Production, Delivery and Playback of 3D Graphics

Thomas Di Giacomo
University of Geneva, Switzerland

Chris Joslin
University of Geneva, Switzerland

Nadia Magnenat-Thalmann
University of Geneva, Switzerland

INTRODUCTION

Three-dimensional (3D) representation is one of the cornerstones of Computer Graphics (CG) and multimedia content. Advances in this domain, coupled with the highly fuelled progression of 3D graphics cards, have pushed the complexity of these representations into a whole new era, whereby a single real-time model can consist of more than a million polygons. Huge architectural buildings, everyday objects, even humans themselves, can be represented using 3D graphics in such detail that it is difficult to distinguish between real and virtual objects. Concurrently, and much towards the other end of the scale, many devices, such as Personal Digital Assistants (PDAs), mobile phones, laptops and so forth, are now "3D capable" to enhance a user's experience and to provide much more depth to the information presented. In many cases, these devices access the same content from the same service provider; for example, providing virtual maps/guides, multi-user games and so forth. It is this broadness of content and the heterogeneity of devices in terms of performance, capability, network connection and more that is the main concern in a continuously expanding market. It is also the concern of users to obtain the best quality for their device; that is, the general expectation of any device of higher performance is that overall the quality of the experience will be better.

Overall, three main stages have to be ensured to meet such requirements within an entire integrated chain. First, 3D media contents, from 3D models to animation data, have to be designed and produced by content providers, designers or artists, for instance. Second is to deliver all these pieces of content at a user's request, and although this kind of data has mainly been stored locally on the system, it is now more likely to be delivered via networks (either via download or streaming), similar to other media types (such as video and audio). Finally, upon reception, this media content has to be consumed by the users; that is, played back on their device. After a brief discussion on the background and concepts required for such goals, those three main steps are presented in detail in the following respective sections.

BACKGROUND

Though 3D technology is often considered, by misconception, as a local storage of data accessed by stand-alone applications – for example, video games, as Joslin, Di Giacomo and Magnenat-Thalmann. (2004) discuss – collaborative virtual environments have opened the way for distributed and networked 3D applications. Such architectures are becoming more and more common, and today's Web graphics are very much evolved. Lau, Li, Kunii, Guo, Zhang, Magnenat-Thalmann, Kshirsagar, Thalmann and Gutierrez (2003) present the emergence of standards that provide generic tools and formats for an even wider availability of such systems. Furthermore, standards enable interoperability and genericity, and extend the usability of graphics-based distributed applications. The MPEG group (see Walsh & Bourges-Sévenier, 2001 for a detailed description on MPEG-4) is one of the most important actors of these standardization efforts. Though it handles many different types of media, and while originally the most commonly used were video and audio, 3D graphics is now receiving a great deal of attention, especially in the use of Binary Format for Scenes (BIFS) and by the Synthetic Natural Hybrid Coding (SNHC) subgroup for Animation Framework Extension (AFX). For example, AFX specifies the case of 3D virtual human animation with FBA and BBA (see Preda & Preteux, 2002).

Throughout this article, we will use the example of animated 3D virtual humans, for two reasons: First, it is a complete and consistent example of a 3D graphics application; second, because the presented concepts are easily extendable to other applications in 3D. One must note that while other standards, such as VRML and X3D, are more dedicated to 3D data than MPEG, our discussion lies in the context of MPEG because MPEG-4 can be considered as a superset of VRML. Overall, MPEG provides a complete framework for multimedia delivery, and recently, with the MPEG-7 and MPEG-21 standards, it even allows for the inclusion of media objects semantics as well as evolved processing, such as Digital Item Adaptation (DIA) or intellectual property management and protection (IPMP).

Moreover, adaptation is probably one of the most important issues considering 3D graphics delivery, and it offers a wide range of possibilities; for instance, by driving single content delivery towards multiple devices. The underlying approach uses a description of scalable encoded data to allow for bitstreams modifications (as illustrated by Figure 4), which is often done by using generic Bitstream Description Language (gBSDL), explained by Amielh and Devillers (2002). Though important work on spatial scalability has been done in Computer Vision by Lindeberg (1994), for instance, adaptation can be processed for many different types of media, as proposed by Gioia et al. (2004). One must focus on the particularity of individual media, such as audio as discussed by Aggarwal, Rose and Regunathan (2001), or video by Kim, Wang and Chang (2003), to ensure an optimum adaptation, both for the quality of content provided and for the accuracy of the adaptation. Recently, graphical adaptations have started to be designed by Raemdonck, Lafruit, Steffens, Otero-Perez and Bril (2002), and by Boier-Martin (2003), but many issues still remain, especially when considering the factor of real human perception, as described by Adelson (1991).

Furthermore, the large and growing variety of today's platform is a very important factor to consider when widely distributing 3D graphics. Some researches are oriented towards optimization for a specific device; other work is taking advantage of some computer graphics (CG) processing, such as Image-Based Rendering of Chang and Ger (2002), to allow 3D playback on light devices. Other approaches use these devices as a display only. For instance, Lamberti, Zunino, Sanna, Fiume and Maniezzo (2003) render 3D scenes with a cluster of machines and then transmit the rendered images to a PDA to be displayed. Recently, important work has been carried on for appropriate graphics API on light devices, the major one being probably OpenGL ES for embedded and mobile devices.

PRODUCTION OF 3D GRAPHICS

To provide immersive virtual experiences, 3D graphics contents must be carefully crafted and produced. Though a detailed description of the entire available production methods and pipeline would require many pages, such processes are required by every 3D-based applications, and thus are briefly discussed here.

The production of 3D graphics roughly consists of two main stages: the design of 3D shapes, and the creation of animation sequences applied on 3D models. Details on these stages follow, with a focus on the production of scalable contents to enable their adaptation in the delivery stage and an efficient playback for the client.

Representation

The standard techniques for designing 3D shapes, often referred as the global term of modeling, range from 3D scanning to user-designed models, with the help of 3D modeling software (e.g., 3DStudio Max, Maya, etc.). With today's technologies, it is also possible to create a virtual clone, using real pictures to create virtual humans according to anthropomorphic parameters and so forth, as shown in Figure 1. Scanners are usually used to produce a first-draft version of a mesh, which is then refined and com-