701 E. Chocolate Avenue • Hershey PA 17033-1240, USA
Tel: 717.533.8845 • Fax: 717.533.8661

Premier Reference Source: The Book Development Process

Next Generation Multimedia on Mobile Devices

Mikel Zorrilla
mzorrilla@vicomtech.org
Vicomtech research centre, Spain
María del Puy Carretero
mcarretero@vicomtech.org
Vicomtech research centre, Spain
Alejandro Ugarte
alexugarte@gmail.com
Vicomtech research centre, Spain
Juan Felipe Mogollón
fmogollon@vicomtech.org
Vicomtech research centre, Spain
David Oyarzun
doyarzun@vicomtech.org
Vicomtech research centre, Spain
Igor García Olaizola
iolaizola@vicomtech.org
Vicomtech research centre, Spain

ABSTRACT

The multiplatform consumption of multimedia content has become a crucial factor in the way of watching multimedia. Current technologies such as mobile devices have made people desire access to information from anywhere and at anytime. The sources of the multimedia content are also very important in that consumption. They present the content from many sources distributed on the cloud and mix it with automatically generated virtual reality into any platform. This chapter analyzes the technologies to consume the next generation multimedia and proposes a new architecture to generate and present the content. The goal is to offer it as a service so the users can live the experience in any platform, without requiring any special abilities from the clients. This makes the architecture a very interesting aspect for mobile devices that normally do not have big capabilities of rendering but can benefit of this architecture.

INTRODUCTION

Multimedia Content landscape is changing very quickly. Social networks are enabling a prosumer and collaborative attitude from the community. Social networks and other popular sites let users create and share information with other people. Their popularity in recent years has increased the active participation and content production of the web because these systems often contain blogs, photo galleries and other means for sharing digital content (Bonhand et. al., 2007). In addition consumers are demanding multimedia content access from everywhere and at anytime through any device, being the use of mobile devices the most popular and increasing trend of the last years.

The term SaaS (Software as a Service) is being used to refer to an abstraction devoted to offer distributed services over the cloud. SaaS can be a possible replacement to traditional software, where buyers obtain a permanent license and install and maintain all necessary hardware, software and other technical infrastructure (Choudhary, 2007).

Taking this term as a reference, the term SaaS can be extended to MaaS (Multimedia as a Service) where the user wants to consume multimedia content in “service mode”, adapting the content properties and delivery to users’ needs and context. Moreover, the term can be further extended to everything available on the Internet as XaaS (Everything as a Service). This fact allows universal access and enables the creation and combination of new contents and services.

There are also some emerging technologies, such as 3D content. Integrating these technologies into multimedia content, the user can enjoy a multimedia experience where he/she can find natural interfaces, virtual and augmented reality environments and automatically generated content personalized to each user.

With the aim of supporting this multimedia experience, Quality of Experience (QoE) is a crucial fact which, depending on the context of the content consumption, decides the most efficient Quality of Service (QoS) values.

Therefore, the main goal of this chapter is to propose an architecture that allows the access to the next generation multimedia on any device without needing a dedicated hardware in the client side where mobile devices can benefit of this architecture.

BACKGROUND

What is Multimedia Technology?

According to Vaughan (1993), “multimedia is any combination of text, graphic art, sound, animation, and video that is delivered by computer”. In addition, Vaughan explains that when users are allowed to control what and when these elements are delivered, it is called interactive multimedia. Besides, providing a structure of linked elements through which the user can navigate, interactive multimedia becomes hypermedia.

In common usage, the term “multimedia” refers to an electronically delivered combination of media including video, images, audio, text and 2D or 3D animations that can be accessed interactively.

Nowadays the expression multimedia is frequently linked to natural interaction because it provides different ways to access the content and thanks to interactive multimedia, users can decide the means to inform themselves.

The scope of multimedia is very extensive. The influence of multimedia content from education to business makes people be attracted to more catchy and interactive information.

The Evolution of the Multimedia Content

Most of the problems of the multimedia are related with how to manage big quantity of data so the compression of the multimedia is a crucial point in order to reduce the storing and transferring data.

The main access to video content occurred when the VHS won the Videotape format war. The first Video Cassette Recorder appeared in 1972 by Philips. However, the first system that succeeded was Sony’s Betamax in 1975. Immediately the Japanese enterprise JVC presented Video Home System (VHS) and in 2000 Philips developed the Video 2000. Instead of developing a single format with the benefits of the different systems, a format war started and VHS became the most popular. This made possible to have a video recorder in most of the houses and popularize it.

Very quickly the Compact Disc became very popular using an optical disc to store digital data. The CD-ROOM that was created by Philips and Sony in 1979 became a popular media to distribute multimedia applications and contents. At the moment there is a new format war with the high definition media format over an optical disc, between Blu-ray Discs and HD-DVD.

The Internet and its increasing bandwidth have influenced on the way multimedia is consumed. Internet brought many advantages to people providing them access to any kind of information instantly. Web pages are not only composed by text but also by images, audio and video. In addition several pages include animations in order to attract the visitor’s attention. The variety of media offering information has been the key to the success of multimedia content.

Web 2.0 has brought collaborative content creation where people share information through different manners such as blogs, wikis, social networks, etc. In addition, new technologies and devices like mobile phones have given society access to this information everywhere and at anytime thanks to ubiquitous technologies.

Nowadays, new current technology brings the opportunity to access information at anytime and anywhere and using any device. For example, people can get information outside their homes using mobile devices. Mobile technology is being developed very fast and users can have nearly the same experiences with their mobile devices than while using their personal computers. Multimedia content is frequent at mobile devices. For example people can communicate with others thanks to MMS (Multimedia Messaging System) adding images, videos or music to the written text.

Finally, in recent years, talking about web 3.0 new formats and methods of interaction are being integrated on the Internet. As an example 3D graphics are demanded by users. Virtual worlds and social networks are more frequent in our daily lives and they integrate multimedia to the information that they bring. Users and developers are demanding more immersive interaction. What is more, the Law demands universal accessibility in order for everybody to have the same rights. This is the reason why natural interfaces and personalization as well as virtual and augmented reality are fundamental technologies for the next generation multimedia.

Figure 1 shows the evolution of the multimedia content and its most important terms over time.

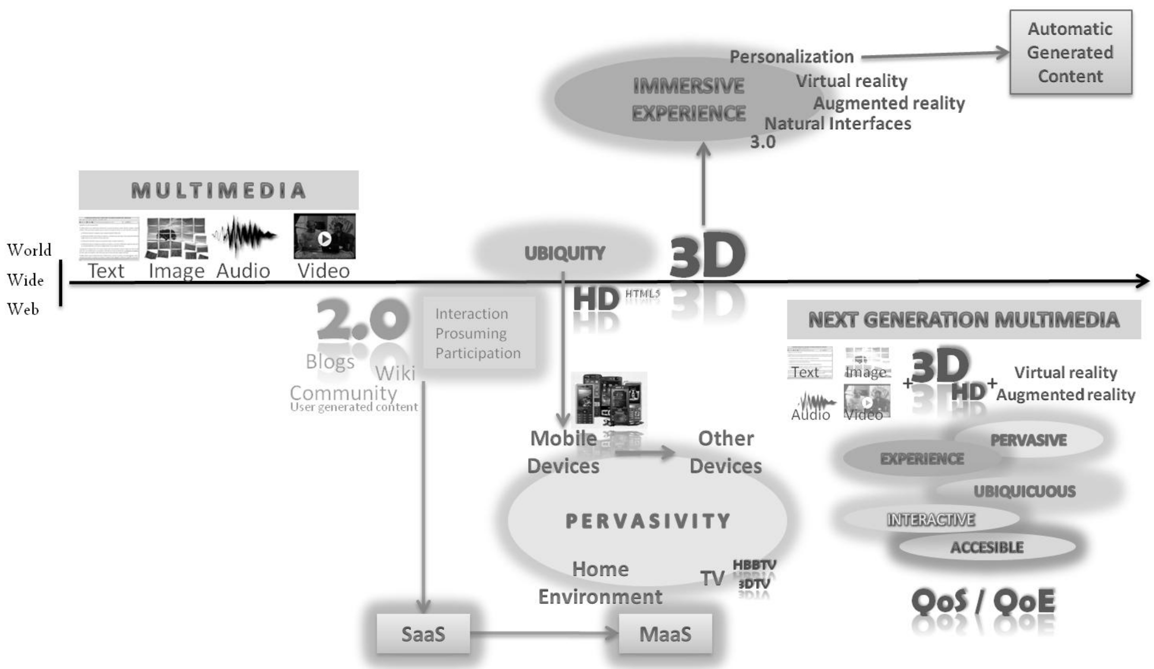

Figure 1: The evolution of the multimedia content and its most important terms

Limitations in current multimedia consumption

Multimedia offers the possibility of using different types of content as a complement to the given information. Thanks to multimedia, users can read, watch, listen and view this information and interact with it. However there are some lacks or disadvantages that should be resolved in order to allow a more immersive interaction.

Although multimedia content gives the opportunity of having an interactive experience with the content or the information, the user is often forced to interact in a certain way. The interaction is not always natural and the user has to adapt oneself to interact with the system. That is a drawback when users have a handicap and they have no choice to select the most appropriate interaction mode for them. Even though it has been said that multimedia is often linked to natural interaction, there are still some limitations to be resolved in order to offer natural interaction to each person.

Another main deficiency is that multimedia applications usually require specific hardware. There are still technological issues to achieve ubiquitous and context aware multiplatform publication infrastructures. To reach this hardware independent many-to-many approach (from many content sources, to many end user devices), a new architectural design is needed.

With the aim of improving the explained limitations, the purpose is to offer a more immersive experience while users interact with multimedia content. There are several new technologies that can be integrated as multimedia content that allow the user to have interaction. Virtual and augmented reality, personalization and natural interfaces are technologies that can be integrated with multimedia content that let users have a more immersive experience while they interact with any content, system or device.

MULTIMEDIA EXPERIENCE

Collaborative content creation

Collaborative content can be created by several ways (Traunmüller, 2010), for example as wikis which are knowledge collections built by collaborative editions; as blogs where people create notes open to be commented by others; or tagging that makes sharing information possible.

Besides, social networks include collaborative content. When a computer network connects people, it is called social network (Wellman, 1997). A social network is a set of people or a social entity connected by a set of socially-meaningful relationships.

Apart from the business behind social networks, their purpose is to communicate or interact with other people in addition to share information with them. Well known social networks like Facebook, YouTube or MySpace allow communication with others by multiple media, where users can share information or/and their thoughts or opinions by text, pictures, videos, music, etc. As Bonhand et. al., (2007) explains, some of the most successful sites let users create and share content, and users can connect and communicate with others and thus provide a richer and more interactive user experience.

Virtual worlds are also social networks, for example Second Life or Imvu. Thanks to these 3D social networks people can communicate with others in a more immersive way. Developers of virtual worlds are integrating multimedia content to offer multimedia experience.

The role of multimedia content on the collaborative Internet is very important. It has allowed people to share their information through different media, enriching the users' experience. But also Internet has helped building (to build) collaborative multimedia applications (Nicol et. al., 1999).

Wrapping Multimedia Content in Services (MaaS)

According to Turner et. al., (Turner et. al., 2003) SaaS (Software as a Service) "is one of a demand-led software market in which services are assembled and provided as and when needed to address a particular requirement".

The use of this term is being very extended on the Internet community. Developers of software and business professionals have a tendency to associate the term SaaS with business software as a probably lower-cost method for businesses. Through a well-designed implementation and proper licenses, SaaS provides benefits without associated complexity and high cost to equip devices with applications they might not need.

Taking all these ideas into consideration, the term SaaS can be extended to MaaS (Multimedia as a Service) where users consume multimedia content as a service. For this purpose, it is necessary to adapt the content properties to users' needs as well as to the context. The presentation of multimedia has to be a completely transparent service where the user does not need to know underlying details like which are the sources that have been used to create the content or where the rendering process happened which are giving him an immersive multimedia experience.

Additionally, this term can be extended to XaaS (Everything as a Service) that means everything available through this high layer abstraction. The XaaS paradigm moves the Internet towards the universal access of every digital resource.

Immersive Experience

According to several authors as Dede (2009) shows, immersion is "the subjective impression that one is participating in a comprehensive, realistic experience". Taking this into consideration, providing immersive experience to the users is to allow people to have a realistic experience using new technologies.

The term immersion is often linked to virtual and augmented reality. However, there are other new technologies that allow offering immersive experience to the users possible while they are accessing to multimedia content so as to complement it and enrich users' experience with the content.

Immersive experience complements the access to the multimedia content giving a more complete idea of the content and allowing the users to have other or a complementary perspective of it. That means that not only can the user watch, read and/or listen to the content but users can also feel immerse on it. This enriches users' experience giving them the illusion for example of interacting with an object something that otherwise cannot be possible.

The use of immersive experience can be linked to any scope, not only is it connected to games but it can also be applied to work, education, entertainment, people care, etc. Nowadays, people are demanding new technologies and a more immersive experience in order to make their life easier and comfortable.

The following technologies have been detected as very significant to offer immersive experience to the user: virtual/augmented reality, personalization and natural interfaces.

Virtual/Augmented Reality

Nowadays technologies like virtual and augmented reality are at their very peak. Although they are more associated to games and entertainment, new applications try to make use of them to enrich their characteristics no matter which is the intention of these applications. But, what are these technologies and what is their purpose?

One of the given definitions of virtual reality (VR) is "the use of computers and human-computer interfaces to create the effect of a three-dimensional world containing interactive objects with a strong sense of three-dimensional presence" (Bryson, 1996). The purpose of

VR is to simulate places or situations that can take place in a real life but in a virtual environment.

Azuma (1997) described augmented reality (AR) as a variation of virtual reality. According to his definition, VR technologies "completely immerse a user inside a synthetic environment. While immersed, the user cannot see the real world around him. In contrast, AR allows the user to see the real world, with virtual objects superimposed upon or composited with the real world".

Thus virtual and augmented reality technologies help to improve the experience of the user while they interact with the system, application or device because these technologies allow the users to live a situation that emulates a real one or complement it with additional information.

Features and advantages

Some authors talk about immersive virtual reality where the user becomes totally immersed in an artificial and three-dimensional world that is completely generated by a computer.

The characteristics of immersive VR can be resumed as:

- Views are three-dimensional perspective and often stereoscopic which enhances the perception of depth and the sense of space.
- The point of view is often egocentric. The views change matching the physical position of the users and the direction of his/her gaze. That provides the immersion of the user in the virtual environment.
- The environment is usually dynamic.
- Multi-sensory techniques are employed to interact with the user. User input can be multimodal. That means using whatever voice, text, gestures, etc.
- Realistic interaction can be possible thanks to gloves or similar devices that provide manipulation and control of virtual objects.

According to Nooriafshar et., al.(2004) the advantages of VR is notable on situations where:

- Access to the real object or environment is difficult or impossible.
- Using the real object is unsafe or has a health hazard for people.
- Experimenting with the real object is too expensive.

An augmented reality system has the following three characteristics (Azuma, 1997):

- it combines virtual and real (reality);
- it is interactive in real time;
- and it is registered in 3D.

Virtual and augmented reality technologies are applied in a wide scope of applications from e-learning systems to surgery training tools.

Natural Interfaces

People’s communication is made through gestures, expressions and movements. According to Valli (2004) natural interaction is made through systems that understand these actions and engage people in a dialogue. In addition a principal feature is that it is not necessary that people wear any special device or learn any instructions while they interact.

Natural interfaces are those in which the user can interact with the interface through gestures, speech, touch, vision and/or smell based interaction. In other words, natural interaction involves “actions and sensations that refer to our daily life experiences” (Mignonneau et. al., 2005).

As Mahoney (2000) explains, the most effective user interface is one that is comfortable and natural. And the most comfortable and natural interface is one that is nearly invisible. That means that there is no separation between control and the presentation of the information. One of the characteristics of a successful natural interface is the reduction of the cognitive load on people while they interact with it (Valli, 2006).

An example of the advantages of Natural Interfaces is that it can enhance visitors’ experience at a museum or exhibition (Alisi, 2005). At their work, the authors show a system which allows visitors to have a natural interaction with works of art. Visitors do not need to wear a special device to interact with the system, tested at a museum in Florence.

Allowing natural interaction between human and objects in a virtual world provides reality (Murakami, 1991). Besides, as Mankoff et. al., (2002) explains natural interfaces are particularly useful in settings where a keyboard and mouse are not available such as a very small or very large display and with mobile and ubiquitous computing.

Not forcing the consumer to interact in a specific way would satisfy user experience. In other words, it would be appropriate that the user could choose the interaction mode that he/she need or prefer. Nowadays, it is frequent to use multimodal interaction which implies visual, touch, gesture and voice interfaces. However it is often difficult to translate all of these manners to mobile devices because of the current technologies. The architecture explained in this chapter solves this problem and makes it possible to access to natural interfaces on any device without a dedicated hardware.

Personalization

With the purpose of universal accessibility, it is necessary to personalize and adapt multimedia content to users, not only to their preferences but also to their needs.

There are two terms associated to personalization: customization and adaptation. The first one, customization is the personalization requested directly by the user. Adaptation means personalization that is automatically performed by the interface or the system (Weld et. al., 2003). As an example to view the difference between customization and adaptation is Miis and the game Wii Fit from Nintendo. A Mii is a virtual character that represents the player, who can decide or customize the appearance of his/her Mii. The system (game) by itself adapts the figure of the Mii according to the real weight of the player. In this case the user cannot decide if his/her Mii is slim or fat, it is adapted by the game according to reality.

Customization is not a new concept. There are many desktop user interfaces that allow the user select the tools they prefer on the menus and toolbars. However nowadays, current technology gives the opportunity to access to nearly the same content from different devices or environments. The problem is that the appearance of the users interface can change because of the characteristics of the device or the environment. For example, the different size of the screen between PCs and mobile devices makes the user interface to be different in many cases and the customization made on PC cannot be possible otherwise.

Regarding users' preferences, it is frequent to think about colors, fonts, size, etc. However, the preferences can involve the order of the content, users' tastes or the place from where people access to the multimedia content. That means that for example, if the user is at home, s/he may prefer access to the content through audio. Conversely, if s/he is outside, possibly the best way could be the access through reading text so nobody else can listen to the information.

As for users' needs, a simple rule that any content should fulfill is that everybody should have access to it. That means that people who have a handicap could have the same rights and/or benefits than the rest of the society. For that reason, multimedia content has to be defined taking into account possible handicaps. In addition the information should be given in different formats, such as video, audio, text, etc. so it can be accessible from different ways. Moreover, contents have to be adapted to assistive technologies to transmit the information to the users. In this way, everybody can access to the information and interact with new technologies.

Thereby personalization should be linked to natural interfaces so that people were able to interact with the interface or the system in a natural way according to their skills and capabilities or their desires. By this way, users could access and interact with multimedia content in a natural way according to their personal interface.

PROPOSSED ARCHITECTURE

Architecture to address the problem of generating and presenting the next multimedia content is presented in order to get the content from many sources in combination with virtual reality and present it in any platform. This solution is device independent and generates and presents multimedia content without a dedicated hardware on the client side.

Due to the lack of support to render computer graphics in most of the platforms, a solution of rendering the multimedia content (augmented reality) in a server and delivering the video using streaming technologies has been adopted.

This architecture only needs a few render capabilities at the client side, such as an RTSP video client that most of the devices have, and it delegates all the technological challenge to the server.

Another main characteristic of the architecture proposed is its device independence. That means that each type of device that is connected to The Internet (no matter the connection medium: 3G, Ethernet, Wireless, etc) should be able to receive the content. It is not dependent of the codecs that are supported on the device, its screen size, the bandwidth of the network connection, the needs of the person that is accessing and watching the content, etc. This is possible thanks to the server makes all the work as well as differences among different devices. Thus, it takes parameters from the context, the device and the user to select the stream that fits for that connection (video codec, bit rate, resolution, complexity of the augmented reality content, etc).

Implemented architecture

Current implemented architecture (see Figure 2) is based on three main parts;

- a Web Service that communicates with a Client device via HTTP,
- a Gstreamer RTSP server that is connected to these Web Services,
- and the Client device that communicates with the Web Service to start a communication and receives the content via a RTSP stream maintaining an interaction via.

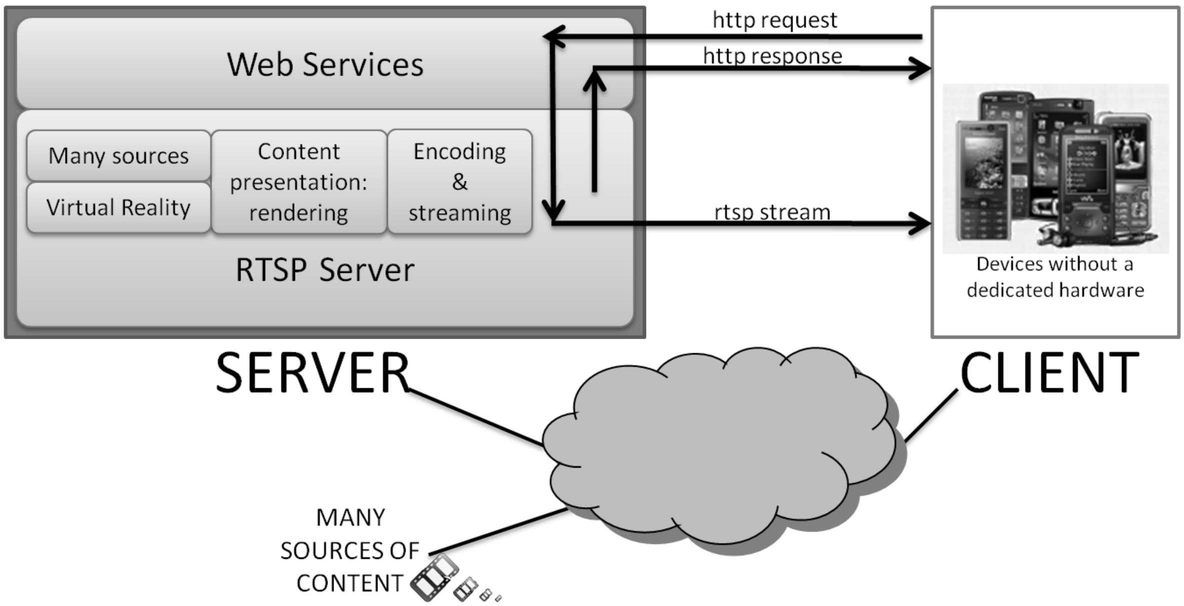

Figure 2: Implemented architecture's diagram

Client

The Client device does not require advanced rendering capabilities. It only needs to be able to render an RTSP stream that should be already fitted by the Server depending on the device. This point makes the mobile phones and other mobile devices ideal for this architecture because they give the user the possibility to consume content anywhere but they normally do not have capabilities to render 3D computer generated content. Thanks to this solution, the user has access to a high characteristic multimedia content anywhere and with a common device.

Web Service

The Web Service server is the responsible of receiving the requests from the Clients. It includes different calls that can be used to request new content and interact with it. For each devices connection it starts a new slot with the content generation, rendering and codifying it, to send to the user.

Several services have been implemented as a Web Service that can be called by the Client. For example a Web Service called *prepareConnection* provides the RTSP URL that gives the stream to the client. This call receives different arguments to receive information about the needs of the user that is asking the connection, the context and the device which is using in order to adjust the parameters of the stream to the best values of the situation. This call returns to the client the RTSP URL that can be understood and showed by the device.

Other services are developed in order to let the user set values about the content which is being watched at that moment and the way to interact with it in real time. This makes it possible that during the receiving of the stream the user is able to call these services and watch immediately the changes in the stream.

RTSP Server

This module presents the content to the user composing a content created from many sources on the cloud and adding 3D computer generated content. It renders it following the user's needs and encodes the stream according to the clients device needs.

It is based on Gstreamer and has the capability to modify the properties of each component during running time, providing interactivity with the different devices.

There are two main parts inside this RTSP Server:

- A Gstreamer module that provides the integration of the 3D computer generated content to the main stream and allows that content to mix with other multimedia content on the cloud in order to get the desired Augmented Reality. Its properties can be modified at running time, so it is possible to interact with that content in real time and change its parameters.
- Encoding and streaming module. This module is responsible of encoding the already generated media in the appropriate way for each device and the connection context.

Architecture's Workflow

In Figure 3 the sequence diagram of the architecture is presented from a device connection request to the interaction with the media stream.

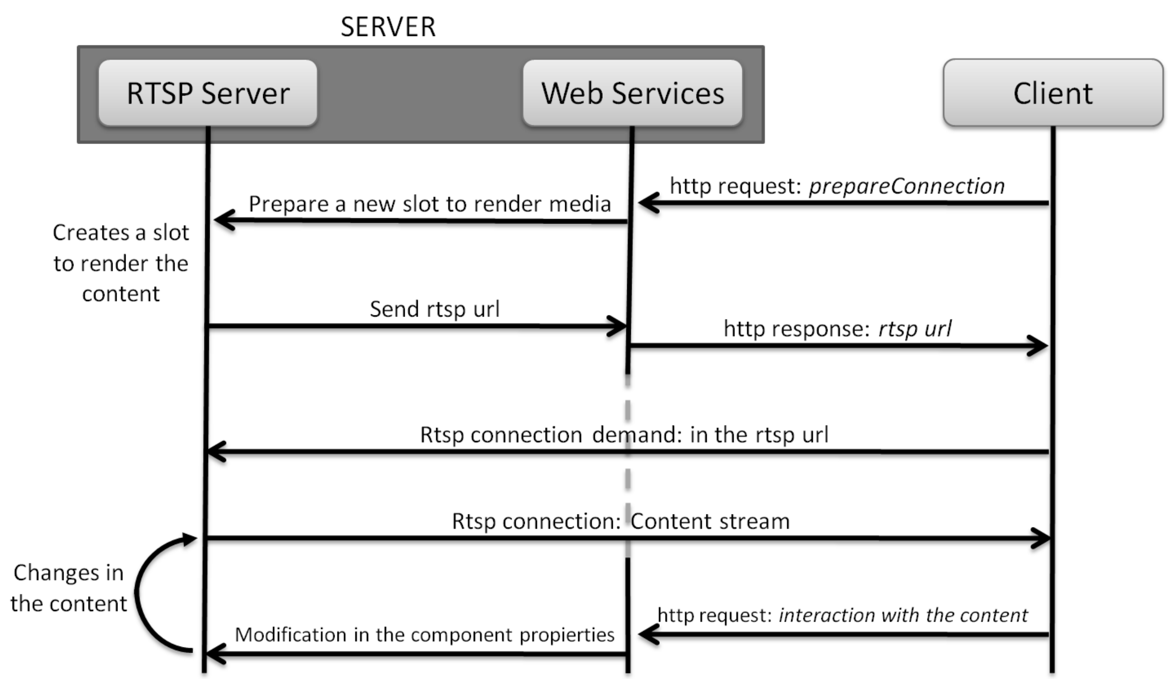

Figure 3: Implemented architecture's sequence diagram

Validation of the Implemented Architecture

A scenario for the validation of the mentioned architecture has been developed. Therefore a 3D computer generated multimodal virtual character was used with the intention of being users' assistant while they are navigating on web pages.

PHP files were developed where HTML web pages are created and an RTSP player is embedded in. PHP asks the Web Service to prepare a Virtual Character with specific characteristics (female or male, eyes color, age, voice, etc) and the Web Service answers a RTSP URL with the video stream. That is shown in the embedded player together with the other entire HTML page. During the navigation of the user, PHP calls more services included in the Web Services such as *speakAvatar*, sending the text and the language of the sentence that wants to hear as parameters. Other services are implemented to provide animation to the Virtual Character (face and body expressions, voice changes, etc). This guides the user during the experience of the navigation into the web pages.

Different tests have been done to consume these web pages including the Virtual Character stream from PCs and from mobile devices, such as mobile phones and PDAs with satisfactory results.

During the validation the researchers realized that the architecture proposed could be used as a preliminary good solution. However some points that needed to be improved were checked. On one hand, the server needs to render and codify an avatar for each client, so one server can provide service to a limited value of clients. A distributed server farm has to be design in order to have a scalable capability to answer to a lot of clients in a real implementation. On the other hand, the interaction reaction time must be improved to be able to watch immediately the requested changes in the stream in our device. A possible solution for these points is given in the next section as a scalable proposal of architecture.

Scalable Architecture Proposal

As it was explained in the previous section two main issues were detected to improve the implementation developed of the architecture: the scalability on a real implementation and the response time of the interaction requests.

In this section an extended architecture is proposed, taking into account the detected drawbacks in the implemented architecture (See Figure 4):

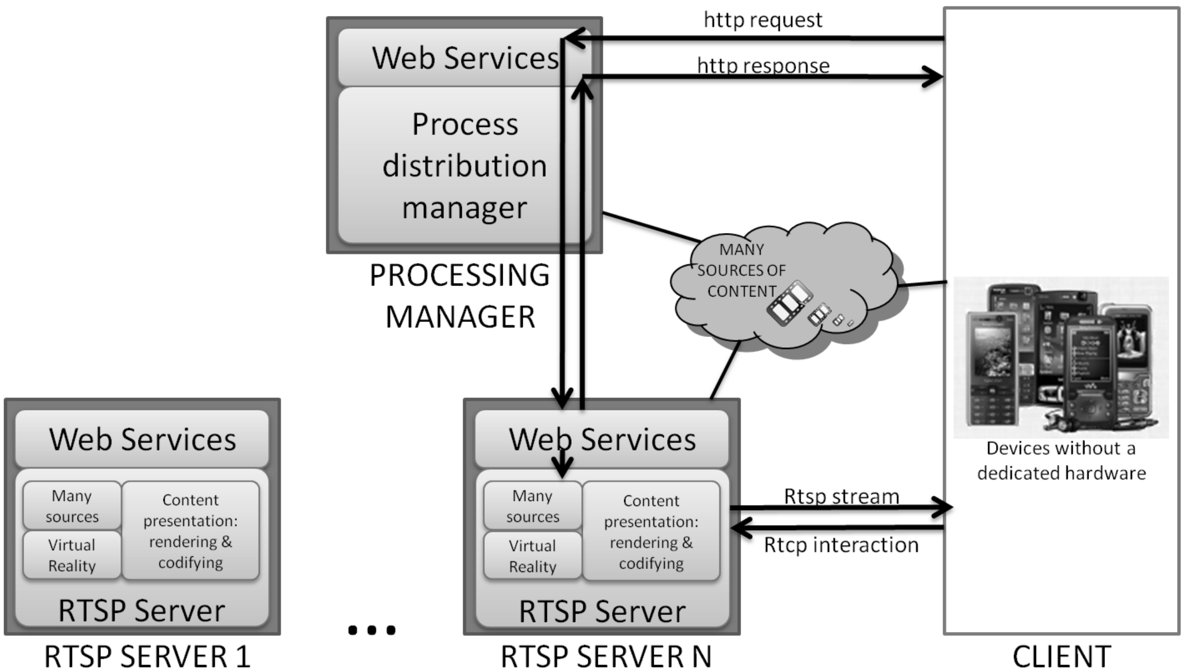

Figure 4: Scalable architecture's diagram

For this scalable proposal there is a completely new module called Processing Manager. This module integrates the Web Services shown on the implemented architecture but does not include any RTSP server. This module receives the Web Service requests and manages these requests to redirect the processing requested to the appropriate RTSP Server available on the cloud. So there will only be a Processing Manager to reply all the requests, but there will be unlimited RSTP Servers to render and encode all the multimedia content. The Client will communicate at the beginning with the Processing Manager, but afterwards it will be redirected to an RTSP connection with one of the RTSP Servers. The different modules are explained with more detail.

Processing Manager

This module has to manage the client's requests by forwarding them to the RTSP Server. The Processing Manager should have to know which processing capabilities that request needs and know which RTSP Server is available to render that process. This module has to manage the number of RTSP Servers on the cloud and their processing status in order to be able to forward the requests.

RTSP Server

This module includes the Gstreamer-based RTSP Server presented on the implemented architecture with a new layer of Web Services that make it possible for it to communicate with the Processing Manager. This way it is possible to add as much RTSP Servers as necessary on the cloud and only informing the Processing Manager about its availability they can start to work, even in run time. This new Web Service layer would not be public and it would only be known by the Processing Manager.

The RTSP Server receives new connection requests from the Processing Manager and informs the RTSP URL where to give the stream. The Processing Manager is the one that informs the Client the address of the RTSP stream but after that any other communication occurs between Client and RTSP Server.

The RTSP Server also communicates the Processing Manager their processing status after each modification in order to notify their availability for further processing.

The interaction channel can also be directly opened between the Client and the RTSP Server using Real Time Control Protocol (RTCP). That way the response time for the interaction is improved. The RTCP has been developed to provide a feedback of the Quality of Service (QoS) of the RTP communication, but it can also be used to transmit additional information over the stream.

<u>Client</u>

The changes shown on the scalable architecture does not have any effect on the client . It asks the Processing Manager for a new connection and the only difference is that the RTSP URL received is not located in the same physical server even if it belongs to one of the other RTSP Servers.

Once the connection is established, the client should interact with the content using the RTCP protocol.

In Figure 5 a sequence diagram of the presented scalable architecture is shown.

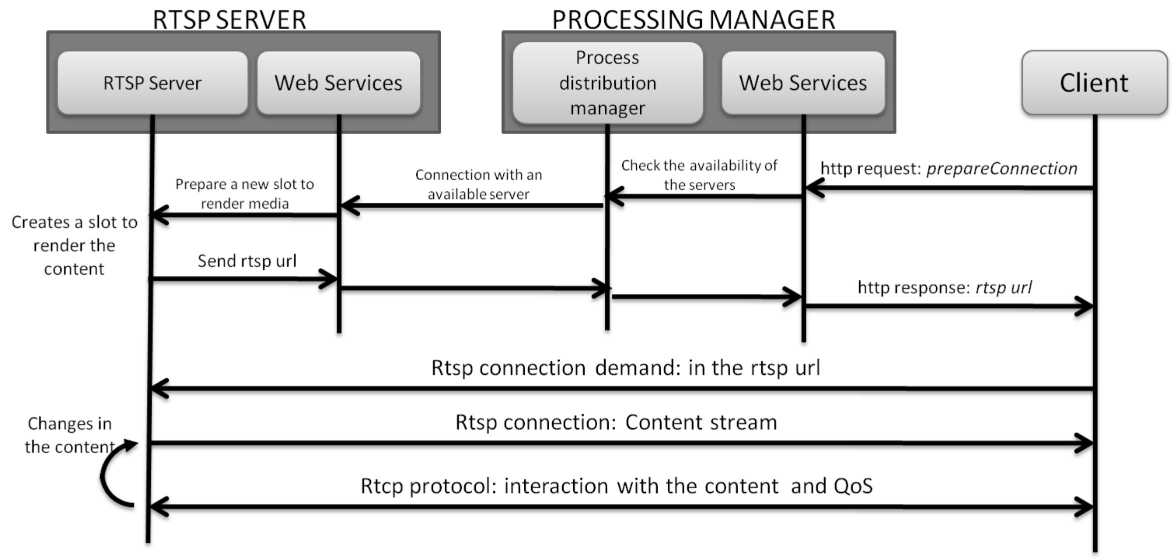

Figure 5: Scalable architecture's sequence diagram

Benefits of the architecture

Two architectures have been presented here; an implemented one and an improved architecture that tries to fix the weaknesses detected on the implementation. However both of them pursue the same goals. First of all, it provides multimedia content as a service on the cloud approaching the Multimedia as a Service (MaaS) concept previously mentioned. The client does not have to worry about codecs, does not need specific hardware on the client's device to render advanced multimedia content (powerful video cards, etc), it only has to ask for the service and the Server provides the best context adapted content.

FUTURE RESEARCH DIRECTIONS

This architecture provides a multimedia consumption service to users, that allows them to access to a multimedia world using any device. A dedicated hardware is not necessary on the client and this brings several benefits to the users. For this purpose, all the data processing for the rendering of the automatically and semi-automatically generated multimedia content must be on the server's side.

An inversion of server farms is needed in order to give a massive service to the user. That is why the atomization of the render process into different and distributed servers is very important. This is the main point in the future research in order to improve the architecture.

Furthermore it is necessary to research the best way to integrate the interaction orders from the client to the RTSP Server over the RTSP stream. An adaptation of the RTCP protocol for this use could provide a very good response time providing the user a better experience. Thus it will not be important anymore whether the render of the content is being done locally or remotely.

CONCLUSIONS

This chapter makes a revision of the multimedia content, analyzing its evolution and evaluating the consumption limitations it nowadays has. A necessity of including immersive experience has been detected in order to empower users' experience while they interact with multimedia content.

In order to integrate immersive experience, new technologies are needed. Not only virtual and augmented reality technology has been detected, but also personalization and natural interfaces. The multimedia content can be the interface providing automatically generated and personalized content with an efficient interaction channel. All of these technologies can improve users' experience and allow them to consume next generation multimedia content from many sources through any device.

A new architecture to present the next generation multimedia has been presented. The main characteristic of this architecture is that it does not require a dedicated hardware in the client side. As a result, less powerful devices such as mobile devices can get a profit of this architecture.

To conclude, next generation multimedia is appearing and industry is offering solutions to enrich user's experience and give them a service to consume it from any sources, anytime and in any platform.

REFERENCES

Alisi, T. M., Bimbo, A. D., & Valli, A. (2005, July). Natural interfaces to enhance visitors' experiences. *IEEE Multimedia Vol. 12, Issue 3* (pp. 80–85).

Azuma, R. (1997). A Survey of Augmented Reality. Presence: *Teleoperators and Virtual Environments, Vol. 6, No. 4*, (pp.355-385).

Beier K.P. (2004). Virtual Reality: A Short Introduction. Retrieved September 2, 2010 from http://www-vrl.umich.edu/intro/

Bryson, S. (1996). Virtual Reality in Scientific Visualization. *Communications of the ACM Vol. 39(5)*, (pp. 62-71).

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science*, 323(5910), 66–69.

Choudhary, V. (2007, January). Software as a Service: Implications for Investment in Software Development, *International Conference on System Sciences*, (pp. 209)

Mahoney, D. P. (2000, February). Innovative Interfaces. *Computer Graphics World. Vol. 23(2)*, (pp. 39-42, 44).

Mankoff, J., Abowd, G. D. & Hudson, S. E. (2000). OOPS: A Toolkit Supporting Mediation Techniques for Resolving Ambiguity in Recognition-Based Interfaces. *Computers and Graphics Vol. 24 (6)* (pp. 819-834).

Mignonneau, L. & Sommerer, C., (2005). Designing emotional, methaforic, natural and intuitive interfaces for interactive art, edutainment and mobile communications. *Computer and Graphics, Vol. 29* (pp. 837-851).

K. Murakami & H. Taguchi,(1991). Gesture Recognition Using Recurrent Neural Networks, Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology,(pp. 237-242) New Orleans, Louisiana, United States.

Nicol, J. R., Yechezkal, S. G., Paschetto, J., Rush, K. S., & Martin, C. (1999). How the Internet helps build collaborative multimedia applications. *Communications of the ACM*, 42(1), 79-85.

Nooriafshar M., Williams R. & Maraseni T.N. (2004). The Use of Virtual Reality in Education, *The American Society of Business and Behavioral Sciences (ASBBS) Seventh Annual International Conference,* Cairns, Queensland, Australia, 6th-8th August.

Traunmüler, R. (2010). Web 2.0 Creates a New Government. *Electronic Government and the Information Systems Perspective. Lectures Notes in Computer Science, Vol 6267/2010*, (pp. 77-83)

Turner, M., Budgen, D. & Brereton, P. (2003, October) Turning Software into a Service. *Computer, vol. 36, no. 10,* (pp. 38–44).

Valli, A. (2006). The design of natural interaction. *Multimedia Tools And Applications, Volume 38, Number 3,* (pp. 295-305).

Valli, A. (2004) Notes on natural interaction. Retrieved August 20, 2010 from http://www.citeulike.org/user/eckel/article/4324923

Vaughan, T. (Ed.). (1993). *Multimedia: Making It Work.*, Osborne/McGraw-Hill, Berkeley.

Weld D, Anderson C, Domingos P, Etzioni O, Gajos K, Lau T & Wolfman S (2003) Automatically Personalizing User Interfaces, *Proceedings of the International Joint Conference on Artificial Intelligence, (IJCAI03),* Acapulco, México

Wellman, B. (1997). An electronic group is virtually a social network. In S. Kielser (Ed.), *Culture of the Internet* (pp. 179–205). Mahwah, NJ: Lawrence Erlbaum Associates.

Key Terms & Definitions

Keyword: Next generation multimedia, mobile devices, Multimedia as a Service, 3D content, natural interfaces, personalization, virtual/augmented reality, multisource, multiplatform.