC’s case, the licence-payer - then we need to establish new techniques, procedures, and creative practices, for producing its content.

BBC Research and Development has been actively pursuing the technical side to this topic since 1996, when we began to explore new developments of cur-

rent Virtual Production technology [1]. This paper briefly reviews that work in section 2. We then move on to describe recent ‘works-in-progress’: the ‘Virtual Media Lounge’ which explores the use of a computer game engine to mix streamed audio and video content with 3D interactive content; and the ‘Massively Multiplayer Online Event’, which makes use of an interesting feature of massively multiplayer online games. These are described in sections 3 and 4. Finally, possibilities of future work are described in section 5.

2 BACKGROUND

Television production is increasingly making use of 3D models, in applications including post-produced animations and Virtual Production. These models are rendered to produce 2D images during the production process. However, with the ever increasing power of 3D graphics processors in home PCs, and new developments in 3D technology, BBC Research and Development has been considering how the broadcaster can maintain content in its 3D form all the way through the programme chain. Hence in 1996, BBC Research and Development teamed-up with 7 academic and industrial partners on a 3 year collaborative project known as PROMETHEUS [2], within the DTI’s LINK programme. The aim of the project was to prove the feasibility of an end-to-end 3D programme chain, from content production, through delivery, to fatigue-free 3D display.

The project built upon previous work in Virtual Production, where tools were developed to allow the

scenery in conventional television production to be replaced with a 3D virtual environment. In order to achieve 3D television production from this starting point, the actors and their interactions with the world around them also had to be 'virtual'. In other words, the actors had to be realistically modelled in 3D, and the resulting models then placed into the 3D virtual environment.

In PROMETHEUS, actor models were created using a range of techniques:

- Texture-mapping of live video onto 3D geometry (planes, simple curved geometry, and rough actor models);
- Animation of avatars, which are created by a 3D 'photo booth', and animated according to data obtained from marker-free, vision-based motion capture methods (for both face and body);
- Simulation of virtual clothing.

In order to deliver 3D television content to the viewer, it needs to be encoded in a way which preserves the model-based nature of the content, so that the viewer can independently control viewpoint and, if they wish, view it with a 3D display. In PROMETHEUS, delivery of 3D content used MPEG-4 BIFS [3]. Display of the decoded and rendered MPEG-4 scene used a 3D display, based on the principle of Integral Imaging [4], to provide a glasses-free, full-parallax display, which can be viewed simultaneously by several people. A more detailed review of the project is given in [5].

The PROMETHEUS project was concluded in September 2002. Since then the Virtual Production team at BBC Research and Development has been considering how this body of work could be built upon in the future. In doing so, the following points have been borne in mind.

- PROMETHEUS was intended to exploit both the computational and networking capabilities of the future home platform. However, it partially failed from inception by not considering the 'return-path' (ie from viewer to broadcaster) and interactivity (beyond movement of viewpoint).
- Very few authoring tools and players which implement the BIFS and other 3D specifications of the MPEG-4 standard are commercially available. Those that are available are very new to the market and not particularly well established even today (throughout PROMETHEUS, we were unable to find a fully functional player which supported the MPEG-4 specifications we required). In contrast, commercial computer game engines are very mature and well established, and they

generally implement most of the features that we would like to exploit on the home platform.

As a result of this, we are exploring the use of computer game platforms as the basis for the future 'broadband platform'. The planned outcomes of this exploration are demonstration applications, showing a variety of simple, cost-effective ways in which this technology could be exploited. The remainder of this paper discusses two of these applications.

3 VIRTUAL MEDIA LOUNGE

The idea of the Virtual Media Lounge is to exploit 3D game hardware and software to enhance the TV viewing (or radio listening) experience. The intention is for this application to run on either a PC or a game console. The concept of the application is as follows.

The user connects to a broadcaster-hosted 'lounge-server' with a (freely available) 'MediaLounge' application. The application is based on a 3D game engine (we are currently using the 'Crystal Space' open source game SDK [6]). The application downloads a 3D virtual world from the lounge-server, which consists of buildings - perhaps a city - and natural features. Each of the rooms in each of the buildings is a 'virtual media lounge'.

When the MediaLounge connects with the lounge-server, it also connects with a local DVB-server process that is hosted on the home-networked set-top box - at its simplest, this would be a DVB card installed in a PC attached to the home network. The DVB server streams all broadcast audio and video content requested by the MediaLounge application, over the local home network. Each virtual media lounge within the downloaded 3D world contains a virtual TV or radio, which is an appropriately shaped polygon with a tag to identify its A/V source - essentially, the tag is a channel ident. The MediaLounge application detects each tag, and maps the appropriate streamed video and audio content into the virtual environment, at the relevant locations. The user selects the desired lounge, and hence the desired channel, by navigating through the virtual world.

Each of the virtual media lounges could be created according to designs by celebrity interior (or exterior) designers. It could also be made possible for users to create their own lounges, and submit them to the server for general use.

The virtual world is downloaded and stored locally so that each MediaLounge application uses a private copy of that world. In other words, it is not a single, massively shared virtual world. However, the MediaLounge application functions a lot like a networked

game engine. It can be instructed by the user to act as a 'server', to allow a small group of elected 'friends' to connect and share the same virtual world. Alternatively, it can be instructed to be a 'client', where it searches on the internet for all other MediaLounges that are in 'server' mode, and reports those that it has permission to connect with - ie it checks whether the user has been elected as a 'friend'. Hence, the broadband connection allows the user to interact with other occupants of the virtual world, creating a shared media experience beyond the boundaries of the user's real room. The interactions could use any type of available interface, not just text or speech. For example, it might be interesting to use a biofeedback device to share emotional responses to dramatic events in the embedded media.

The embedded media stream is not restricted to just TV and radio. For example, the home DVB server could be much more general purpose, serving 'time-shifted' audio and video content (as with a personal video recorder), pre-recorded content from CD, DVD, etc, or even home-made content such as photograph slideshows or videos. Indeed, there is no reason why the virtual world should not contain portals to 3D games.

Each occupant of a room is assigned an avatar. At any point in time, the avatar is located according to the location of the viewpoint that the corresponding occupant has adopted. The avatar itself would be some form of 3D humanoid model. In its most simplest form, this model could consist of a blank rectangle, showing the user's name or a recent photograph, and the most recent text input from them. In its most elaborate form, it could be an accurate, animated model of the actual person it is representing (obtained using whatever scanning technology is available to the user). In between these extremes, it could be some kind of stylised model that has been authored by the user.

From our perspective, the idea of adapting the video texture-mapped polygon actor model technique, as used in PROMETHEUS, is the most interesting. This technique works by creating simple geometry (a rectangle, or a semi-spheroidal polygon) and texture-mapping it with a live video. In this case, the video input could come from a webcam pointing at the corresponding user, which is encoded into a very low bit-rate video stream. This could be accompanied with a low bit-rate audio stream from a microphone, or with text.

A virtual world could also contain artificial characters, with in-built AI behaviours. For example, it might be fun to watch a pop programme, and dance with the other virtual room occupants with the aid of a 'dance mat', and the artificial characters could be 'Pan's People' style computer-generated dancers.

The images in figures 1 and 2 are snap-shots of the MediaLounge application in its current state of development. Here, we see different viewpoints of a Virtual Media Lounge in a wooden floored, stone building. The lounge has a virtual TV, which is tuned to BBC2.

4 MASSIVELY MULTIPLAYER ONLINE EVENT

The Massively Multiplayer Online Event (MMOE) concept builds upon recent developments in online gaming, streamed multimedia content, and domestic broadband. Current Massively Multiplayer Online Games (MMOG), such as EverQuest [7], PlanetSide [8] or The Sims - Online [9], take place in so-called persistent worlds, gaming spaces where a user can log in at any desired time to play the game. The content is typically that provided by the game environment (creatures, dungeons, objects, etc) or by other players (social interactions). There are rarely any 'live' events in such games, and such live events are considered bonuses, rather than core components of the game.

The MMOE concept turns this model upon its head. The gaming environment exists only as a substrate for orchestrated events controlled by some third party. In order to demonstrate this, we have developed a broadband gameshow application (called 'Manhunt Island') using the 'Flashpoint' game engine [10]. The player experience is described as follows.

The player, with the gaming software installed on his home platform, logs onto the event website. When the event begins, the player is presented with a window displaying an audio/video stream from a studio. The streamed content from the studio shows a presenter (host) describing the event rules. In this case, the player takes the role of a bounty hunter on a large island. The aim of the game is for the player to track down and capture one of a number of escaped convicts that are being controlled by guests on the show. The presenter is then shown introducing the guests and transforming them into avatars, which are then inserted into the game-world. When this introduction is complete, the game launches.

The players find themselves in the game-world and must follow a series of clues in order to capture one of the escaped convicts within some allotted time. Meanwhile, the studio guests controlling the convicts are trying to avoid being captured. While the game is underway, an audio commentary is streamed to the players describing the state of the game. The game ends either when all the convicts have been captured, or after some specified period of time. The audio/video stream from the studio is presented to the player again and the presenter congratulates the winners and hands out

prizes.

5 WHERE NEXT?

In 8 years we have progressed from conventional Virtual Production to the stage of exploring prototype services which draw-upon the emerging commodity technology of the broadband-enabled home. The prototype Virtual Media Lounge application described in section 3, will continue to be developed - we are planning to set it up as an open source project in the near future. We plan to port it to a suitable games console, so that it can be demonstrated as a shared experience on several platform types.

Beyond this, we anticipate that further work would involve development of more elaborate artificial characters for insertion into lounges, so that the experience is more compelling. For example, we could have Virtual Lounges which play out music of a selected genre, and are not only occupied by avatars that are controlled in real-time by other users, but also artificial characters, who facilitate activities within the virtual room, along the lines of a club or a party.

Developing this idea further, the entire virtual world could be given behaviours which adapt the colours, geometry, and physics of the surroundings according to characteristics of the music. It could also be used for a TV music channel (such as 'The Hits', 'MTV', etc), where the music videos are played-out on a large virtual video-wall.

As mentioned in section 3, the idea of using affective feedback from users to provide input to the overall environment would also be an interesting line of study. Again, there are many ways in which it could be applied. Affective input could be biofeedback, or it could be derived using vocal intonation and gestural information [11]. However, of more interest would be a device which can acquire bio-signals remotely.

There is also the possibility of using the Media Lounge as the underlying platform for the MMOE. The introductory video could be played out on a large window, which is also tagged as a portal into the shared virtual world of the event. The user views the introductory video and then traverses the portal into the game event. If the MMOE Lounge is entered when no event is scheduled to begin, the video window could play-out a kind of newsreel, showing related news articles, action replays of exciting game-play, celebrity interviews, and award events for the most skilled players.

There are a number of technical as well as aesthetic trade-offs in considering all of these options. Avatars, AI, 3D rendering, stream encoding and decoding, etc, all require processing power on the local host. Net-

work latency effects need to be handled gracefully. This is particularly pertinent with any kind of shared dancing experience. This would need to be handled with buffering so that the resulting lag is fixed according to the beat of the music. Possible differences in hardware need to be accommodated through scalability. There are obviously issues regarding security, and how the system would work with in-place security measures (eg fire-walls). There are also issues regarding the handling of the sheer volume of users of the MMOE. However, these issues cannot be tackled until we have a reasonable involvement from interested clients.

6 SUMMARY

This paper introduced the idea of the 'broadband platform', a generic term we have coined to refer to any device that is connected to the broadband-enabled home network. Our past investigations on the development of our Virtual Production technology for use with the 'broadband platform' were reviewed. Then we described our current work, where we are exploring the use of computer game technology for streamed 3D interactive mixed media.

7 ACKNOWLEDGEMENTS

We acknowledge the contributors of the 'Crystal Space' project - especially Jorrit Tyberghein and Eric Sunshine - for their ongoing technical support on the use of the 'Crystal Space' game SDK.

BIOGRAPHY



Marc Price: Marc completed an Electronic Engineering Technician apprenticeship at 'Klark Teknik', in 1986, and after a period in industry, obtained a BEng Electronic Engineering (with first class honours) from the University of Central England in

1992. Marc then studied for a PhD at King's College London, where he investigated and developed signal processing theories for synthesising and analysing several novel classes of digital filter. Shortly after receiving his PhD in 1996, he took a position with the BBC R&D department. Initially, Marc worked on archive retrieval, on which he has a patent pending. In 1998, Marc took a sabbatical postdoctoral position at King's College London, to follow-up some of his earlier research with applications in high quality speech coding. Marc returned to BBC R&D in October 1999, to work in the general area of virtual production. Since then, Marc has contributed to the PROMETHEUS and

ORIGAMI projects, and pursued his interests in multimodal 3D interactive media.



Bruce Weir: Bruce graduated from Sheffield University in 1995 with a degree in Physics. He then worked at Smith & Nephew Group Research Centre in York for 12 months. Following this, Bruce studied for a PhD at the Cavendish Laboratory in

Cambridge, on the crystal structures of electroluminescent polymers. After receiving his PhD in 2000 he joined BBC R&D, working on virtual production technologies and thinking about how narrative in computer games can be improved. He is currently designing AI control systems for robotic cameras.

REFERENCES

- [1] O. Grau, M. Price, G. Thomas, 'Use of 3D Techniques for Virtual Production', *Proceedings of SPIE Photonics West Symposia*, Vol4309, January 2001.
- [2] <http://www.bbc.co.uk/rd/projects/prometheus/>
- [3] P. Woodward, Y. Paker, A. Pearmain, 'Implementing virtual studios in MPEG-4', *Proceedings of IEEE CAS Workshop on MPEG-4*, 2001.
- [4] M. Lippmann, 'Epreuves Reversibles Donnant La Sensation Durelief', *J. Phys*, No. 821, 1908.
- [5] M.Price, et al, 'Real-time Production and Delivery of 3D Media', *Proceedings of International Broadcasting Convention*, 2002.
- [6] J. Tyberghein, et al, <http://crystal.sf.net>
- [7] <http://everquest.station.sony.com/>
- [8] <http://planetside.station.sony.com/>
- [9] <http://thesimsonline.ea.com>
- [10] <http://www.flashpoint1985.com/>
- [11] M.Pantic, L.Rothkrantz, 'Toward Affect-sensitive Multimodal Human-computer Interaction', *IEEE Proceedings*, Vol.91, No.9, September 2003.



Figure 1: A TV Lounge: View from the Door



Figure 2: A TV Lounge: View from Inside