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

M. PRICE, B. WEIR

BBC Research and 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 a demonstration application which is being developed to provide a better understanding of what can be achieved with streamed 3D interactive mixed media.

1 Introduction

The networked 'home platform' that connects to broadband (ie the 'broadband platform') will offer an unprecedented range of features. The idea of being able to interconnect a TV set-top box, a PC, a games console, and on a wider scale, home networks of other households, brings to life a shared experience of streamed 3D interactive mixed media. Hence, the 'broadband platform' will require new kinds of user interfaces, for new kinds of content of a more elaborate nature than is available via HTTP. 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. 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. However, the fact that the 'broadband platform' represents a convergence of media technologies means that the broadcaster would be negligent to ignore it. If the evolution of the platform needs to be driven according to the needs of the user - in the BBC's case, the license-payer - then we need to establish new techniques, procedures, and creative practices, for producing its content.

BBC Research and Development has been actively pursuing this topic since 1996, when we began to explore new developments of our current Virtual Production technology [1]. This paper briefly reviews that work in section 2. We then move on to describe a recent 'work-in-progress', which is exploring the use of a computer game engine to mix streamed audio and video content with 3D interactive content. This is described in section 3. Finally, possibilities of future work are described in section 4.

2 Background

Television production is increasingly making use of 3D models, in applications including postproduced 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, 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.
1
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 must also be 'virtual'. In other words, the actors must be realistically modelled in 3D, and the resulting models 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 [2]. Display of the decoded and rendered MPEG-4 scene used a 3D display, based on the principle of Integral Imaging [3], 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 [4].

The PROMETHEUS project was concluded at the end in September 2002. Since then the Virtual Production team at BBC Research and Development have 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.

_ PROMETHEUSwas intended to exploit both the computational and networking capabilities of the future home platform. However, it partially failed from inception by not considering the 'returnpath' (ie from viewer to broadcaster) and interactivity (beyond movement of viewpoint).
_ MPEG-4 as a standard does not provide all of the functionality necessary to fully exploit the computational and networking capabilities of the future home platform. To move forward, we either need to amend the standard, or use a different one.
_ Authoring tools and players which implement the BIFS and other 3D specifications of the MPEG-4 standard 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). However, 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 outcome of this exploration is several demonstration applications, showing a variety of simple, cost-effective ways in which this technology could be exploited. One of these applications - the Virtual TV Lounge - is described in the following section.

3 Virtual TV Lounge

The idea of the Virtual TV 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 application functions as follows.

The user connects to a broadcaster-hosted 'lounge-server' with a (freely available) 'TVLounge' application.

The application is based on a 3D game engine (we are currently using the 'Crystal Space' open source game SDK [5]). The TVLounge 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 TV lounge'.

When the client 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 TVLounge application, over the local home network. Each virtual TV lounge within the downloaded 3D world contains a virtual TV, which is an appropriately shaped polygon with a tag to identify its A/V source - essentially, the tag is a channel ident. The TVLounge application detects each virtual TV 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 TV 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.

Since the virtual world is downloaded, each TVLounge application uses its own local, private copy of that world. In other words, it is not a single, massively shared virtual world. However, the TVLounge 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 TVLounges 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 TV viewing 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 (eg [6]) to share emotional responses to dramatic events in a TV programme or film.

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 users 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, in a similar way to mobile video-phones. This could be accompanied with a low bit-rate audio stream from a mic, 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.

Also, each of the virtual TV lounges could be created according to designs by celebrity interior (or exterior) designers. Or, it could be made possible for users to create their own lounges, and submit them to the server for general use.

The images in figure 1 are snap-shots of the TVLounge application in its current state of development (31st December, 2003). Here, we see different viewpoints of a Virtual TV Lounge in a wooden floored, stone building. The virtual TV is tuned to BBC2.

4 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 service described above, the Virtual TV Lounge, will continue to be developed. We plan to port it to a suitable games console, so that it can be demonstrated as a shared experience on several platform types. At this stage, we will be in a position to publicly demonstrate the application, and hopefully generate positive feedback on how to shape the idea further.

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 also have Virtual Radio 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 what might happen in 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. The list of ideas along these lines is endless.

As mentioned in the previous section, 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 [7]. However, of more interest would be a device which can acquire bio-signals remotely.

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. Network latency effects need to be handled gracefully. Possible differences in hardware need to be accommodated through scalability. There are obviously concerns regarding security, and how the system would work

with in-place security measures (eg fire-walls), and so on. However, these issues are not really of concern to us until we have a reasonable involvement from interested clients.

5 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. We reviewed our past investigations on the development of our Virtual Production technology for use with the 'broadband platform'. Then we described our current work, where we are exploring the use of a computer game engine to mix streamed audio and video content with 3D interactive content.

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] P. Woodward, Y. Paker, A. Pearmain, 'Implementing virtual studios in MPEG-4', *Proceedings of IEEE CAS Workshop on MPEG-4*, 2001.

[3] M. Lippmann, 'Eppeuves Reversibles Donnant La Sensation Durelief', *J. Phys*, No. 821, 1908.

[4] M.Price, et al, 'Real-time Production and Delivery of 3D Media', *Proceedings of International Broadcasting Convention*, 2002.

[5] http://crystal.sf.net

[6] http://www.ibva.com

[7] M.Pantic, L.Rothkrantz, 'Toward Affect-sensitive Multimodal Human-computer Interaction', *IEEE Proceedings*, Vol.91, No.9, September 2003.

Figure 1: Views of a TV Lounge

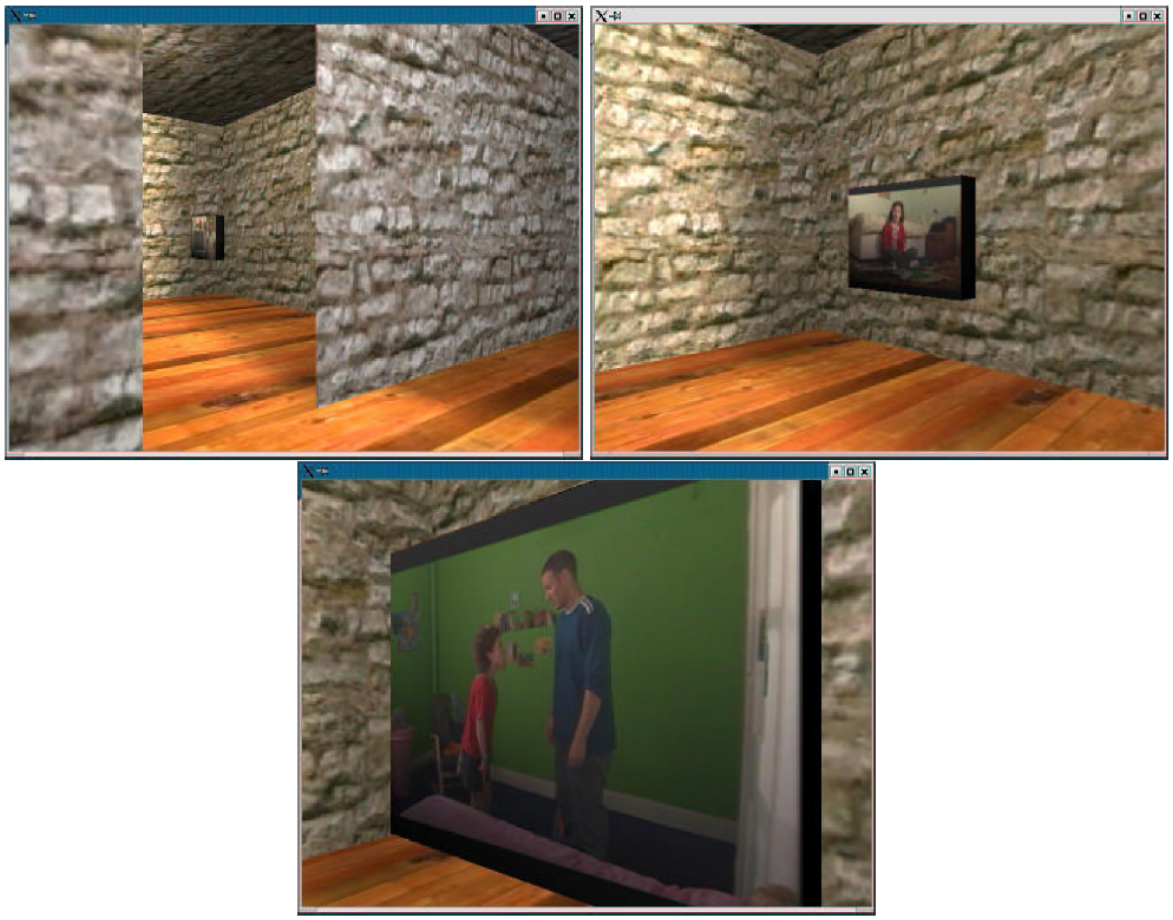

Figure 1: Views of a TV Lounge