Multimed Tools Appl
DOI 10.1007/s11042-013-1516-7

HTML5-based system for interoperable 3D digital home applications

Mikel Zorrilla · Angel Martin · Jairo R. Sanchez · Iñigo Tamayo · Igor G. Olaizola

Abstract Digital home application market shifts just about every month. This means risk for developers struggling to adapt their applications to several platforms and marketplaces while changing how people experience and use their TVs, smartphones and tablets. New ubiquitous and context-aware experiences through interactive 3D applications on these devices engage users to interact with virtual applications with complex 3D scenes. Interactive 3D applications are boosted by emerging standards such as HTML5 and WebGL removing limitations, and transforming the Web into a real application framework to tackle interoperability over the heterogeneous digital home platforms. Developers can apply their knowledge of web-based solutions to design digital home applications, removing learning curve barriers related to platform-specific APIs. However, constraints to render complex 3D environments are still present especially in home media devices. This paper provides a state-of-the-art survey of current capabilities and limitations of the digital home devices and describes a latency-driven system design based on hybrid remote and local rendering architecture, enhancing the interactive experience of 3D graphics on these thin devices. It supports interactive navigation of high complexity 3D scenes while provides an interoperable solution that can be deployed over the wide digital home device landscape.

Keywords Home device interoperability · Digital home applications · Computer graphics · 3D virtual environments · Interactivity · Hybrid rendering system

M. Zorrilla · A. Martin (✉) · J. R. Sanchez · I. Tamayo · I. G. Olaizola
Vicomtech-IK4, Mikeletegi pasealekua 57, 20009 Donostia-San Sebastián, Spain
e-mail: amartin@vicomtech.org

M. Zorrilla
e-mail: mzorrilla@vicomtech.org

Q1
Q2

1 Introduction

Users are becoming more accustomed to improved experiences that provide interactive 3D applications exploiting the technology in immersive environments. Thanks to the advent of low energy-consumption Graphic Processing Units, interactive 3D applications are currently running in most digital home devices. Connected TVs, smartphones and tablets are being fitted with graphic capabilities providing users an enhanced experience on top of interactive and 3D applications, and pushing the market to new advanced 3D applications with complex interactive virtual environments.

The landscape of digital home devices has changed last years completely with the introduction of smartphones and tablets in the home network bringing secondary displays to foster customized media, together with the evolution of the TV to Smart Connected TV. Moreover, these kind of devices are running over application-based Operating Systems. Most popular are Android and iOS [22] for smartphones and tablets and different proprietary platforms (Samsung, Philips, etc.) for Smart TVs. However big companies such as Google or Apple offer a Connected TV solution which could nearly provide a full digital home approach through the different devices and their Operative System. Each solution facilitates a framework and an SDK (Software Developer Kit) to exploit native assets providing the hardware features of the devices: connectivity, motion and voice control, camera, GPS, graphic capabilities, etc. However, the deployment of the applications from one OS to the others implies major changes and specific adaptation. This platform heterogeneity at the OS level generates an important interoperability problem.

The rapidly increasing use of the Web as a software platform with truly interactive applications is boosted by emerging standards such as HTML5[1] and WebGL[2] that are removing limitations, and transforming the Web into a real application platform middleware to tackle the interoperability problem. Following this trend, the new HbbTV[3] standard for broadcasting environment interactivity is also based on a specific HTML browser.

HTML5 provides devices the capability to run rich web applications accessing the entire device features on a web browser. It comes together with CSS[4] and JavaScript which provides an appropriate framework for the content interactivity and universal access to different APIs. WebGL is the API oriented to 3D graphics in the HTML5 canvas element. It is easier to craft innovative user experiences using powerful HTML5 layout and WebGL rendering engines than current native IDEs.

Digital home browsers are rapidly adopting HTML5 features on a tough race just after the desktop browsers. The standard has won a prominent place as a horizontal approach to reach interactive multimedia applications on home devices. HTML5 applications can be packed for the different execution environments providing an interoperable application with minor changes through different OSs. That is why HTML5 is being strongly promoted by the standardization bodies and a sector of

[1]Html5 standard specification (May 2011) http://www.w3.org/TR/html5/
[2]Webgl website (Mar. 2011) http://www.khronos.org/webgl/
[3]HbbTV 1.5 specification (April 2012) http://www.hbbtv.org
[4]Cascading style sheets (css) standard specification (May 2011) http://www.w3.org/TR/CSS/

the market to achieve a HTML5 marketplace instead of the different proprietary ones, such as Android Market, iPhone App Store, Samsung Apps Market, Net TV Apps, etc.

Digital home applications are changing how people experience and use these devices. The incoming pioneer interactive 3D applications for mobiles are inciting users to discover new ubiquitous and context-aware experiences through smartphones and tablets and show the feasibility to access this rich media apps through the Smart TV. User requirements are involved in the mentioned tough race demanding power efficient techniques together with advanced interactive virtual applications with complex 3D scenes on digital home devices as they do on PCs.

The introduction of the canvas element into HTML5 enables 3D rendering on the Web while WebGL technology brings hardware-accelerated 3D graphics to the Web Browser without plug-ins turning HTML5 into the promising solution to cope with such fragmented device market by universal developments for device-independent applications and services. This paper provides a complete state-of-the-art of the current browser capabilities of the digital home devices using HTML5. We present performance results concluded by experiments carried out in representative set-top boxes, smartphones and tablets. The current limitations to run advanced interactive 3D applications are also explained in the article giving rise to a system proposal to overcome the detected handicaps to be able to run advanced interactive 3D applications using HTML5, making thin devices suitable for a wider range of applications. A system architecture called *3DMaaS* is detailed to provide complementary rendering capabilities to these devices, adding to their own capabilities the chance to push to the cloud complex 3D rendering tasks. A technical validation of *3DMaaS* is done emphasizing on the overcoming of the limitations detailed on the state-of-the-art.

2 Digital home device software platforms

The TV is still the main device for watching media content in the digital home. Nevertheless in the same way that mobile phones have gone from thin to smartphones and tablets, providing access to all kind of services and contents, home television is evolving from a passive device for multimedia content consumption to the so called “SmartTV”. Worldwide shipments of Internet-connected televisions have reached 25 % of total units in 2011 and it is expected that it will the 70 % by 2016.[5]

However, the Connected TV platforms are very heterogeneous and based on proprietary approaches, where the interoperability is a problem. TV manufacturers have developed their own frameworks, providing a SDK to develop specific applications and with a own marketplace. Samsung Smart TV provides a SDK to develop Flash-based or JavaScript engine-based applications. These applications are located by Samsung in their marketplace called Samsung Apps. Philips Net TV provides a CE-HTML browser with a index page to access to the Net TV apps. Connected set-top boxes are also very heterogeneous with different web browser such as Opera Mobile, specific OS such as Boxee[6] or burgeon Linux/Android devices[7] to transform not-

[5] May 2012. IMS Research

[6] http://www.boxee.tv/

[7] http://www.raspberrypi.org/

connected TVs into a full connected devices. Moreover, Google and Apple have their TV solutions, Google TV[8] and Apple TV[9] respectively, but they are not positioned yet as a market leader as they do on mobile systems.

Smartphone and tablet market penetration is going faster than Connected TVs. In Q3 2012, global smartphone shipments jumped to 179 million units.[10] It was a rise of 45 % from last year beating the annual growth rate. Growth continues but it is slowing down as most of the developed markets come close to 80–90 % penetration. Meanwhile, global tablet shipments reached 20 million units in Q3 2012.[11] As tablets and smartphones get faster, allowing a quicker transfer of data, integrating new connectivity and interactivity paradigms along with fancy graphics, users have developed a habit for downloading applications. This pushes mobile application market to a rapid evolution shifting the business landscape and to a competitive environment. The research firm Gartner recently forecast that mobile application stores will deliver 310 billion downloads internationally in 2016 and $74 billion in revenue.[12] A key difference of each platform is the application market, App Store, Android Market and Windows Phone Marketplace. Hereby, Gartner claims that an integrated cross-device experience will help fuel this demand.

The Android Market is open [3], whereas others are gated. This means, Android foster developers to self-publish created applications into the Android Market, whereas Apple or Microsoft decide what gets published keeping the application approval right before they become available in the Marketplace.

The different marketplaces availability responds to the change on the mobile phone landscape, playing their correspondent OS, such as iOS, Android, Windows Mobile or Research In Motion (RIM) BlackBerry OS, a prominent role in the applications development [22]. The market penetration of Android and iOS is increasing strongly and both are becoming the two major OSs to take into account. While on January 2011 the market share of Android and iOS was 61 % in Europe, on December 2012 it has raised up to 85 % thanks to the big increase of the Android OS. This trend is more representative in North America, where on January 2011 the Android and iOS mobiles represented the 61 % of the market, and two years later they are over the 90 % (Fig. 1).

The Android operating system is built from a modified Linux kernel. Previous specific versions for tablets and smartphones, entirely designed for devices with large screens and thinner devices respectively, converge in the version 4 that brings together phones and tablets easing the multi-device development and interoperability. The software stack contains Java applications running on a virtual machine, and system components are written in Java, C, C++, and XML. In order to develop Android applications the SDK can be integrated in different environments such as Eclipse.

Apple developed the iOS for its products catalogue. The operating system is derived from Mac OS X and is built on top of the Darwin foundation and XNU

[8] http://www.google.com/tv/

[9] http://www.apple.com/uk/appletv/

[10] Kang, T.: Global smartphone vendor market share: Q3 2012. International Data Corporation (Oct. 2012)

[11] Mawston, N.: Global tablet vendor market share: Q3 2012. Strategy Analytics (Oct. 2012)

[12] Market Trends: Mobile App Stores, Worldwide. Gartner (Sept. 2012)

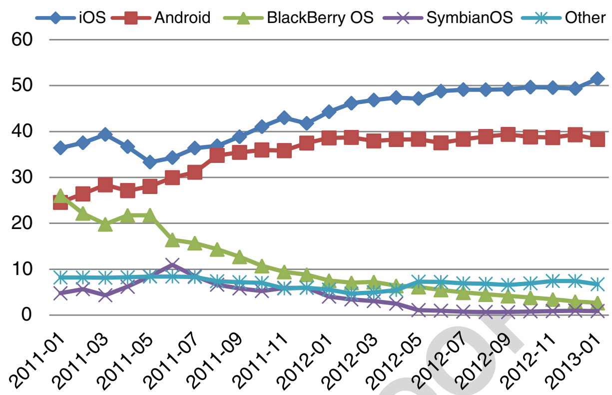


Fig. 1 Top mobile OSs in North America from Jan 2011 to Apr 2012 (StatCounter Global Stats)

kernel. XNU combines the Mach 3 microkernel, elements of Berkeley Software Distribution (BSD) Unix, and an object-oriented device driver API (I/O kit). iOS frameworks are written in Objective-C. In order to tackle application development for iOS, Apple provides a Xcode development environment and a iOS Simulator to test applications.

To sum up, the current digital home platform ecosystem is heterogeneous, with several operating systems, programming languages, and interfaces, resulting in more complex software cross-platform development and testing processes. Digital home devices increasingly depend on reliable software to offer a fresh user experience. Hence the current trend in developing interoperable applications lays on using Web technology instead of platform-specific APIs.

3 The web as a software platform

According to the wide landscape of digital home application frameworks described in the previous section, developers need to carefully determine how and where to invest their time and effort before tackling an application development project. Writing native applications requires developers expertise and background in specialized IDEs.

However, Smart TV, smartphone and tablet trend is to be always connected to the Internet. Application developers should not ignore advantages of moving from desktop computing to web-based applications [1, 20, 21]. On the one hand, applications provided on the Web as services do not require installation or manual upgrades, easing the software life cycle management while inherit web security and privacy policies. On the other hand, in terms of monetizing an application, another relevant advantage lays on the deployment and sharing of Web applications that can be instantly worldwide, with no middlemen or distributors. This way the application monetizing strategy do not have to obey marketplace policies enabling a free design of the business model. Last but not least, the potential of the web-based applications can support user collaboration over the Internet, deploying virtual spaces where users interact and share application experience and data, fostering new paradigms of interactivity and social networking.

Developers can also benefit from web-based solutions saving time and effort. They can apply their knowledge of designing web applications to smartphone, tablet or TV application design, removing learning curve barriers. Web-centric approach for digital home applications enables not only rapid prototyping, but also unified integration with Web services. It requires access to the hardware resources of the digital home devices through JavaScript that always lags behind the new capabilities that manufacturers introduce. In order to mitigate this limitation new W3C's HTML standard provides device orientation, speech recognition and geolocation management bridging from native features to web-centric development.

From the viewpoint of the developers, the key for transition towards web-based software is the ongoing evolution of web development technologies, specifically HTML, CSS and JavaScript. This way, development turns more efficient to face interoperable and innovative mobile user experiences exploiting powerful HTML layout and rendering engines than native IDEs.

Emerging standards such as HTML5 and WebGL will play a crucial role removing the remaining limitations and transforming the Web into a horizontal software platform. They will significantly shift the perception of the web browser and web applications capabilities to a fully featured web-centric operating system and to a fully interoperable application respectively.

W3C HTML5 standard specification[13] defines the core language of the World Wide Web. New features and elements are introduced paying an special attention to improve interoperability.

HTML5 provides many capabilities enabling developers to combine video, audio, 3D, and 2D into one seamless application. HTML5 embraces multimedia by means of built-in audio and video support through *<audio>* and *<video>* tags that allow media files to be played without third party browser plug-in components. Moreover, specifically for live multimedia streaming Dynamic Adaptive Streaming over HTTP formats rise as the solution to provide high quality video streaming on the Internet thanks to enabled adaptivity. To sum up, adaptive HTTP streaming is a promising technology to overcome access to media consumption through home network devices facing the bitrate and resolution adaptation to each singular context while manage seamless underlying network topology. Thereby, Google Chrome, Opera, Safari or Firefox browsers bet on Adaptive HTTP Streaming formats including them on their development roadmaps. There are different proprietary implementations such as Microsoft Smooth Streaming, Apple HTTP Live Streaming (HLS) or Adobe HTTP Dynamic Streaming. But MPEG-DASH had been accepted by ISO as an International Standard with the purpose to converge all the proprietary approaches into the standard.

HTML5 also brings relevant features fostering new paradigms of interactivity and user experience. The Canvas API provides salient 2D drawing capabilities for interactive graphics. Moreover, HTML5 specification provides numerous additions and enhancements such us realtime message based, speech recognitions, device orientation awareness or drag&drop action to be applied to the HTML objects.

[13] Html5 standard specification (May 2011) http://www.w3.org/TR/html5/

CSS3[14] brings lots of possibilities that boost creativity such as transitions, opacity definition and native columns. It also provides much more flexibility enabling 3D effects such as zoom, pan, rotation, transformations and animations.

But the most relevant features that turn HTML5 into a interoperable software application platform are:

- Offline operation. The HTML5 contains several features that address the challenge of building web applications that allow to operate even when an active network connection is not available.
- Local storage. HTML5 brings a persistent cache based on local SQL database, allowing data to be stored locally in the device. It also provides a filesystem API in order to manage read and write actions.

Moreover, HTML5 development can be easily transformed in an application package ready to be provided in the Android marketplace or in other App stores. PhoneGap[15] is an open source framework for creating mobile web applications in HTML5, JavaScript and CSS3 while still taking advantage of the core features of native applications in some platforms such as iOS and Android devices.

Applications often engage users through 3D visual interfaces. They are more effective, attractive, and are considered as a key factor to add value to applications improving the overall user experience. For the Web, WebGL takes the role of enabling technology as a solid foundation for 3D graphics applications [15].

WebGL[16] is a cross-platform web standard for hardware accelerated 3D graphics API developed by the Khronos Group that includes among others Mozilla, Apple, Google and Opera. WebGL brings to the Web the support to display and manipulate 3D graphics natively in the web browser without any plug-in components. WebGL performs 3D graphics on top of the HTML5 canvas element and is accessed using Document Object Model (DOM) interface. WebGL allows communication between JavaScript applications and the OpenGL software libraries, which accesses the graphics processor of the device. This makes possible to exploit hardware capabilities to render 3D content. WebGL is based on OpenGL ES 2.0, and it uses the OpenGL shading language GLSL.

4 HTML5 interactive 3D applications

4.1 Advanced 3D application requirements

Recently the use of 3D graphics in many industrial fields and applications such as games, advertisement products interaction, serious gaming/simulation for more effective training, financial and medical data analysis, and CAD design are increasing more and more. Often applied data 3D applications interfaces exploit 3D graphics to support professional user productivity or represent data that could not be done otherwise such as Google Earth. But 3D graphics is also used for making more visually attractive interfaces.

[14] Cascading style sheets (css) standard specification (May 2011) http://www.w3.org/TR/CSS/

[15] Phonegap website (Jan. 2012) http://www.phonegap.com

[16] Webgl website (Mar. 2011) http://www.khronos.org/webgl/

Due to the mobility of users and professionals involved in these applications, it is mandatory to provide access through mobile devices tracking the variety of contexts of the user. Facing the emerging trend in consumer technology for delivering 3D content to the mainstream user via digital home devices [16], new solutions must promote the communication between the TV and mobile devices of the digital home. This provides users access to 3D graphics applications through Connected TVs, smartphones and tablets having a great experience interacting with 3D virtual environments. Demand for 3D visualization is increasing in these devices as users expect more realistic immersive experiences. So 3D graphics combines immersion and interactivity fostering creativity for new envisaged applications and information navigation interfaces.

Mobile games are one of the fastest growing segments of the application industry. Bringing together the social gaming paradigm and the internet connection capability of most of the digital home devices, users will embrace the online interaction trend from PC.

User interfaces based on 3D graphics let users interact with virtual objects, environments, or information but the experience can be improved with the inclusion of real media sources around the user. According to [2] definition, virtual worlds technologies completely immerse a user inside a synthetic environment. In contrast, augmented reality allows the user to enjoy the real environment, with virtual objects superimposed upon or composited with the real world providing extra information or interactivity about what is around. This requires the fusion of very heterogeneous media sources in a concept called 3D Media [4, 23]. 3D Media is composed of different audio and video sources, static images and 3D objects enabling enhanced experiences.

3D rendering pushes the visual boundaries and interactive experience of rich environments, but 3D interfaces, virtual worlds and augmented reality applications require high 3D graphical features. The more complex the 3D scenes are, the higher the hardware requirements are. Although connected TVs, set-top boxes, smartphones and tablets are rapidly improving their graphic capabilities thanks to the integration of low energy-consumption Graphic Processing Units, the capabilities are below the user expectation. Users are demanding experiences they are used to in powerful devices such us PCs, mixed with the new characteristics that the digital home devices provide, such us using the camera of the smartphone for an ubiquitous augmented reality experience.

4.2 Limitations of the browsers in digital home devices

WebGL API coupled with JavaScript engines are boosting increasing capabilities of the web browsers making possible to develop complex computational environments including 3D graphics. Therefore platform-independent applications are directly performed through the web browser on different devices without the need to install additional software or plug-ins bringing the accessibility and interoperability of the web. However, constraints to render complex 3D environments are still present in digital home devices. It is necessary to define the hurdles, in terms of performance, that a developer will face when creating a web browser-based software for 3D interactive applications on top of HTML5 and WebGL. Here, we introduce not only the limitations around complex applications that require 3D graphics technology according to a set of evaluations, but also the clues about the bottlenecks origin that would enable the work to be done to remove the detected barriers.

We have chosen different devices that provides a wide representative landscape of the current digital home platforms, in order to detect the browser capabilities and limitations. On one hand, we have selected two high performance set-top boxes with browsers that support HTML5 and WebGL:

- **Innout Media Center 4Gs HD Set-top Box**: Opera Mobile 12.0 browser supporting HTML5 / WebGL profile and HbbTV profile.
- **Gigabyte GN-SB100 series**: Android 2.2 OS, Opera Mobile 12.0 browser.

These set-top boxes support WebGL but do not have specific hardware to run it, so they can not deal with it at all. However, Opera Mobile has announced[17] that its TV browser with WebGL runs on the recently launched Intel Atom Media Processor CE5300. Mitsubishi Electric is also working on a set-top box with a high performance TV browser called Espial[18] with WebGL applications support.

On the other hand, according to the mobile devices in the digital home, two of the selected devices are Android and the other two are iOS. In order to track the market trend, where the tablets have an increasing presence, the evaluations consider two smartphones and two tablets:

- **Samsung Galaxy S**: GT-I900 smartphone with Android 2.2.1 firmware.
- **Samsung Galaxy TAB**: GP-P1000 tablet with Android 2.2 firmware.
- **iPhone 4**: iOS5 smartphone.
- **iPad**: iOS5 tablet.

Android and iOS Safari default browsers do not support WebGL yet. Neither Opera Mini nor Google Chrome Beta version for Android 4 do, but all of them have included it in their roadmaps. Here, for the Android devices, the Mozilla Firefox 4.0 browser have been employed for the tests. Firefox has WebGL support and it can be installed from the Android Market. Other browsers such as Opera Mobile 12.0 also support WebGL for Android devices. But for the iOS devices, a specific application which runs a webkit based browser called GoWebGL[19] provides WebGL capabilities.

In order to measure the frame rate achieved for each device, a simple 3D scene is composed using rotating cubes. In each test the total number of 3D objects is increased as well as their polygonal complexity, ranging from 1 to 80 objects and from 12 to 200k polygons per object. The geometry of one single cube is loaded into a vertex buffer which is drawn multiple times using a different transform matrix for each cube. The performance is measured as the average time per frame sampled over 50 frames for each object and polygon configuration. Figure 2 shows the results of these tests as the maximum number of polygons that can be rendered in interactive time (15 fps) in function of the object quantity in the iPad.

As seen in the plot, the maximum number of polygons drops exponentially with the number of 3D objects. Given that the geometry is loaded as a vertex buffer and

[17] March 2012. IP&TV World Forum in London

[18] March 2012. http://www.espial.com/company/press_item/id745

[19] https://github.com/gauthiier/GoWebGL

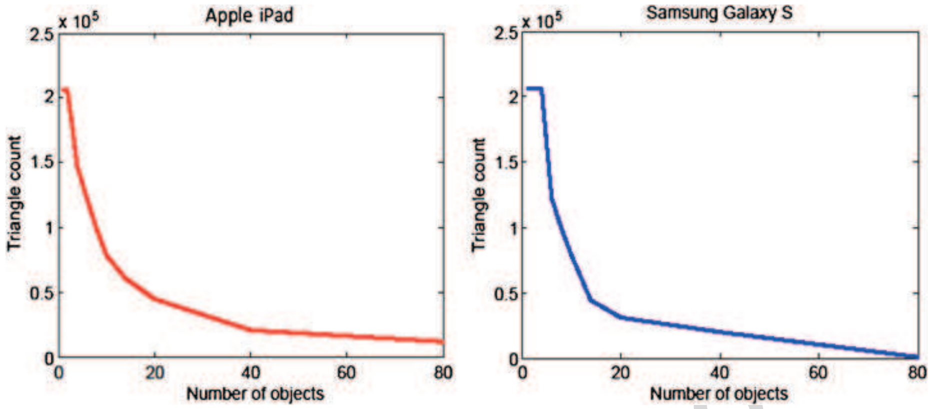


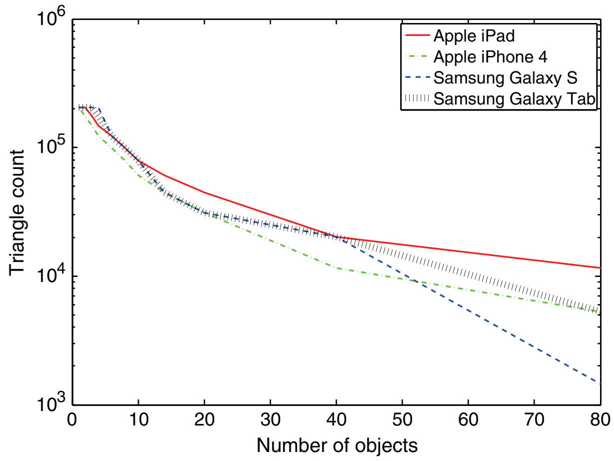
Fig. 2 Maximum number of polygons that can be rendered in interactive time (15 fps) in function of the object quantity in the iPad and Samsung Galaxy S

provided that the total number of polygons is maintained constant, these results can be explained by two reasons. On the one hand, floating point operations are very CPU demanding in JavaScript. Since an additional cube means an additional matrix rotation, the overall performance is significantly affected. On the other hand, the new transform matrix must be transferred to the GPU overloading the CPU-GPU communication bus. Although the amount of data is quite small, the bus latencies of these small devices can have a negative impact which means a notable bottleneck.

From these results it can be derived that current WebGL subsystems can support a good performance for simple scenes composed by small amounts of objects, regardless of its polygonal complexity. This limitation brings an important drawback hindering scene-graph based rendering engines, since each object in the graph must be transformed recursively with respect to its parent.

Figure 3 compares the performance of the different devices in terms of the maximum number of polygons that can be rendered against the number of objects while keeping 15 fps target frame rate. The trend is quite similar through the four

Fig. 3 A comparative of the maximum number of polygons that can be rendered in interactive time (15 fps) in function of the object quantity in the different devices



devices. The iPad and iPhone have the same 3D processing behavior while increasing the number of objects, being the performance of the iPad slightly better than others. The Android devices achieve almost the same throughput while increasing the number of objects and the responsiveness is very close to the iPad. However, the capabilities drop from 40 objects, specially in the Samsung Galaxy S.

These results become evident the need to improve the performance of 3D applications over the digital home browsers. However, these measures where unachievable some months ago, and the rapid adoption of HTML5 features on the mobile browsers let us think that these results are going to be improved very fast removing barriers in terms of WebGL compliance. Android and iOS browser will be able in the near future to run WebGL in the same way that other mobile browsers will do it (Google Chrome, Opera Mini, etc.) and will be accessible from these platforms. The WebGL performance itself needs to be improved by a better integration of the JavaScript capabilities of the browser and the architecture of the device. However, remaining throughput limits closely related to GPU potential would not disappear quickly due to life battery technological constraints.

In terms of HTML5 and WebGL support for Connected TVs and set-top boxes, different initiatives such as Espial or the Opera Mobile for TV highlight the relevance of these technologies on the roadmaps of the TV browser developers.

Anyway, according to the obtained results, and even if the capabilities of the devices are going to increase rapidly, users are already demanding advanced 3D applications on digital home devices. The proposed *3DMaaS System* faces all the previously described issues responding those who are not willing to upgrade their devices as fast as the market moves. It also optimizes the development investment of a new application turning it suitable for any device with video streaming support. Moreover, these results establish thresholds to define 3D performance profiles to support local/remote rendering distribution decisions according to the 3D scene complexity. This way *3DMaaS System* can mitigate local 3D processing stress of the device by taking care of full or partial 3D rendering in a remote resource that is real-time encoded and streamed inside a video.

Section 5 shows the related work on different approaches to increase the capabilities of the devices to render 3D content and introduces the *3DMaaS System* proposed in Section 6, which allows to extend the capabilities of the devices pushing to the cloud complex 3D rendering tasks and combining it with its hardware possibilities on a hybrid system.

5 Related work

A solution based on remote rendering performed by a high processing cloud server with enough network bandwidth resources can keep the target performance while achieve interoperability widen the audience. The server would manage all the 3D Media involved in order to render the 2D result according to the user actions. Last but not least, standard mechanisms to adapt the video stream to the network capacity can solve bandwidth problems. However, this solution delegates the final performance to interaction latency. Different approaches driven by the described solution, face digital home device's applications to overcome the current limitations in terms of 3D processing and rendering.

The gaming sector is the main driver for graphics computation. Recently launched cloud hardware solutions, such as by Nvidia Grid product,[20] brings promising cloud-based video streaming technology ready to deliver gaming content to consumer devices, including smartphones, tablets, PCs, and TVs enabling up to 36 concurrent HD-quality video streams with low latency from a single server using NVIDIAŠs GPU virtualization technology. Nvidia Grid faces GaaS boosting such as OnLive[21] or Gaikai,[22] lately incorporated to Sony, to overcome scalability and performance issues.

The concept of Gaming as a Service (GaaS) is presented on [13] where the quality of experience and the latency are key factors of success. These features can be dramatically enhanced when combining the computational load of the local machine with remote rendering by sending complex calculations to a remote server using proprietary approaches. Laikari et al. [9] proposes the Games@Large System oriented to set-top boxes on home networks and for enterprises such as hotels. Fechteler and Eisert [5] extends the Games@Large System with the main idea to calculate motion vectors directly from the 3D scene information used during rendering of the scene.

Similar hybrid computation approaches also tackle visualizing 3D objects on other sectors. These solutions consist on sending graphical commands such us roto-translation parameters from the end client to the server. This way the server can calculate the strictly necessary data that the end client needs and stream it offering a progressive reconstruction of the polygons. These solutions are valid for a mere combination of 3D objects, but not extensible for 3D Media based applications.

Lamberti et al. [10] proposes a remote rendering scenario for mobile devices like PDAs running a dedicated application called Mobile 3D Viewer. This approach is based on the Chromium software [8]. Marino et al. [12] presents an approach sending 3D graphical commands in a stream from the server to the client and it is based on WireGL [7].

SHARC System [18] is an approach for enabling scalable support of realtime 3D applications in a cloud computing environment. It is based on service virtualization with tools like VNC. This solution extends VNC as a video streaming platform. VNC and similar virtualization tools are also used on [19] and [11].

Nadalutti et al. [14] presents a MobiX3D mobile player for access 3D content through mobile devices using OpenGL ES.

Contrary to the above described solutions our approach does not require a specific player o application on the client side, running on a HTML5 browser to overcome interoperability. In Section 6 we present the *3DMaaS System* which exploits the potential of WebGL, based on OpenGL ES 2.0, leveraging 3D processing on mobile devices by delegating 3D WebGL rendering to a remote server. *3DMaaS System* enables the 3D Media content based applications by means of adaptative video streaming from the server side to the end device.

[20]http://www.nvidia.com/object/cloud-gaming.html

[21]http://www.onlive.com/

[22]http://www.gaikai.com/

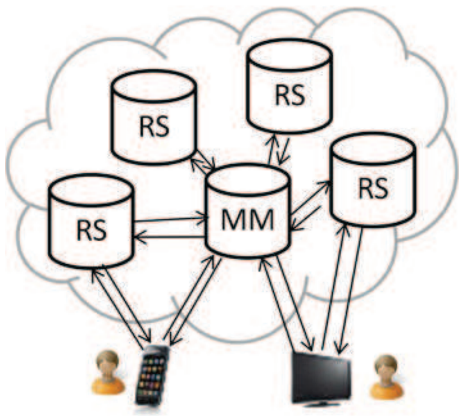


Fig. 4 General infrastructure of the *3DMaaS System*

6 3DMaaS system design and experiments

6.1 Design

There are three main actors on the *3DMaaS System* [24] (Fig. 4): MaaS Manager (MM) which monitors the computational load of the resources pool and dispatchs the device request to one of them to achieve a target QoS through load balancing strategies; Rendering Server (RS) the remote rendering resource; and their communication with the end devices. The features that *3DMaaS System* requires are really affordable for any kind of end device. Moreover, the cloud rendered stream is adapted to the different codecs and parameters to represent the media content at the different end devices (set-top boxes, smartphones, tablets, etc.). A block diagram of the general architecture is shown in Fig. 5 and all the modules are more deeply explained below.

RS is the core module of the *3DMaaS System*. Figure 5 shows the different blocks of the RS and its communication with MM and the end device:

- **Web services with MM**: MM reports end device context to the RS.
- **Internal manager**: It manages the requests and creates the streams.
- **3D Media & Render**: According to the real-time captured context such as object user interaction it generates the rendering for the composition.

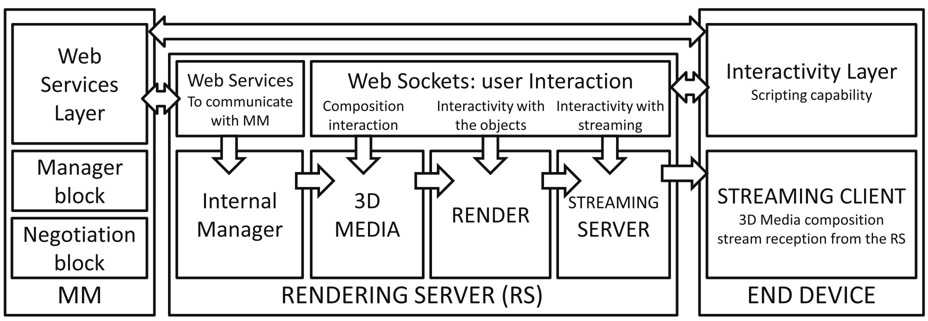


Fig. 5 The block diagram of MM and RS and their communication with the end device

- **Streaming server**: It deals with real-time encoding and the streaming session with end device taking into account the context profile captured/negotiated by the MM: Device features (supported streaming protocols and codecs, screen size, etc.); Connection context (network bandwidth, etc.); and User preferences (objects in the composition, their size, etc.).
- **Web sockets for user real-time interaction**: TCP web sockets are used for real-time communication. The content user interaction is translated to: changes on the composition (add new elements, delete them, move their position, resize them, etc.); modifications over an object (3D movements, texture changes, stop or rewind a video or audio, etc.); and adjustments of the streaming parameters (video resolution, bitrate, codec, etc.).

The *3DMaaS* System aims a wide range of video streaming formats in order to fit in very different devices. To achieve it *3DMaaS* provides a complete set of streaming formats [25], dealing with RTSP and Dynamic Adaptive Streaming over HTTP [6] such as HLS and MPEG-DASH. The *3DMaaS* Streaming Server must launch a suitable pipeline according to the previously negotiated format because each alternative is supported depending on the browser implementation.[23] Open Source frameworks provides the pillars to the 3DMaaS Streaming Server. Being more specific, Gstreamer performs RTSP server and some plugins[24] [25] bridge HLS communication, while GPAC[26] and DASH-JS [17] JavaScript- and WebM-based DASH library for Google Chrome hold MPEG-DASH compliance. Last but not least, x264 tune options[27] accomplish the required ultra low latency that keeps a good interaction latency to guarantee the quality of experience of the user.

In terms of achieving low latency, the main solutions deployed lays on: Web socket for application logic communication and system awareness of user interaction; video codec tuning to push the streaming processing time to the minimum; multimedia encapsulator set up to minimize the buffering requirements; RTCP session, for those suitable streaming protocols, in order to perform quality of connection measures enabling dynamic streaming parameter settings to keep QoS.

Concerning scalability, *3DMaaS System* size is a critical factor because it must provide enough RS resources to satisfy the incoming demand of remote rendering service. To face it, *3DMaaS System* has been designed to ease the rapid deployment of new RS instances but an automatic elastic behavior according to usage forecasts is out of focus.

Regarding the end device, the capabilities required by *3DMaaS System* for the HTML5 application of the client are really affordable for most of the common digital home devices. It only has to include *video* tag with the video streaming address provided by the RS and scripting capabilities to send HTTP interaction parameters. Their target is twofold: establish a new connection with *3DMaaS System* on a initial negotiation through MM; and for delivery of TCP web socket requests for low-latency interaction once the streaming communication is running with the RS.

[23] http://www.longtailvideo.com/html5/
[24] http://gitorious.org/ylatuya-gstreamer/gst-plugins-bad/commits/hlswip
[25] https://github.com/ylatuya
[26] http://gpac.wp.mines-telecom.fr/2012/02/01/dash-support/
[27] http://mewiki.project357.com/wiki/X264_Settings

6.2 Experiments

The critical performance metric of remote rendering solutions is the experienced latency for delivering a frame after graphics rendering update driven by user interaction. Our approach based on a video streaming server for 3D interactive application overcomes latency challenge. It is tackled by the hybrid solution proposal combining remote rendering of background 3D objects, where latency does not have a high impact on the user experience, with local browser WebGL rendering of foreground 3D objects which require low latency. The application server processes the user input and renders new screen frames and transmits them to the device in real time. Moreover, [25] presents *3DMaaS System* results for low latency streaming applications achieving 27.84 ms latency score. The hybrid strategy minimizes the number of objects that the browser have to render optimizing performance. For this, various experiments were carried out in order to assess the efficiency of the proposed architecture for visualization of 3D scenarios through digital home browsers. Users interact with the 3D rendering applications running on an accelerated graphics back-end for remote rendering and web browser for local rendering, allowing highly interactive experiences regardless of the complexity of the scene being considered.

Here, a low quality connection of the end device would have a negative impact on latency. To mitigate it and keep the Quality of Experience, the streaming session is monitored and dynamically modified in terms of bitrate and framerate. Since visualization framerate experienced at the mobile client constitutes the main limitation of 3D web based applications, especially when considering complex 3D scenes, framerate driven analysis tests have been designed in order to accurately quantify critical parameters of our hybrid visualization system, thus providing an effective measure of the performance of the proposed architecture.

Unfortunately, none of current available TV sets are not able to deal with 3D rendering tasks. These devices cannot perform 3D WebGL applications due to lack of specific hardware but can also benefit from the *3DMaaS System* pushing to the cloud the whole rendering scene instead of building an hybrid rendering approach. Therefore, the experiments performed to define the performance thresholds on hybrid scenarios have been focused on mobile devices.

The tests have been done over the same devices described in Section 4.2 in order to measure the frame rate achieved for each device with the *3DMaaS System*. But in this case two superimposed HTML5 canvas have been involved. The one on the front is the simple 3D scene described on Section 4.2, composed by rotating cubes. The canvas in the back is a <*video*> tag receiving a live video stream from the remote rendering server with the 3D background.

In order to set up the tests, the same range that defined in Section 4.2 has been employed for the number of 3D objects as well as their polygonal complexity in the front canvas. This way the performance combining the 3D local rendering capabilities and video stream reception on the different mobile devices is compared with the obtained measures on Section 4.2 with a mere local 3D rendering.

Figure 6 compares the frame rendering time for a number of polygons performed in a Samsung Galaxy TAB including the 3D object canvas and the live video stream visualization, with the measures obtained for the same 3D scene without the background video stream. The aggregation of the remote rendered live video stream does not have a considerable impact on the performance adding just an extra constant CPU demand. This way, rendering time for advanced applications

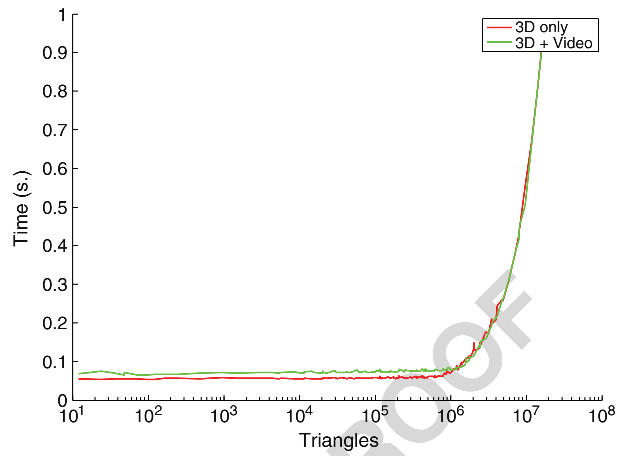


Fig. 6 A comparative of the frame rendering time and the number of polygons rendered in the Samsung Galaxy TAB, for 3D contents and added remote rendered live video stream to the 3D contents

with demanding 3D capabilities are not penalized by the added video stream. The GPU turns into a bottleneck from 10^6 triangles for this simple 3D scene, so this barrier settle the complexity that can be afforded by the device GPU without performance drawbacks. From this point remote rendering would make possible complex scenarios with no GPU overhead keeping the interactivity performance of the application. So this approach provides the application enriched 3D rendering capabilities, extending the device's hardware through remote rendering.

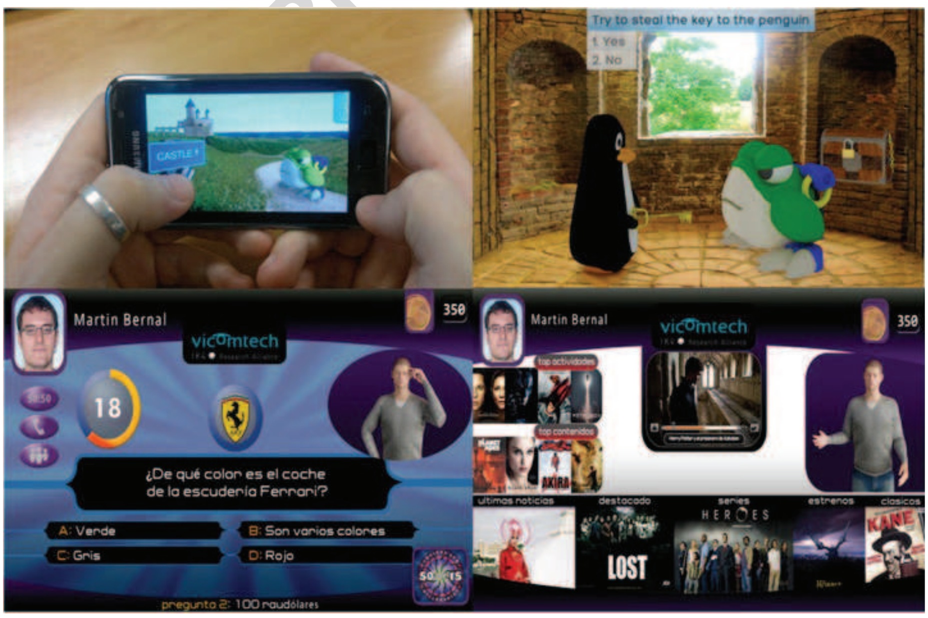


Fig. 7 Set of 3D Media contents delivered through 3DMaaS

The results obtained by the proposed architecture for hybrid remote and local rendering enhance the interactive experience of 3D graphics on digital home devices, proving the feasibility of interactive navigation of high complexity 3D scenes while provides an interoperable solution that can be deployed over the wide device landscape. However, this approach transfers responsibilities related to synchronization and OpenGL state consistency of local and remote 3D scenes to the application.

Figure 7 depicts different services and games from pure 3D object interaction in games, for e-learning purposes, to on demand content delivery services, specifically driven to e-inclusion and entertainment, which have been deployed on top of the *3DMaaS* infrastructure comprising the system portfolio.

7 Conclusions

Digital home application is a very disruptive market overcoming the potential that the Internet has and the incorporation of mobile devices together with the evolution of the TV to Smart TV in the digital home. These devices increasingly depend on reliable software to offer a good user experience. However, the current digital home platform landscape is highly heterogeneous, with different operating systems resulting in barriers to achieve cross-platform development and testing processes for digital home applications. New envisaged applications could engage with information and services exploiting the context. However, context awareness for pervasive applications introduces new challenges for ensuring that the desired user experience is achieved. The hardware and software of the devices vary so many that it is difficult to achieve portability feature across platforms. Hence the current trend in developing interoperable applications is to use web technology instead of platform-specific APIs.

HTML5 and WebGL are fully aligned with this trend by providing the Web as a software platform for interoperable applications. They offer device orientation, geolocation management and 3D rendering, bringing from native features to web-centric development. However, constraints to render interoperable complex 3D environments are still present especially in digital home devices such as TVs, set-top boxes, smartphones and tablets. Results described around the browser limitations to render 3D scenes of these devices, become evident the need to improve the performance of 3D applications over the digital home browsers to satisfy the prospects of the users, even if the these devices are being fitted with improved low energy-consumption GPUs.

In order to overcome this problem, the *3DMaaS* approach introduced in this paper, deploys remote servers performing the remote rendering of complex 3D scenes and then sending the frame results to a digital home device. This video streaming server approach pushes part of the graphics generation logic to the cloud and, in essence, turns the end device into a thin terminal. Driven by latency constraints, our approach proposes a hybrid solution combining remote rendering of background 3D objects, where the latency does not have a high impact on the user experience, with local browser WebGL rendering of foreground 3D objects which require low latency. Synchronization and 3D scene consistence challenges must be managed by the HTML5 application and the related complexity depends on its domain. Experiments show the results obtained by the proposed system for hybrid remote and local rendering enhance the interactive experience of 3D graphics on digital home devices proving the feasibility of interactive navigation of high complexity 3D scenes while provides a interoperable solution that can be deployed over the wide device landscape.

Acknowledgements The authors would like to thank the Spanish Ministry for Industry, Tourism and Commerce, and the European FEDER funds for supporting the research activities involved in this paper.

References

1. Anttonen M, Salminen A et al (2011) Transforming the web into a real application platform: new technologies, emerging trends and missing pieces. In: Proceedings 2011 ACM symposium on applied computing, pp 800–807
2. Q3 Azuma R (1997) A survey of augmented reality
3. Q4 Butler M (2011) Android: changing the mobile landscape. IEEE Pervas Comput 4–7
4. Q4 Daras P, Alvarez F (2009) A future perspective on the 3d media internet. Media 303–312
5. Fechteler P, Eisert P (2010) Accelerated video encoding using render context information. In: IEEE 17th International conference on image processing (ICIP) 2010, pp 2033–2036. doi:10.1109/ICIP.2010.5651183
6. Hofmann I, Farber NFH (2011) A study of network performance with application to adaptive http streaming. In: BMSB 2011 IEEE international symposium
7. Humphreys G, Eldridge M, Buck I, Stoll G, Everett M, Hanrahan P (2001) Wiregl: a scalable graphics system for clusters. In: Proceedings of the 28th annual conference on computer graphics and interactive techniques, SIGGRAPH '01. ACM, New York, pp 129–140. doi:10.1145/383259.383272
8. Humphreys G, Houston M et al (2002) Chromium: a stream processing framework for interactive rendering on clusters. ACM Trans Graph (TOG) 10 Q3
9. Laikari A, Laulajainen JP et al (2009) Gaming platform for running games on low-end devices. In: Proceedings of ICST personalization in media delivery platforms, p 4
10. Lamberti F et al (2007) A streaming-based solution for remote visualization of 3d graphics on mobile devices. IEEE Trans Vis Comput Graph 14 Q4
11. Lu Y, Li S, Shen H (2011) Virtualized screen: a third element for cloud mobile convergence. Multimedia IEEE (2):4–11. doi:10.1109/MMUL.2011.33 Q4
12. Marino G, Gasparello P, Vercelli D, Tecchia F, Bergamasco M (2010) Network streaming of dynamic 3d content with on-line compression of frame data. In: Virtual reality conference (VR), 2010. IEEE, pp 285–286. doi:10.1109/VR.2010.5444762
13. Moreno C, Tizon N, Preda M (2012) Mobile cloud convergence in gaas: a business model proposition. In: Hawaii international conference on system sciences, pp 1344–1352
14. Nadalutti D, Chittaro L, Buttussi F (2006) Rendering of x3d content on mobile devices with opengl es. In: Proceedings of the eleventh international conference on 3D web technology, Web3D '06. ACM: New York, NY, USA, pp 19–26. doi:10.1145/1122591.1122594
15. Q4 Ortiz S (2010) Is 3d finally ready for the web? Computer (1):14–16. doi:10.1109/MC.2010.15
16. Q4 Ortiz S (2011) Bringing 3d to the small screen. Computer (10):11–13. doi:10.1109/MC.2011.303
17. Rainer B et al (2012) A seamless web integration of adaptive http streaming. In: Proceedings of the 20th European signal processing conference 2012. Bucharest, Romania
18. Shi W, Lu Y, Li Z (2010) Scalable support for 3d graphics applications in cloud. In: IEEE 3rd International conference on cloud computing, p 8
19. Simoens P, De Turck F, Dhoedt B, Demeester P (2011) Remote display solutions for mobile cloud computing. Computer (8):46–53. doi:10.1109/MC.2011.70 Q4
20. Taivalsaari A, Mikkonen T (2011) The death of binary software: and user software moves to the web. In: Ninth international conference on creating, connecting and collaborating through computing, pp 17–23
21. Taivalsaari A, Mikkonen T (2011) The web as an application platform: the saga continues. In: 37th EUROMICRO conference on software engineering and advanced applications, pp 170–174

22. Tarkoma S et al (2011) Arching over the mobile computing chasm: platforms and runtimes. IEEE Comput 22–28 Q4
23. Zahariadis T, Daras P (2008) Laso-ballesteros I, towards future 3d media internet. NEM Summit 13–15 Q4
24. Zorrilla M et al (2012) Next generation multimedia on mobile devices, mobile technology consumption: opportunities and challenges. IGI Global. doi:10.4018/978-1-61350-150-4.ch011
25. Zorrilla M et al (2012) End to end solution for interactive on demand 3d media on home network devices. In: BMSB 2012 IEEE international symposium, pp 1–6

Mikel Zorrilla is with the Department of Digital TV & Multimedia Services, Vicomtech-IK4 GraphicsMedia.net. He received his Telecommunication Engineering degree in 2007 from Mondragon Unibertsitatea, Spain.

He focuses on the research lines related to multimedia services and interactive media technologies. During his studies he worked in the area of communications of IK4 Ikerlan Research Centre (www.ikerlan.es) (2002–2006). Afterwards, he developed there the End of Degree Project about The Transport of Multimedia Traffic With an "Industrial Ethernet" Communication Bus (2006–2007). Since 2007 he is working at Vicomtech-IK4, where he designs, develops and leads projects of the Digital Television and Multimedia Services area.

Angel Martin is with the Department of Digital TV & Multimedia Services, Vicomtech-IK4 GraphicsMedia.net. He received his engineering degree in 2003 from University Carlos III, Spain.

He starts his Phd in the Communications and Signal Processing Department of the University Carlos III in the video coding research area in 2003. At the same time he collaborates with Prodys developing a standard MPEG-4 AVC/H.264 codec for DSP. In 2005 he starts to work on Telefonica

I+D in projects related to multimodal interactive services for Home Networks. Also in Telefonica I+D, he goes deeper into image processing area in terms of 3D video and multiview coding. From 2008 to 2010 he worked in Innovalia as R&D Project consultant related with smart environments and ubiquitous and pervasive computing. Currently he is on Vicomtech-IK4 managing and developing R&D projects around multimedia content services.

Q5 **Jairo R. Sanchez**

Q5 **Iñigo Tamayo**

Igor G. Olaizola is the head of Digital TV & Multimedia Services Department in Vicomtech-IK4 GrpahicsMedia.net, Spain. He received his Engineering degree from University of Navarra, Spain (2001).

He developed his Master thesis at Fraunhofer Institut für Integrierte Schaltungen (IIS), Erlangen—Germany in 2001 and currently he is preparing his PhD in Computing Science and Artificial Intelligence at University of Basque Country. He has participated in many industrial projects related with Digital TV as well as several European research projects in the area of audiovisual content management. His current research interests include multimedia content analysis frameworks and techniques to decrease the semantic gap.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

Q1. Please check authors' name if captured correctly.
Q2. Please check authors' affiliation if captured correctly.
Q3. Please provide complete bibliographic data for reference items [2] and [8].
Q4. Please provide volume number for reference items [3], [4], [8], [11], [15], [16], [19], [22] and [23].
Q5. Please provide bios and photos for "Jairo R. Sanchez and Iñigo Tamayo".