Video Analysis in Interactive Application of Digital Television

CRISTIANE ZAKIMI CORREIA PINTO, WAGNER LUIZ ZUCCHI
Department of Electronic Systems Engineering – Polytechnic School
University of São Paulo
Av. Prof. Luciano Gualberto, Trav. 3, No. 158, Sala C2-23, São Paulo/SP, 05508-010
BRAZIL
cristiane.zcp@gmail.com   wzucchi@lps.usp.br   http://www.lps.usp.br

Abstract: - A digital television system can provide sounds and images of better quality than an analog broadcast system. But that is not all. The Brazilian digital terrestrial television system adopted in 2006 also makes it possible to receive TV signals on portable devices and mobile television receivers besides enabling interactivity. This paper consider digital terrestrial television, which is the digital television broadcast over the air, free of charge to all population, and presents a video analysis based in the Quality of Experience (QoE) that was made measuring the user perception with an interactive application of digital terrestrial television. Such an application was developed using NCL (Nested Context Language), which is the standard declarative programming language of Brazilian digital television system called *Integrated Services Digital Broadcasting - Terrestrial Brazil* (ISDB-TB). In that application a secondary video is loaded through a broadband Internet access simultaneously with the main video being received through broadcasting. A test platform was created where IP packet loss was introduced in a controlled way affecting the secondary video, as it is expected to occur in a real network. Video quality was assessed for each loss level with objective metrics in order to compare QoE in each situation. The objective of this research is to verify if it is feasible to use the broadband Internet as a return channel of an interactive application of Brazilian digital television.

1 Introduction

At the same time that transmedia productions are becoming more common, legal changes are occurring in Brazil that bring an increase on daily interactivity use.

The growth of the broadband access to Internet [1], and the adoption of a digital terrestrial television system that yields sounds and images of better quality when compared to legacy analog broadcast system and also makes it possible to receive TV signals on portable devices and mobile television receivers are factors enabling interactivity.

Digital terrestrial television is the digital television broadcast over the air, free of charge to all population. It is different from digital cable television or digital satellite television that are paid TV services.

In Brazil, it can be stated that television is a major media platform. According to data from the Brazilian Institute of Geography and Statistics [2], the television set is present in 96.9% of Brazilian households, what is more than the number of households with a refrigerator (95.8%). Only the stove is found in more homes than a television set (98.6%).

Considering the number of potential users of digital terrestrial TV, the use of interactive applications including this media platform is an important aid to the execution of a transmedia narrative.

One of the possibilities in the Brazilian Digital Terrestrial Television standard is to use the broadband Internet as the main interactive channel [3]. Considering this possibility, this study of Quality of Experience (QoE) perceived was made based on the assessment of the video quality obtained when using an interactive transmedia application of digital terrestrial television. Such an application was developed using NCL (Nested Context Language), which is the standard declarative programing language of Brazilian digital television system. In that application a secondary video is loaded through a broadband Internet access simultaneously with the main video being received through broadcasting.

The use of a broadband Internet as the main interactive channel is relevant because the delay of loading the application and its corresponding data is

Cristiane Zakimi Correia Pinto
Wagner Luiz Zucchi

critical to synchronize the application and the broadcast programming.

A test platform was created where IP packet loss was introduced in a controlled way affecting the secondary video, as it is expected to occur in a real network. The tests were made with the premise that the main video does not suffer interference, keeping its original quality.

The quality of the videos obtained in each case of packet loss was analyzed according to objective metrics and the obtained data allowed to assess what would be the QoE perceived in each case.

The objective of this research is to verify if it is feasible to use the broadband Internet as a return channel of an interactive application of Brazilian digital television.

2 Transmedia

There has been a lot of talk about multimedia and crossmedia. More recently, with the popularization of the Internet and the appearance of new technologies that facilitated the interaction between people (like social networks), the concept of crossmedia evolved to transmedia, as defined for the first time by Henry Jenkins in [4].

Multimedia is often define as the existence of more than one means of communication of the same content, just a copy of the content in different media (e.g., a soap opera that has been produced and shown on television and then released in DVD and blu-ray) [5].

Crossmedia is the possibility of a content to appear in different means of communication to share some idea, but there is no connection between the parts of the story that are developed in each platform (e.g., a movie and a comic book with the same characters) [5].

Transmedia is the term used when a story unfolds across multiple media platforms with each new text making a distinctive and valuable contribution to the whole [4],[5].

The Matrix phenomenon is an example of transmedia: three movies ("The Matrix", produced in 1999, "Matrix Reloaded", produced in 2003, and "Matrix Revolutions", also produced in 2003), a collection of short animation films detailing the Matrix universe ("Animatrix", produced in 2003), a series of comic books, and two games ("Enter the Matrix" video game, and a multiplayer online game set in the universe of "The Matrix"). In The Matrix, several situations that occur on a media platform are part of the story on another media platform.

Always when there is a transmedia narrative, each platform gives its contribution to the narrative, encouraging the engagement of the audience. The audience makes the connection between the platforms, contributing to the development of the story.

In Fig. 1, a visual comparison between multimedia, crossmedia, and transmedia connections is shown (adapted from [5]).
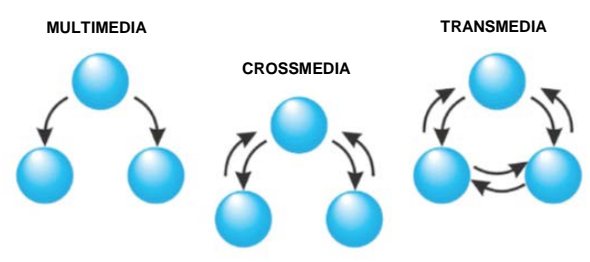

Fig. 1 Multimedia, crossmedia, and transmedia

3 Digital Terrestrial Television

Digital terrestrial television is the digital television broadcast over the air, free of charge to all population.

The differences between the analogue terrestrial television and digital terrestrial television are quite significant. In the case of Brazil, the digital television gives the possibility to receive high definition video and high quality audio, to receive TV signals on portable devices and mobile television receivers, to have more than one television program per channel and to have interactivity combined with the television programming.

A digital television programming is composed of a main audio and a main video, and may also contain additional data. This ability to transmit data, including applications relating the various media objects defined in these data, makes possible the offering of interactive services. The audio and video are delivered to the digital encoders that generate a main video stream and a main audio stream, both compressed. These streams along with the data stream are multiplexed into one signal called the Transport Stream (TS). Then, the TS is modulated into a transmission channel, in frequency, and transmitted. At the reception, the signal is received and demodulated, being delivered to the demultiplexer which separates the main streams of audio and video from the data streams. The data streams are then delivered for processing. If used,

the interactivity channel can provide a channel for user data to interact with the process [6].

3.1 Digital Television Systems

The digital television systems consist of a set of standards that make it possible processing, transmitting and receiving digital signals. Currently, in the world, there are the following digital TV systems: the American system ATSC (Advanced Television Systems Committee), the European system DVB-T (Digital Video Broadcasting – Terrestrial), the Japanese system ISDB-T (Integrated Services Digital Broadcasting – Terrestrial), the Chinese system DTMB (Digital Terrestrial Multimedia Broadcast) and the system adopted in Brazil ISDB-TB (Integrated Services Digital Broadcasting – Terrestrial Brazil), which is the Japanese system with modifications.

The choice of the ISDB-T system to be the base of the Brazilian system was made after a period of tests between the three existing systems at the time (ATSC, DVB-T and ISDB-T), in 2006 [7].

The Fig. 2 shows an overview of the architecture of the Brazilian digital terrestrial television system (ISDB-TB).
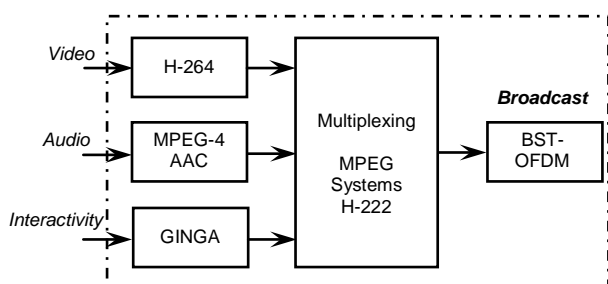

Fig. 2 Overview of the ISDB-TB

3.2 Data Carousel

The data carousel aims to allow transmission of data such as applications, which can be processed by the receiver.

The digital terrestrial television broadcast is unidirectional (from the TV station to the receiver), and is possible to turn on the television or tune the channel of the TV station when the application is already being transmitted. Because of that, the DSM-CC (Digital Storage Media – Command and Control) protocol was created.

The DSM-CC is part of the MPEG-2 System, and transmits the data cyclically, like a carousel (see Fig. 3). Thus, the receiver can receive data as soon as the channel is tuned. Only when all the

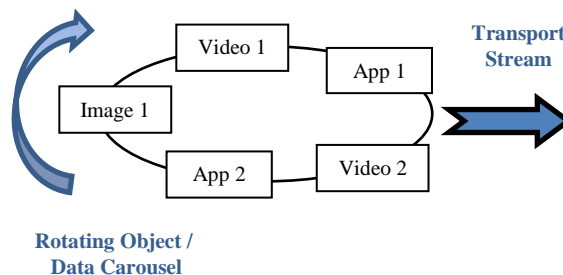information is received by the receiver, the application becomes available to the viewer.

Fig. 3 Data carousel

3.3 Interactivity

Nowadays interactivity is very common. The popularization of the Internet has caused the people to get used to interact not only with the computer through websites but also with other people through social networks. The broadcasters need to keep the development of new digital products and services to maintain the viewers tuned in its channel and to keep their attention. [8].

The interactivity through digital television is different from the experienced when using the computer and accessing the Internet, for example.

A computer has large processing capacity and data storage and it is often operated by people who have a certain technical knowledge to install programs. On the other hand, the digital television receiver has limited resources, little memory, small or non-existent storage capacity, and the user cannot install components in the equipment. Therefore, the applications developed for interactive digital TV must consider the limitations of the system and be appropriate to the context.

It is important to clarify that the interactivity offered by digital terrestrial TV differs from the interactivity found in so-called "smart TVs" or "connected TVs". The "smart TVs" applications that provide access to interactive content (e.g. videos, tools for social networking, communication) are part of proprietary solutions that are the result of a partnership between content providers and manufacturers of television sets [9].

The *middleware* is the responsible for the interactivity in the digital terrestrial television, since it allows the applications and services to be always accessed in the same way, regardless of the platform they run on (e.g. [10]). In another words, the middleware is an intermediate layer between the common applications of digital TV system (such as electronic program guides and services offered by the TV station) and the applications that are defined

by the receiver manufacturer, which is non-standard (such as the operating system used by the receiver).

Ginga is the middleware adopted in the ISDB-TB [11].

4 Quality of Experience

The Quality of Experience (QoE) can be defined as a dependent parameter from the point of view of the user, which combines personal sensations and perceptions to judge whether the QoE is satisfactory or not [12], [13].

In this paper it was considered that the QoE is based on evaluation of video quality received when interacting with the application of digital TV.

5 Video Quality Assessment Methods

The video quality assessment can be made through subjective metric or through objective metric.

When a video is evaluated by subjective metrics, the judgment is based on human perception.

Subjective tests for determining quality are very time consuming, so there are studies to develop a software for subjective quality estimation of all tested sequences using the recommendation ITU-T P.910 (subjective video quality assessment methods for multimedia applications) [14]. The recommendation ITU-T P.910 is suitable for testing different multimedia applications.

The objective metrics use mathematical models to estimate user opinion.

Objective metrics can be classified according to the availability of the original video, which is considered high quality or free of distortion. It is called *FR* (Full Reference) when the not degraded reference video signal is fully available, *RR* (Reduced Reference) when the original video signal is not fully available, and *NR* (No-Reference) when the reference video is not available [15].

The video analysis in this paper is the type *FR* because the original video and the degraded video are used.

The following objective metrics were used to evaluate the videos obtained by the digital interactive TV application for this paper.

5.1 PSNR (Peak Signal to Noise Ratio)

The PSNR defines the relationship between the maximum possible power of a signal and the noise that affects the representation of the signal between the frames of the original video and the degraded video [15]:

$$PSNR = 10\log_{10} X$$

$$X = \frac{L^2}{\frac{1}{MNT}\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{t=1}^{T}\left[I(m,n,t)-\hat{I}(m,n,t)\right]^2}. \quad (1)$$

In (1), L is the dynamic range of pixel values, M and N are the width and the height of the original video and the degraded video sequences, respectively. T is the number of frames containing the sequences, $I(m,n,t)$ and $\hat{I}(m,n,t)$ represent the pixel in position *(m,n)* from t frame for the original sequence and the degraded video sequence, respectively.

The PSNR range goes from 0 to 100 (the best possible PSNR value is 100).

5.2 SSIM (Structural SIMilarity)

This metric is based on the human visual model and it assumes that the images are highly structured, and these dependencies contain very important information relating to the structure of the object.

The SSIM algorithm estimates the similarity between the original video and the degraded video, comparing the brightness *l(x,y)*, contrast *c(x,y)* and structure *s(x,y)* of the original video *x* and the degraded video *y*.

The SSIM is represented by the following equation [15]:

$$SSIM(x,y) = \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)} \quad (2)$$

$$C_1 = \left(K_1 L\right)^2 \quad (3)$$

$$C_2 = \left(K_2 L\right)^2 \quad (4)$$

$$L = 2^{\#bits/pixel} - 1 \quad (5)$$

$$K_1 \ll 1 \; ; \; K_1 = 0.001 \text{ (usual value)} \quad (6)$$

$$K_2 \ll 1 \; ; \; K_2 = 0.03 \text{ (usual value)} \quad (7)$$

In (2), μ_x and μ_y is the average of x and y, σ_x^2 and σ_y^2 are the variance of x and y, respectively; σ_{xy} is the covariance of x and y; C_1 e C_2 are constants which stabilize the division with the peak denominator, avoiding the onset of instabilities when the denominator approaches zero. L is the dynamic range of the video's pixel values.

The SSIM range goes from 0 to 1, with 1 being the best possible value.

Cristiane Zakimi Correia Pinto
Wagner Luiz Zucchi

5.3 VQM (Video Quality Metric)

This method uses the Discrete Cosine Transform (DCT), and uses the original video and the degraded video as inputs to evaluate the video quality [16].

This metric is based on simplified human spatial-temporal contrast sensitivity model [17]. The distortion of a compressed video is calculated in four steps [18]:

I.   For every frame the model performs DCT for 8 x 8 pixels blocks $b_i(x,y,t)$ of the original video frame $p(x,y,t)$, and for blocks $b_i'(x,y,t)$ of the degraded video frame $p'(x,y,t)$.

$$DCTb_i(u,v,t) = DCT(b_i(x, y, t))$$
$$DCTb_1'(u,v,t) = DCT(b_i'(x.y.t)) \quad (8)$$

II. The DTC coefficients are converted by the model to local contrast values $LC_i(u,v,t)$ using DC component of each block. On the same way model obtains $LC_i'(u,v,t)$ of the degraded video.

$$LC_i(u,v,t) = \frac{DCTb_i(u,v,t).\left(\dfrac{DC_i}{1024}\right)^{0.65}}{DC_i} \quad (9)$$
$$DC_i = DCTb_i(0,0,t)$$

III. The model converts $LC_i(u,v,t)$ and $LC_i'(u,v,t)$ to just noticeable difference values, $JND_i(u,v,t)$ and $JND_i'(u,v,t)$, by using static and dynamic spatial contrast sensitivity function (CSF).

IV. The coefficients of original and degraded sequences are subtracted to produce a difference values $Diff_i(t)$. The VMQ score is obtained by:

$$Dist_{Mean} = 1000.mean(mean(Diff_i(t)))$$
$$Dist_{Max} = 1000.\max(\max(Diff_i(t)))$$
$$VQM = Dist_{Mean} + 0.005.Dist_{Max} \quad (10)$$

The higher is the quality of the video, the lower is the value of VQM, with 0 being the best possible value.

6 Test Platform

The aim of the test platform is to provide interactivity between the digital TV receiver and the TV station using the broadband access to Internet. The following premises are used:

a.   Television set has a digital TV receiver with interactivity, according to Brazilian standard [19].
b.   Television set uses a broadband Internet access as the interactivity channel.
c.   The TV station has access to broadband Internet to provide data to the interactive application.
d.   There is no interference in the main video reception.
e.   The user QoE is related only to the quality of the received video over the auxiliary channel.

In this scenario, the TV station broadcast the interactive application along with the audio and video signals via data carousel [20]. The television set receives the data, displays on-screen programming and gives the possibility of interaction for the user (viewer). The user can interact and can use the interactive channel (also called return channel), which is the broadband Internet (see Fig. 4).

To minimize the delay caused by loading data to compose the interactive application, the data will be accessed via Internet through the interactive channel rather than being transmitted over the air. The Internet content would be required by the running application.

The platforms "television" and "Internet" become interconnected with the described environment, creating the necessary structure for the user to go through the various media and to have a full transmedia experience.
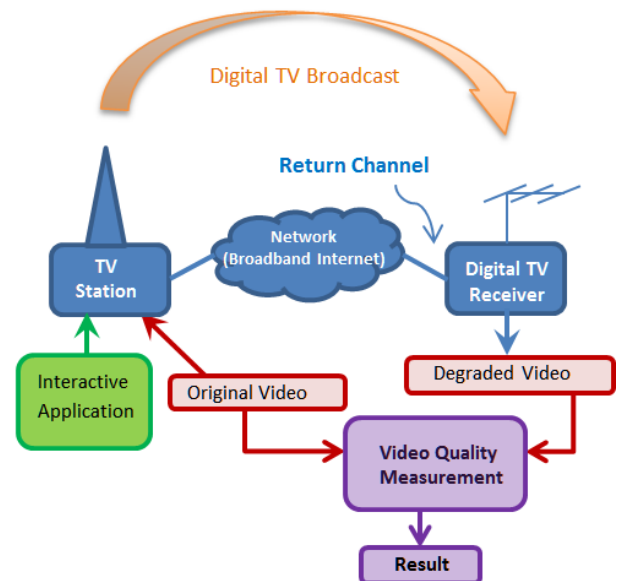

Fig. 4 Scenario to be analyzed by the test platform

The first step to implement the test platform was the preparation of a transmedia interactive application that utilizes the media platforms considered in this work. The application was written in NCL, which is mandatory for digital TV receivers.

The interactive application was used by GINGA-NCL digital TV set-top-box emulator, available free on the Internet [21] (the Fig. 5 shows the emulator's first screen).

Fig. 5 GINGA-NCL digital TV set-top-box emulator

Two computers were connected via UTP cable, where one of them was representing the TV station and the other was representing the digital television set.

The software WampServer [22] was installed in the TV station computer in order to transform it into a web server with the secondary video to be accessed. It also was installed the SoftPerfect Connection Emulator software [23] to introduce packet loss into the communication channel.

In the other computer that represented the digital television set the following software were installed: digital TV receiver emulator with interactivity GINGA-NCL [21], video recording software CamStudio [24], video editing software AVI Trimmer [25] and MSU Video Quality Measurement Tool [26].

The MSU Video Quality Measurement Tool determines the video quality with objective metrics. The tool is free to be used in standard-definition videos but it is necessary to purchase the PRO version license for high-resolution video.

The main video was loaded into the computer that represented the digital television set because it was considered that would be no interference in its reception.

With this test platform setting (see Fig. 6), it was possible to use the application of interactive digital TV, simulate various conditions of packet loss, and evaluate the quality of the received video.
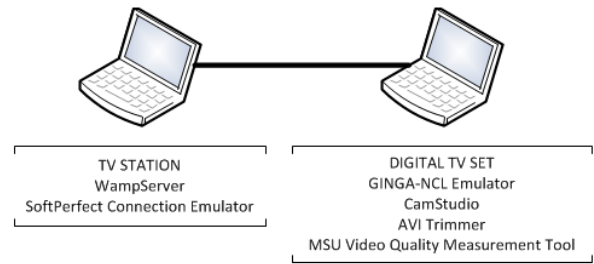

Fig. 6 Test platform

7 Results

The results that have been obtained are shown in Table I. The corresponding graphics of each metric are shown in the Fig. 7, Fig. 8, and Fig. 9.

Table I Packet loss effect on the quality of the received video

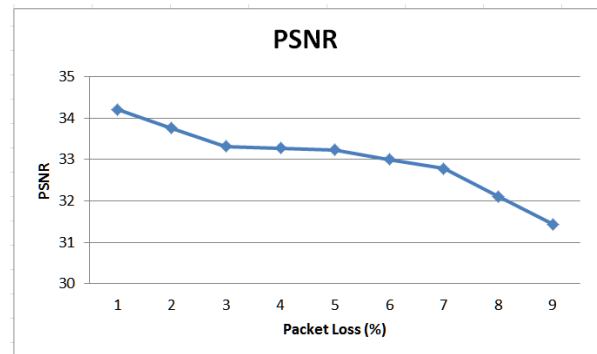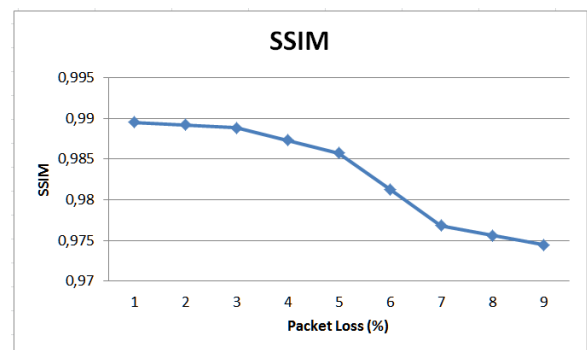| Packet Loss (%)	PSNR (dB)	SSIM	VQM
1	34,20398	0,98952	3,31409
2	33,75884	0,98918	3,75706
3	33,31370	0,98885	4,20004
4	33,27305	0,98731	4,99218
5	33,23240	0,98577	5,78433
6	33,00510	0,98130	6,00716
7	32,77781	0,97684	6,22999
8	32,10996	0,97564	6,42314
9	31,44211	0,97445	6,61630

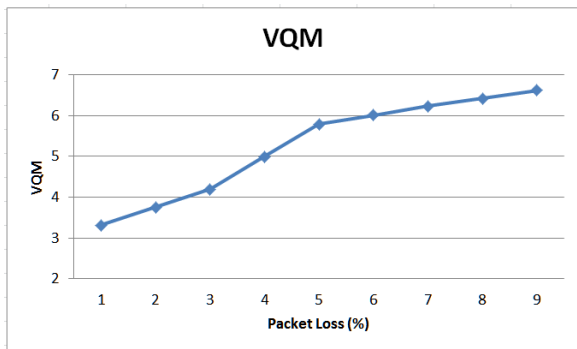Fig. 7 PSNR Graphic

Fig. 8 SSIM Graphic

Fig. 9 VQM Graphic

8 Conclusions

The obtained results show that the quality of the resulting degraded video for each situation of packet loss is as expected, i.e., the higher the packet loss, the worse the quality of the video displayed to the user and therefore worse the QoE.

It is remarkable that the video quality in this study with the interactive application of digital TV resembles the case of receiving the video via streaming that was studied in [15].

Analyzing the PSNR metric, the video quality loss was found directly proportional to the packet loss until to reach to 3% packet loss. Between 3% and 5% packet loss, the video quality remained almost the same, returning to worsen over 5% packet loss.

When analyzing the graphic of SSIM values, it appears that up to 3% packet loss video quality hardly changed. Between 3% and 5% packet loss, the video quality has deteriorated slowly, and to greater than 5% packet loss, the video quality fell rapidly.

To the VQM metric, the loss of video quality has increased almost linearly with increasing packet loss.

Considering all the three metrics, as the graphics show, it can be considered that with less than 3% of packet loss the secondary video still shows a QoE compatible with the kind of application be considered.

So, it is possible to receive the secondary video and use the broadband Internet as a return channel of an interactive application of Brazilian digital television, as far as the losses can be controlled inside these limitations.

References:

[1] Secretaria-Executiva do Comitê Gestor do Programa de Inclusão Digital. (2010). Programa Nacional de Banda Larga [Online]. Available:

http://www.planalto.gov.br/brasilconectado

[2] G1 Economia. (2012, Sep. 20). Número de casas com TV supera o das que têm geladeira [Online]. Available: http://g1.globo.com/economia/noticia/2012/09/numero-de-casas-com-tv-supera-o-das-que-tem-geladeira.html

[3] *Digital Terrestrial Television - Interactive Channel. Part 1: Protocols, Physical Interfaces and Software Interfaces*, ABNT NBR 15607-1:2011 EN, 2011.

[4] H. Jenkins, *Cultura da Convergência*, Aleph, 2009.

[5] R. D. Arnault *et al.*, Era Transmídia. *Revista GEMInIS*, ano 2 n.2, Jul./Dec. 2011, pp. 259–275.

[6] L. F. G. Soares, S. D. J. Barbosa, *Programando em NCL 3.0: Desenvolvimento de Aplicações para Middleware Ginga, TV Digital e Web*, Elsevier, 2009.

[7] Brazil, Decreto n° 5.820 de 29 de junho de 2006, *Diário Oficial da União*, 2006, Jun. 30, pp. 51.

[8] B. Zuell, D. Boras, N. M. Preradovic, The Digital Change of Television in a Social Media Environment: The Case of the BBC, *Recent Advances in Information Science*, WSEAS Press, 2012.

[9] A. C. B. Angeluci, R. D. Lopes, M. K. Zuffo. Estudo Comparativo entre TV Digital Aberta e TV Conectada no Brasil, *Proc. XXXIV Congresso Brasileiro de Ciências da Comunicação – INTERCOM*, 2011, pp. 1–16.

[10] C. S. Oliveira, E. B. Dos Santos Jr, H. C. De Oliveira, Synchronizing Brazilian Sign Language (LIBRAS) on Videos for Open Digital TV with Educational Purposes, *Applied Mathematics and Informatics*, WSEAS Press, 2010.

[11] *Digital Terrestrial Television – Data Coding and Transmission Specification Broadcasting. Part 1: Data coding specification*, ABNT NBR 15606-1:2013 EN, 2013.

[12] R. Jain, Quality of Experience, *MultiMedia, IEEE*, vol. 11, no. 1, Jan.-Mar. 2004, pp. 95–96.

[13] Y. Lu, B. Fallica, F. A. Kuipers, R. E. Kooij, P. Van Mieghem, Assessing the Quality of Experience of SopCast, *Int. J. Internet Protocol Technology*, vol. 4, no. 1, 2009, pp. 11–23.

[14] Z. Miličević, Z. Bojković, K. R. Rao, An Approach to Interactive Multimedia Systems Through Subjective Video Quality

Assessment in H.264/AVC Standard, *WSEAS Transactions on Systems*, Vol. 11, Is. 8, 2012, pp. 305-314.

[15] D. C. Begazo, D. Z. Rodríguez, M. A. Ramírez, Avaliação de Qualidade de Vídeo Sobre uma Rede IP Usando Métricas Objetivas. *Revista Iberoamericana de Sistemas, Cibernética e Informática*, 2011 [Online]. Available: http://www.iiisci.org/journal/risci/Abstract.asp?var=&id=HCA940MZ

[16] D. Z. Rodríguez, G. Bressan., Video Quality Assessments on Digital TV and Video Streaming Services Using Objective Metrics. *Latin America Transactions, IEEE (Revista IEEE America Latina)*, vol. 10, no. 1, Jan. 2012, pp. 1184–1189.

[17] F. Xiao, *DCT-based Video Quality Evaluation*, 2000 [Online]. Available: http://www.compression.ru/video/quality_measure/vqm.pdf

[18] M. Vranjes, S. Rimac-Drlje, D. Zagar, Objective Video Quality Metrics, *49th International Symposium ELMAR-2007 Focus on Mobile Multimedia*, 2007, pp.45–49. DOI: 10.1109/ELMAR.2007.4418797

[19] *Televisão Digital Terrestre – Receptores*, ABNT NBR 15604:2007 Versão corrigida: 2008, 2008.

[20] *Digital Terrestrial Television – Data Coding and Transmission Specification Broadcasting. Part 3: Data transmission specification*, ABNT NBR 15606-3:2012 EN, 2012.

[21] GINGA-NCL. [Online]. Available: http://www.gingancl.org.br/en

[22] Wampserver, a Windows web development environment. [Online]. Available: http://www.wampserver.com/en/

[23] SoftPerfect Connection Emulator. [Online]. Available: http://www.softperfect.com/products/connectionemulator/

[24] CamStudio. [Online]. Available: http://www.camstudio.org

[25] SOLVEIGMM. [Online]. Available: http://www.solveigmm.com/pt/products/avi-trimmer-mkv/

[26] VIDEO GROUP. [Online]. Available: http://compression.ru/video/quality_measure/video_measurement_tool_en.html