Experimenting with Multimedia Advances using GPAC

Jean Le Feuvre, Cyril Concolato, Jean-Claude Dufourd, Romain Bouqueau, Jean-Claude Moissinac
Multimedia Group, Telecom ParisTech, LTCI UMR 4141
Paris, France
{ lefeuvre, concolato, dufourd, bouqueau, moissinac }@telecom-paristech.fr

ABSTRACT
Multimedia applications are challenging software and require collaboration of very different components such as networking, rendering, or scripting to provide a nice user experience. GPAC is an open source multimedia framework that implements a vast number of components and helps experimenting with different types of multimedia applications. It provides a multimedia packager, some servers and a player for dynamic and interactive multimedia content. In this paper, we present the recent additions to the project which allow experimenting with new applications such as dynamic home networking or rich broadcasting of interactive content on the latest devices, including with 3D displays. Such experiments can be useful in research work and, as illustrated, in academic environment for educational purposes.

Categories and Subject Descriptors
H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems –*Animations, video;* H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *graphical user interfaces, standardization;* H.5.4 [**Information Interfaces and Presentation**]: Hypertext/Hypermedia – *Architectures.*

General Terms: Algorithms, Experimentation, Standardization, Languages.

Keywords: Multimedia, Interactivity, Broadcasting, Stereoscopy, Home Networking, MPEG, BIFS, SVG, Web3D, Streaming.

1. INTRODUCTION
GPAC [1] is a cross-platform open source multimedia framework distributed under the LGPL license. Since its creation the project has been known for its strong focus on standard technologies for the coding, delivery and playback of dynamic and interactive multimedia content, such as those developed within MPEG, W3C or Web3D standard bodies. More recently, the project has opened the way for experimenting with new multimedia technologies such as HTTP adaptive streaming, Home Networking (UPnP, DLNA) or stereoscopic multimedia, as will be described later on in this paper.

GPAC offers a collection of tools for interactive media packaging, distribution and playback, all built upon a single library, called libgpac; this library contains all core functionalities of the framework, enabling a simplified reuse of GPAC by third-party projects.

The main packaging application of GPAC is called MP4Box. Among its numerous functionalities, we can cite:

- Multiplexing many audio, video and text formats, multiplexed or not, in ISO Base Media Files (e.g. MP4),
- Encoding and decoding scene description formats such as MPEG-4 Binary Format for Scenes (BIFS) and Lightweight Application Scene Representation (LASeR),
- Preparing the content for HTTP download or streaming with the Real-Time Transport Protocol (RTP).

MP4Box has a large user base within the audio and video community (e.g. default MP4 packager for x264[1]), and is used by many academics[2]. GPAC also contains several servers for streaming with RTP and with MPEG-2 Transport Stream (MPEG-2 TS). These tools will be further detailed in the next section.

The main aspect of GPAC is however its multimedia player, MP4Client/Osmo4. With its focus on interactive content, the player differs from traditional media players in many ways, and is close to web browsers, as it enables playback of:

- Interactive, animated 2D vector graphics such as BIFS or W3C Scalable Vector Graphics (SVG),
- Interactive, animated 3D graphics such as MPEG-4 BIFS or Virtual Reality Modeling Language (VRML) or X3D, the XML successor of VRML,
- Multiple controllable video, sound or subtitles objects in 2D or 3D content,
- 2D content integrated in 3D content and vice-versa,
- Javascript for BIFS/SVG/VRML, including AJAX.

One particularity of the player is its extensibility through a large set of plug-ins covering media formats demultiplexing, media decoding, javascript extensions or user event filtering. The player already has support for a wide range of multimedia delivery protocols (HTTP including IceCast, ShoutCast, RTP streaming, MPEG-2 TS). This makes the player quite suitable for research experiments, as new functionality can be added without modifying the player code. The player can also be seen as a good candidate for testing new ideas for interactive multimedia. For instance, most (if not all) aspects of the upcoming HTML Media specification[3], such as buffer monitoring, media control or selection, are covered and demonstrated on the GPAC web site [1]. Advanced aspects such as integration of 3D content in Web content such as envisaged in [3]

[1] x264, http://www.videolan.org/developers/x264.html

[2] GPAC References, http://gpac.wp.institut-telecom.fr/about/references

[3] http://www.w3.org/TR/2011/WD-html5-20110405/video.html

can also be tested. Although quite versatile, one weakness of the project is its lack of support for HTML/CSS.

The GPAC project was already selected for the 2007 edition of the competition, and an in-depth overview of the software can be found in [2]. This paper presents, in Section 2, the major evolution of the platform in the recent years and in section 3, how it is used for educational purposes.

2. ADVANCED MULTIMEDIA USAGES

The field of multimedia applications is always evolving at a fast pace, with new technologies appearing every year or two. The GPAC projects tries to keep this pace and to integrate new software components in order to facilitate experimenting with these new technologies.

2.1 Home Networking

Among the foreseen new usages, the convergence between the internet world, the broadcast world and the home is seen as a promising source of new multimedia applications. Standards such as those developed by the Digital Living Network Alliance (DLNA) are being deployed, enabling devices in a home network to discover each other and to communicate. Connected TV sets are being shipped to propose Internet content in the home. Within this context, we have added the Universal Plug and Play (UPnP) support, through the Platinum library[4], and the ability for the GPAC player to act as a remote control for other DLNA media players on the network, selecting both target screens and source content.

Additionally, the scene description capabilities of the GPAC player have been leveraged to expose the information and control offered by DLNA within the scene. This work was done within the scope of the standardization of widget technologies. In the recent years, standard bodies such as MPEG or W3C have worked on Widgets, which are packaged (e.g. ZIP) multimedia presentations (HTML, SVG, BIFS) used as standalone applications. We have been particularly involved in this work at MPEG, especially on the following topics: communication between such widgets in a home network, and session mobility of such widgets between various devices (TV, laptop, smart phones).We have demonstrated interoperability of MPEG-U Widgets with UPnP-based services, especially DLNA media player and media server. MPEG-U is ISO/IEC 23007-1 and its GPAC implementation was released as the MPEG reference software; some videos demonstrations can be found on the GPAC web site[5]. GPAC is also a conformant implementation of W3C Widgets Packaging and Configuration[6] and The Widget Interface[7].

2.2 Media Transport

As part of its recent developments, new media transport methods, mostly related to broadcast or near-broadcast, were added to GPAC.

In the past few years, the French Government has been investigating the usage of the T-DMB standard [4] for Augmented Digital Radio, bringing high quality audio along with interactivity to the user. The GPAC player was used by many industrial during this phase, because of its native support for the T-DMB system. We have been working on enhancing this support in GPAC, and have introduced live BIFS encoding and live audio stream injection in our MP42TS tool to generate T-DMB compliant radio streams, as described in Figure 1. A video on GPAC web site demonstrates the result of this work[5], which is being integrated in the MPEG-2 Systems standard through a new amendment. Similar developments were made with the live injection of 3GPP DIMS content into MPEG-2 TS in the context of the DVB-H or ATSC-M/H standards were demonstrated at NAB2010.

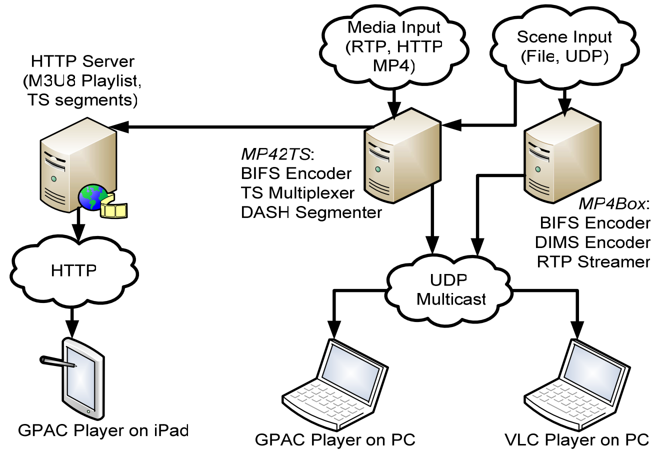

Figure 1 - GPAC Media Server Tools

Finally, GPAC has been extended towards HTTP streaming technologies. Solutions such as the proprietary Apple HTTP Live Streaming (HLS)[8] or standardized MPEG Dynamic Adaptive Streaming over HTTP (DASH) [5] have been added both in the player and in the content production tools. For instance the MP42TS application is now capable of truncating its MPEG-2 TS output in segments compatible with the DASH standard or with HLS, in live or offline mode, as described in Figure 1. The MP4Box tool can produce MP4 files compatible with the MPEG-DASH standard, and the player is capable of processing both types of content.

2.3 3D Screens

A recent trend in digital television industry is the deployment of 3D TV. We have started in 2007 a research activity around usage of auto-stereoscopic screens on mobile platforms, and demonstrated usage of video, slide shows and interactive contents for such devices. In order to test depth effects on other platforms, we have recently extended GPAC to be able to display multimedia content on many types of 3D screens [6]. GPAC now supports anaglyph rendering (stereo only), or any number of viewpoints rendered side-by-side, top-to-bottom, and columns- or row- interleaved output. Custom view interleaving is also possible through OpenGL shaders, and has been tested with a 5 views display built by SpatialView[9], as shown in Figure 2.

[4]Platinum UPnP SDK, http://www.plutinosoft.com/platinum

[5] http://gpac.wp.institut-telecom.fr/gpac-features/videos/

[6] http://www.w3.org/TR/2011/WD-widgets-20110322/

[7] http://www.w3.org/TR/2010/WD-widgets-apis-20100907/

[8] http://tools.ietf.org/html/draft-pantos-http-live-streaming-06

[9] http://www.spatialview.com/

Figure 2 - Depth-based rendering on 5 views 3D display

In multiview mode, the player currently handles 3D content (BIFS/VRML/X3D), 2D scalable content (BIFS/SVG) with some extensions to provide depth information (as explained in [2]), video with depth map as used in MPEG-C [7] and image with depth map. Although still in experimental stage, this makes GPAC the first open-source multimedia browser with support for many 3D displays, including multiview auto-stereoscopic display.

2.4 Scalable Video Coding

Among the new formats that have been developed recently, the MPEG Scalable Video Codec (SVC) [8] provides an interesting ability to represent video content as scalable layers. It is interesting for multimedia applications to adapt to the varying network conditions or display size. As said previously, the player functionalities can easily be extended with various plug-ins, in particular for media processing. To support SVC, the code of OpenSVC project[10] was integrated as a new plug-in and means to select a layer (quality, spatial resolution) from the scene was added. Additionally, the ability to manipulate SVC content within MP4 files was also added to MP4Box.

2.5 Platforms

Although most of the work on GPAC derives from research activities, the project tries to enable experimenting advanced multimedia applications on many platforms, even the latest ones.

For instance, a special effort has been made to produce installer packages for most desktop platforms (Linux-Debian, Windows, Mac OS X), enabling end-users or students to experiment GPAC without having to compile it.

Additionally to the existing Windows Mobile port, mobile versions of GPAC now exist for Apple's iOS and Google's Android. Videos of GPAC running in iPhone or Android devices are available on GPAC web site[5]. Figure 3 shows GPAC on an iPad playing an SVG electronic service guide transcoded from DVB-T broadcast with GPAC. Figure 4 shows GPAC on an ad-hoc digital radio device. The player has also been successfully ported on embedded SH4 platforms, typically set-top boxes.

Finally, one effort in GPAC was the design of a platform-independent graphical user interface based on GPAC-supported scene technologies such as BIFS or SVG and Javascript. This effort enables the deployment of applications on additional platforms more efficiently.

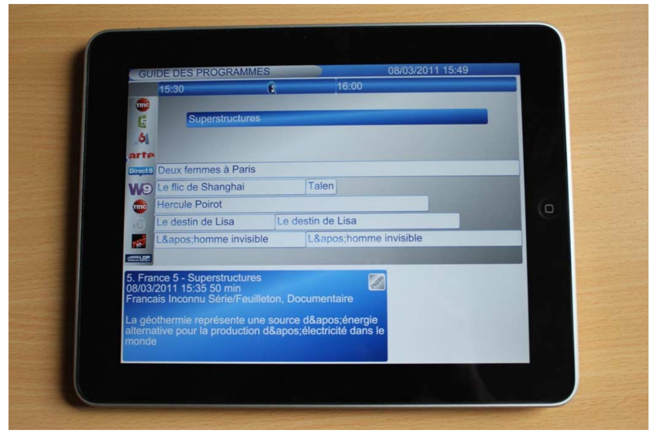

Figure 3 - GPAC on iPad showing an EPG application

Figure 4 - GPAC on a T-DMB kitchen-radio

3. EDUCATION

GPAC is a corner stone of our daily activities, and we use it in research projects as well as educational works. This section presents the effort made in this area.

3.1 Usage of GPAC in student projects

It is our very belief that programming projects play a very important part in the student's learning, as they help the student to acquire in-depth understanding f concepts seen in classes through practical cases. For a long time, we have been involving our students (master, PhD …) with GPAC, taking advantage of its plug-in and core-library based architecture to have them implement complex scenarios without having to program all the complexity of a full-featured multimedia player or delivery system. Depending on work quality and context, these projects may be released to the public.

We usually diversify the topics of our project. One recurrent topic is human-computer interaction modes in GPAC, and students have produced GPAC plug-ins for head-tracking based on OpenCV[11] (released as the *opencv_js* module), Wiimote control based on wiiuse[12] (released as the *wiis* module) or Voice recognition based on HTK[13].

Another recurrent topic is scene description technologies, usually based on BIFS or SVG. Some project only use GPAC tools such as

[10] http://opensvcdecoder.sourceforge.net

[11] http://opencv.willowgarage.com

[12] http://sourceforge.net/projects/wiiuse/

[13] http://htk.eng.cam.ac.uk/

Osmo4 or MP4Box in the content production or playback phase. Such projects include for example a DVD to BIFS transcoder or an Electronic Program Guide in BIFS. Other projects build on top of GPAC or extend it, as was done with a simple authoring tool based on GPAC (released as *V4Studio*) or with a plug-in building an SVG spectrogram out of an audio signal.

In parallel to our classes on multimedia transport, some students have realized an SVG streaming application, a DVB-H burst simulator or have included BIFS carouseling mechanism in some GPAC tools (MP42TS).

We have also proposed more programming-centric projects to our students, out of which we must cite the JPEG2000 decoder plug-in based on OpenJPEG[14] and the web-browser plug-ins for NPAPI-based browsers (Osmozilla) and ActiveX-enabled ones (GPAX). All these projects are now part of the public GPAC source.

Students in 2009 and 2010 used GPAC as a capture tool in projects comparing the playback of SVG files in any browser with a reference image. These results should soon appear in GPAC's upcoming automatic testing platform.

3.2 Other Academic Usages of GPAC

Apart from student projects, we use GPAC in several student classes. Students use MP4Box to construct, inspect and dump MP4 files, and based on box layout dumped by MP4Box, deduce media tracks interleaving properties. They use GPAC RTSP/RTP logging and RTP over RTSP tunneling capabilities to study the behavior of RTSP sessions, timing and network information exchanged and how they match values in RTP packet.

The player is also used in scene description classes, where students have to author simple SVG or BIFS (2D or 3D) content, progressively adding interactivity, scripting and AJAX support. The BIFS part of these classes has been extended into a tutorial in French and English publically available on [1].

We are also planning to use GPAC in an upcoming corporate training session on 3DTV to demonstrate auto-stereoscopic screens, and are proposing to our students subjects related to that topic, namely spatial view interpolation, where GPAC is used as the visualization tool for the generated frames.

Our collaborations with other academics have confirmed that GPAC is also used for lab sessions in other universities, such as Telecom SudParis or by master students in Klagenfurt University or Gent IBBT-MMLab.

Apart from student-related usage, GPAC is used in many of our research projects[15]. GPAC is also being used by academics throughout the world in many papers, and although it is quite hard to quantify due to lack of explicit references, we try to maintain a list of published work referring to GPAC on our web site.

4. CONCLUSION

In this paper, we have presented the GPAC multimedia framework and detailed the features added to the framework since the publication of [2].

We have also detailed how our students use GPAC for lab sessions or technical projects, and attempted to measure the impact of the framework on the academic community. Although GPAC is always in heavy development, we are pleased to see that the academic community is actively using it. Future directions of the project will include hybrid broadband/broadcast tools, stereoscopic enhancement, and advanced user interactions.

5. ACKNOWLEDGMENTS

Part of the work presented in this paper is financed by the following French-funded projects: HybRadio, Calder and EndT2End. The authors would like to thank all the other contributors to the project (from students, to developers and post-docs).

6. REFERENCES

[1] GPAC Web Site, http://gpac.sourceforge.net

[2] J. Le Feuvre, C. Concolato, J.C. Moissinac. "GPAC: open source multimedia framework". In *Proceedings of the 15th international Conference on Multimedia*, Augs*burg, Germany, September 25 - 29, 2007. MULTIMEDIA '07. ACM, N*ew York, NY, 1009-1012. DOI= 10.1145/1291233.1291452

[3] Johannes Behr, Peter Eschler, Yvonne Jung, and Michael Zöllner. 2009. X3DOM: a DOM-based HTML5/X3D integration model. In Proceedings of the 14th International Conference on 3D Web Technology (Web3D '09), Stephen N. Spencer (Ed.). ACM, New York, NY, USA, 127-135. DOI=10.1145/1559764.1559784

[4] G. Lee, S. Cho, K.-t. Yang, Y K. Hahm, and S. I. Lee, "Development of terrestrial DMB transmission system based on Eureka-147 DAB system", IEEE Trans. on Consum. Elec., vol. 51, no. 1, pp. 63-68, Feb. 2005. DOI=10.1109/TCE.2005.1405700

[5] Thomas Stockhammer. 2011. Dynamic adaptive streaming over HTTP : standards and design principles. In Proceedings of the second annual ACM conference on Multimedia systems (MMSys '11). ACM, New York, NY, USA, 133-144. DOI=10.1145/1943552.1943572

[6] N. A. Dodgson, "Autostereoscopic 3-D displays," IEEE Computer, vol. 38, no. 8, pp. 31–36, Aug. 2005. DOI=10.1109/MC.2005.252

[7] A. Bourge, J. Gobert and F. Bruls, "MPEG-C part 3: Enabling the introduction of video plus depth contents," in Proc. of IEEE Workshop on Content Generation and Coding for 3D-television, Eindhoven, The Netherlands, June 2006.

[8] Robert Kuschnig, Ingo Kofler, Michael Ransburg, and Hermann Hellwagner. 2008. Design options and comparison of in-network H.264/SVC adaptation. J. Vis. Comun. Image Represent. 19, 8 (December 2008), 529-542. DOI=10.1016/j.jvcir.2008.07.0

[14] http://www.openjpeg.org/

[15] http://gpac.wp.institut-telecom.fr/about/references/#rd