Video Quality Indicators for ISDB-Tb Free to Air Digital Television

Rafael Sotelo
IEEE Senior Member

Jose Joskowicz
IEEE Senior Member

***Abstract*— The present paper describes the architecture of an ongoing project (called "VQI – Video Quality Indicators") that aims to generate indicators of the quality of video received by the viewers of Free to Air Digital Television. Particularly, it is oriented to ISDB-Tb system so the video codec used is H.264 with both HD and SD resolutions. The final result of the project is to obtain an objective indicator of video quality and direct subjective measures, based on the average opinion of the audience. For the subjective measurements, the concept of *crowdsourcing* for TV is introduced. They are performed through an Interactive TV application using Ginga middleware. The paper includes a brief discussion on the treatment of packet loss for DTV, and a description of a new application for smartphones developed to input the ratings in subjective test sessions which simplifies processing of results and minimizes errors.**

***Index Terms*— Quality of Experience, ISDB-Tb, Ginga, Quality Assessment**

I. Introduction

Digital Television (DTV) has been deployed virtually in all of the developed countries and is being deployed in most of the rest of the world. While many countries have not yet started digital TV transmissions, ITU (through the Geneva 2006 Agreement) has set June 17th 2015 as the date for the end of the digital to analog transition for a vast region, including Africa and Europe. At the end of the transition period, countries may continue to operate analogue broadcasting stations provided that these stations do not cause unacceptable interference to neighbor countries and that they do not claim protection.

Broadcasters always aim to deliver the best possible video quality to their audience. The promise of better video quality has been one of the driving forces for the advent of DTV.

While the above statement is clearly true, some processes involved in DTV, such as digital video encoding and transmission systems, introduce degradations that may result in unsatisfactory perceived quality.

Since video quality depends on many aspects related to the encoding process, to the transmission stage or even to the content itself (see Table I), it is desirable to establish clear and reliable indicators that allow measuring and/or estimating the video quality that is perceived by the audience of a certain broadcaster.

This work was supported by ANII-MIEM/Dinatel FST_1_2012_1_8147.
Rafael Sotelo is with Facultad de Ingeniería, Universidad de Montevideo, Montevideo, Uruguay (Phone: +598 27067630; email: rsotelo@um.edu.uy)
Jose Joskowicz is with Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay (Phone: +598 27110974; email: josej@fing.edu.uy)

TABLE I
FACTORS AFFECTING VIDEO QUALITY

Encoding Process	Transmission Stage	Video Content
Video Codec	Percentage of Packet Loss	Temporal Activity
Resolution	Loss Patterns (i.e. burstinnes)	Spatial Activity
Bit Rate	Delay	
Frame Rate	Jitter	

The effect of the encoding process in the video quality perceived has been previously analyzed by different authors. A parametric model combining the aspects that affect video quality during the encoding process has been proposed in [1]. In this previous work, video resolution was varied from small sizes up to SD (Standard Definition), but HD (High Definition) was not evaluated.

The transmission stage parameters incidence in video quality has often been analyzed from a telecom provider point of view [2]. Packet loss, jitter and delay are often related to IP or similar data networks, were the encoded video is encapsulated in IP packets. This approach differs from a broadcast scenario in some unattended ways, were small 188 bytes packets are broadcasted individually, and different methods of error corrections are implemented, such as Reed Solomon codes. In that sense, specific aspects of packet loss in broadcasting must be considered and analyzed.

The effect of video content in video quality has often been studied associated with the encoding parameters, and some objective video quality models include different aspects of video content in the models. Nevertheless, less work has been done regarding the effect of video content combined with the degradations introduced in the transmission stage, particularly considering DTV.

A review and comparison of more than 10 different video quality estimation models have been presented in [3].

In parallel, Interactive Television (ITV) is gaining more and more attention from researchers in Latin America [4] [5]. ITV is expected to reduce the digital divide. To help the deployment of ITV it is necessary to develop applications that make TV viewers appreciate this new technology as useful.

The present paper describes the architecture and results of an ongoing project that aims to generate both objective and subjective indicators of the quality of video received by viewers of Free to Air DTV. It is oriented to ISDB-Tb system [6] and the video codec used is H.264 for HD and SD. The project is called VQI: Video Quality Indicators [7]

The rest of the paper is structured as follows: Section II introduces DTV quality assessment studying previous work, particularly in relation to the treatment of degradations introduced by noise and our own considerations and approach; Section III presents the architecture and basic modules of the system; The module for objective estimation of video quality is outlined in Section IV, together with the description of the model used and the subjective test performed for calibration, were a specific application was developed to collect ratings; Section V describes the process to obtain a novel subjective indicator, including the Ginga application developed to allow the end TV viewer to rate the video quality of the program he is watching, using the new *crowdsourcing* concept. Finally Section VI includes the conclusions of the paper and the future work of the project.

II. DTV Quality Assessment

Video quality, as perceived by the viewers, mainly depends on the parameters used during the encoding process, and the degradations introduced in the transmission stage. In both cases, the final quality also depends on the content itself, for example if it is a quiet scene or it has a great amount of motion. In order to design and develop a system to measure and estimate the video quality in ISDB-Tb Free to Air DTV, it is necessary to review the specific aspects of this broadcasting system that can affect the video quality.

ISDB-Tb system, which is described in [6], uses the codec H.264 and allows both HD (1920×1080) and SD (720×480 or 720×576) resolutions.

ISDB-Tb allows different modes of transmission. This enables the allocation of different number of programs, with different bit rates in the same frequency band. The system divides the 6 MHz RF channel in 13 equal segments. The central one (number 0) is often reserved for a rugged modulation scheme, oriented to provide content to mobile telephones. The other twelve segments are available for the main HD and SD programs. They can be configured in one or two sets, each one with its own transmission parameters, such as modulation (QPSK, 16QAM or 64QAM), Forward Error Correction (FEC, ranging from 1/2 to 7/8) and guard interval (1/32 to 1/4). Table II shows the available data rate in kbps for each segment depending on its configuration. As can be seen, the selection of different transmitting modes would imply a different total bit rate capacity for the whole RF channel. For example, if the twelve segments are set in a quite conservative configuration as 16QAM, 2/3 and 1/4, there will be a total of 12×748.95 kbps, approximately 9 Mbps available. While if they all are set in 64QAM, 3/4 and 1/16, there will be a total of 12×1321.68 kbps, approximately 16 Mbps available. In each of these cases the broadcaster will have available a quite different amount of bit rate. In case he decides to transmit only one HD program in the available bandwidths, the signal will have 9 or 16 Mbps in each case. If he decides to transmit two HD programs, each one will have 4.5 or 8 Mbps, respectively. Similar considerations apply if he decides to transmit any combination of HD and SD signals.

As can be seen, the bit rate can be configured by the broadcaster with some limitations. As is known, this will determine the video quality of the program transmitted.

Besides, the selection of these transmission parameters will affect the way the signal is propagated and received, particularly, in relation to the robustness to noise.

TABLE II
Data rate of a single segment [**ABNT NBR 15601**].

Carrier modulation	Convolutional code	Number of transmitted TPS	Data rate [a]			
			Guard Interval 1/4	Guard interval 1/8	Guard interval 1/16	Guard interval 1/32
DQPSK QPSK	1/2	12/24/48	280.85	312.06	330.42	340.43
	2/3	16/32/64	374.47	416.08	440.56	453.91
	3/4	18/36/72	421.28	468.09	495.63	510.65
	5/6	20/40/80	468.09	520.10	550.70	567.39
	7/8	21/42/84	491.50	546.11	578.23	595.76
16QAM	1/2	24/48/96	561.71	624.13	660.84	680.87
	2/3	32/64/128	748.95	832.17	881.12	907.82
	3/4	36/72/144	842.57	936.19	991.26	1 021.30
	5/6	40/80/160	936.19	1 040.21	1 101.40	1 134.78
	7/8	42/84/168	983.00	1 092.22	1 156.47	1 191.52
64QAM	1/2	36/72/144	842.57	936.19	991.26	1 021.30
	2/3	48/96/192	1 123.43	1 248.26	1 321.68	1 361.74
	3/4	54/108/216	1 263.86	1 404.29	1 486.90	1 531.95
	5/6	60/120/240	1 404.29	1 560.32	1 652.11	1 702.17
	7/8	63/126/252	1 474.50	1 638.34	1 734.71	1 787.28

[a] Represents the data rate (bits) per segment for transmission parameters data rate (bits) = Transmitted TSP × 188 (bytes/TSP) × 8 (bits/byte) × 1/frame length

In DTV the Bit Error Rate (BER) is often used as a parameter related to the capability of the receiver to reconstruct the transmitted signal. For example, DVB-T [8] defines Quasi Error Free (QEF) reception as less than one uncorrected error event per hour, corresponding to BER = 10^{-11} at the input of the MPEG-2 demultiplexer, or a BER = 2×10^{-4} after Viterbi.

The common approach is that if Carrier to Noise relation (C/N) is below a certain value, there will be a cliff effect (also called brick effect, or 'brick wall' effect or 'fall off the cliff') that will cause an immediate degradation of the signal [9] [10] [11]. It is frequent to see graphs like Figure 1, that shows how much abrupt is the fall of quality in DTV compared to Analogue TV as signal intensity decreases.

Although the quality of the RF transmission link is usually characterized by the BER, this approach is not the best to study the problem on how transmission link noise affects QoE from the user's perspective in DTV. This fall takes place in 1 to 3 dB [12] [13]. This range may include many households in a normal coverage area of a DTV station. C/N varies with environment conditions and varies with time. That increases the number of receivers that can be on the edge of this "cliff effect" in different periods of time. Besides, noise intensity may have different structures during time depending on its origin (i.e. homogeneous or bursts). It is then necessary to model the effect of this degradation on C/N and its different natures in the final QoE perceived by the viewers.

Another concept called the Correct Reception Rate (CRR) has been used in [14] and [15], but it also is used as a threshold (good / bad reception), and does not enable to make any analysis of the perceived video quality as a function of this parameter.

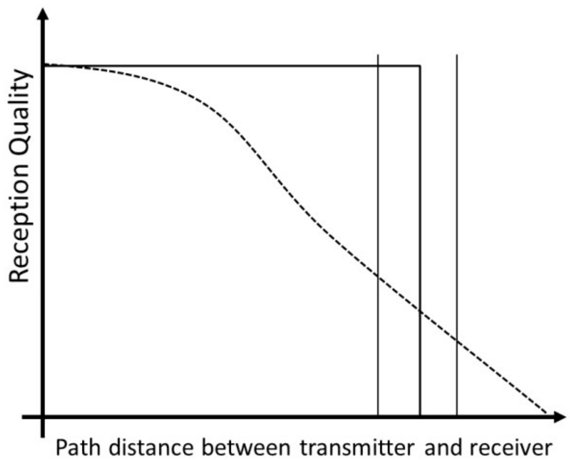

Fig. 1. Cliff Effect in Reception Quality for Digital TV transmission (solid line) vs. Analogue TV (dotted line).

Other works have partially analyzed the effect of the signal fading in the video quality [16]. However this parameter is very difficult to measure at the receiver, and cannot be included in a video quality estimation model.

The last stage in a DTV receiver is the Reed-Solomon (RS) decoder. 188 bytes Transport Stream Packets (TSP) are used, and error corrections are performed at the receiver. In the case of ISDB-Tb or DVB-T the Reed Solomon Code is (204, 188) and can correct errors in up to 8 bytes. If there are severe error-bursts, the RS decoding algorithm may be overloaded, and be unable to correct the packet. In this case the transport_error_indicator bit in the TSP header shall be set [17]. The decoder can decide what to do with the missing information. Three different receivers analyzed by the authors, when recording a Transport Stream file, simply drop the TSPs marked as having had a transport error. That is, checking the continuity counter of the header of the TSP of the TS file recorded, some missing TS may be found.

Thus, for DTV, it is necessary to focus on degradations experienced by packet loss (individual TS packets) due to digital broadcasting transmission. There is much work done regarding the study of IP packet loss when transmitting video (for example, for streaming or IPTV). However, it is different of the case of DTV since the IP packet includes 7 TSPs. In IP networks, TS packets are lost in blocks of 7, while in DTV TS packets may be lost individually.

According to the previous considerations, the objective model developed for this project has to take into account the characteristics of DTV, including SD and HD resolutions, bit rates according to the ISDB-Tb possible transmission modes, and TSP packet losses with homogeneous and burst distributions.

III. VQI ARCHITECTURE

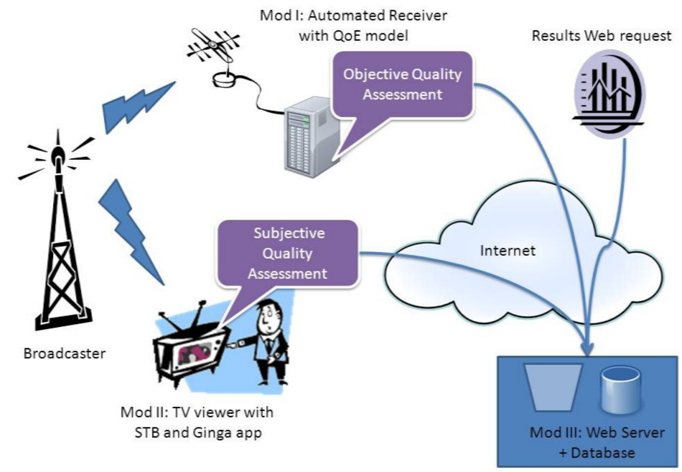

Fig. 2. Architecture of VQI System

The main objective of this project is to generate indicators of the quality of video received by the viewers of Free to Air DTV. Figure 2 outlines the architecture of the designed system, consisting of three different modules. There are two modules that directly receive the normal programming of the broadcasters.

Module I performs an objective video quality assessment. This module is an automated receiver made up of a PC with a TV tuner card that records different TV channels according to a programmed schedule. In this PC also resides the algorithm that implements the QoE estimation model, further described in Section IV. This module, based on the channel recordings, computes the predicted Mean Opinion Score (MOSp) by the system for each channel for each period of time. It sends this information (channel, date, hour and MOSp) to Module III, through LAN or Internet, which stores and publishes the results.

On the other hand, Module II enables a subjective quality assessment through direct qualifications of TV viewers. There is an interactive TV application developed under Ginga middleware [18] that has to be broadcasted to be loaded to users' Set Top Box (STB). In case that the TV viewer is tuning a channel in which the application is being broadcasted, he will be able to launch it with a particular button of the STB's remote control. By doing so, he may qualify the video quality of the program being watched. The corresponding information (user, channel, date, hour and qualification) is sent to Module III, through Internet, which stores and process it. The application is described in Section V with special focus on the role it plays in the system. It can easily be imagined that this module is replicated in each household of the audience that want to adhere to the system.

Module III consists on a Web Server and a Data Base. It compiles the information received from the other modules and publishes both the objective and subjective indicators. The first one is directly received from Module I while the second is processed based on the information received from the (possibly numerous) TV viewers that qualify each program.

These indicators can be published, and made available for the operators, the regulators and even the general public, enabling comparison of quality between the various signals, and thus helping to ensure that broadcasters deliver the best possible quality.

The feedback collected from TV viewers can trigger alarms. For example, if many people from a certain zone in the coverage area that usually qualify acceptable the quality of a certain TV channel, begin to qualify as unacceptable, may be a sign that there is a problem either on the transmitter or in the antenna system.

IV. Objective Quality Assessment

As stated in the previous Section, Module I of the system was designed to perform an automatic objective quality assessment of the video signal. Using a pre-defined schedule, the system performs short video recordings (10 – 20 seconds). These recordings are saved as Transport Stream files (".ts"), and are analyzed using a no-reference video quality model. The system was designed in modular blocks, so any no-reference video quality model can be used, without making significant changes to the whole system. In this first implementation, the video quality estimation is based on the parametric model presented in [1] and [19]. This model considers the degradation introduced in the encoding process, and takes into account the bit rate, the frame rate, the video resolution, and some aspects of the video content, based on the average Sum of Absolute Differences (SAD) and the average of the amplitude of the Motion Vectors (MV).

The average bit rate is obtained from the recorded file. For this implementation, the frame rate is constant, at 50 fps, due the usual configuration of DTV stations in Uruguay. Video resolution can be SD or HD. In order to obtain the average SAD an MV, the original Transport Stream file is converted to an AVI file, and the AVI file is processed.

To include the effects of the degradation introduced in the transmission process a novel approach is under development in the present project. As stated in Section II, noise on DTV affects the service on individual TS packets coming out from the RS decoder. They are either omitted or marked as erroneous. In our approach the overall percentage of packet loss and the burstiness are measured directly from the Transport Stream file. Each TS packet has a header, with a sequential number. The file is inspected to detect missing packets. If the time between missing packets are shorter than a predefined threshold, the losses are considered of the same burst.

In order to calibrate the objective model, subjective tests were performed. Five different video clips were used: "Fox & Bird", "Football", "Concert", "Voile" and "Golf", obtained from [20]. This video clips spans over a wide range of different spatial and temporal activity. The video clips were coded in H.264/AVC, High Profile, Level 4.1 for HD and Main Profile, Level 3.1 for SD, with no more than two consecutive B frames, and key interval of 33 frames. One hundred different degraded video clips were generated in HD, and one hundred in SD, varying the bit rate, the percentage of packet loss with uniform distribution and the number and duration of bursts with burstiness distribution.

The subjective tests were performed according to the general guidelines of Recommendation ITU-R BT.500-13 [21], using the 5 point Absolute Category Rating with Hidden Reference (ACR-HR) scale, as defined in Recommendation ITU-T P.910 [22]. A 42" Led TV was used. The test room allows up to four simultaneous people watching the video clips and voting. A special voting system was developed, allowing the evaluators to use a smart phone application, synchronized with the clip sequences, to enter each qualification. The voted scores are automatically stored in a database, associated with the test session, the specific video clip and the user. The system is depicted in Figure 3.

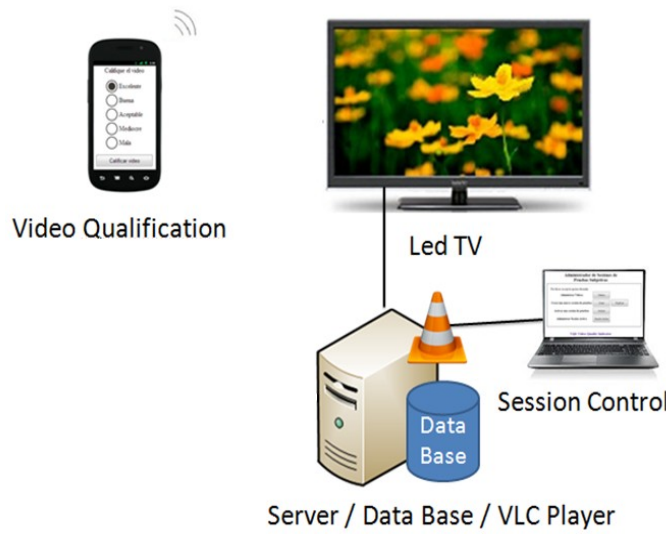

Fig. 3. Automatic qualification system for subjective tests

The system automates the subjective tests of video quality, offering the possibility to create and configure test sessions with their respective videos, which are played automatically. It also facilitates the tasks of collecting and processing the qualifications of evaluators (traditionally taken on paper), which are entered into the system by the evaluators using mobile devices in sync with the playback of the video clips presented during testing. According to our experience the traditional way to collect evaluation, based on filling paper forms caused annoyance to evaluators and is a potential source of errors, since the subject can skip the qualification of a clip, thus causing a shift in all his results.

V. Subjective Quality Assessment

Subjective quality indicators are very accurate but also expensive and hard to obtain.

There is a new and efficient tool for performing subjective tests that is being applied in the Internet. Crowdsourcing QoE assessment usually means to outsource subjective studies to a crowd in the Internet [23]. For example, it is being used to study YouTube QoE [24]. We have expanded this concept to DTV, to allow a TV station audience to perform subjective quality assessments.

According to the system described in Section III, we are taking advantage of the new capability of interactivity from DTV to develop a complementary subjective quality indicator. An interactive interface has been developed, so that the TV viewers can rate the image quality of each digital television signal. This application, called QualifyTV, runs on home Set-Top Boxes (STBs) or TV sets with return channel available as illustrated in Figure 2. The viewer can rate the quality of each program from each TV channel, provided that the station is broadcasting the application. This rating is stored in a central repository, along with the predicted rating obtained with the objective model.

Fig. 4. QualifyTV snapshot: qualification is enabled.

The broadcaster must include QualifyTV in the carrousel when he wants to enable the audience to qualify the quality of his program. A small graphic appears at the screen, showing the viewer that his input is enabled. In case the viewer wants to vote, the application will be loaded. A snapshot can be seen in Figure 4 where the messages announces that qualifying is enabled and can be started pressing the red button in the remote control. When it is pressed the menu shown in Figure 5 is displayed allowing scrolling with the arrow keys to select in five steps from Excellent to Bad. The vote is sent by pressing the blue button in the remote control. The application sends the data to the central repository through the return channel (cable modem, ADSL, FTTH…).

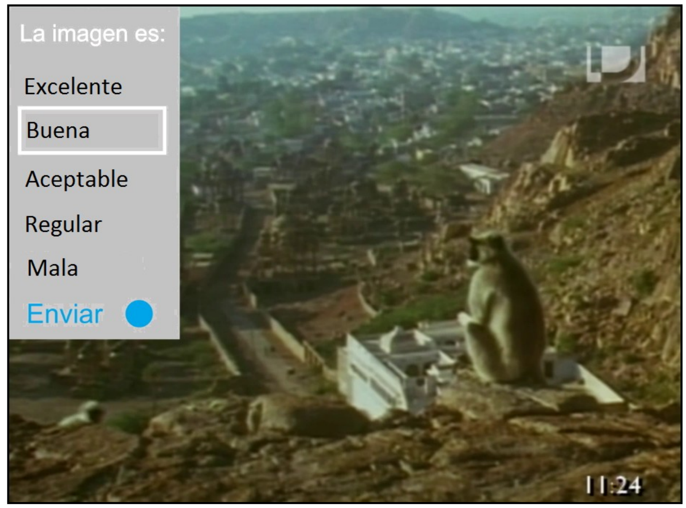

Fig. 5. QualifyTV snapshot: qualification.

The system is designed for ISDB-Tb, so the interactive TV application is developed for middleware Ginga [18]. This application allows the system to have feedback from viewers once deployed, thus resulting in a subjective indicator of quality.

VI. Conclusions and Future Work

This paper outlines VQI (Video Quality Indicators), a system designed to assess QoE in Free to Air Digital Television, consisting on the collections of both an objective and a subjective video quality indicator.

Subjective QoE indicator is achieved through the direct opinion of the audience, using an Interactive TV application, developed on Ginga. This allows bringing the *crowdsourcing* concept to Free to Air DTV in the field of video quality assessment.

Objective QoE indicators are automatically performed, in near real time, making recordings of the different broadcasted signals and using a no-reference video quality estimation model. The model is being developed taking into account the specific characteristics of the encoding and transmission process in ISDB-Tb Digital Television. A special brief discussion has been made on how to deal with noise and packet loss in this broadcasting system. For the objective model calibration, subjective tests were performed, using a smartphone based application specifically developed to improve functionality and confidence on the results of the subjective tests.

The results (objective and subjective measurements) obtained with the developed system can be published, and made available for the operators, the regulators and even the general public, enabling comparison of quality between the various signals, and thus helping to ensure that broadcasters deliver the best possible quality.

REFERENCES

[1] Jose Joskowicz and J. Carlos López Ardao, "Combining the Effects of FrameRate BitRate DisplaySize and Video Content in a Parametric Video Quality Model," *Latin America Networking Conference, LANC 11*, 2011.

[2] Fuzheng Yang and Shuai Wan, "Bitstream-based quality assessment for networked video: a review," *Communications Magazine, IEEE*, vol. 50, no. 11, pp. 203-209, November 2012.

[3] J. Joskowicz, R. Sotelo, and J.C. Lopez Arado, "Comparison of parametric models for video quality estimation: Towards a general model," in *Broadband Multimedia Systems and Broadcasting (BMSB), 2012 IEEE International Symposium on*, Seoul, 2012, pp. 1-7.

[4] G. Baum and L.F.G. Soares, "Ginga Middleware and Digital TV in Latin America," *IT Professional*, vol. 14, no. 4, pp. 59-61, 2012.

[5] V.F. de Lucena, N.S. Viana, O.B. Maia, J.E. Chaves Filho, and W.S.J. da Silva, "Designing an extension API for bridging Ginga iDTV applications and home services," *Consumer Electronics, IEEE Transactions on* , vol. 58, no. 3, pp. 1077-1085, 2012.

[6] Associaçao Brasileira de Normas Tecnicas ABNT, "ABNT NBR 15601 - Digital terrestrial television – Transmission system," 07.04.2008.

[7] VQI: INDICADOR DE CALIDAD DE VIDEO PARA TELEVISIÓN DIGITAL. [Online]. http://www2.um.edu.uy/ingenieria/vqi/

[8] Standards Inst.(ETSI), "300 744 V1. 6.1 (2009-01),“Digital video broadcasting (DVB); framing structure, channel coding and modulation for digital terrestrial television"," Standards Inst.(ETSI), Valbonne, France, ETSI EN 300744, 2009.

[9] F.C.C. De Castro, M.C.F. De Castro, and M Fernandes, "8-VSB channel coding analysis for DTV broadcast," *IEEE Transactions on Consumer Electronics*, vol. 46, no. 3, pp. 539,547, Aug 2000.

[10] T Kratochvil and V Ricny, "Simulation and experimental testing of the DVB-T broadcasting in the SFN networks," *18th International Conference Radioelektronika*, 2008.

[11] W Fischer, *Signals and Communication Technology, 3rd ed. - Digital Video and Audio Broadcasting Technology (Chapter 12).*: Springer-Verlag Berlin Heidelberg 2010, 2010.

[12] S.R. Martin, "RF Performance of DTV Converter Boxes—An Overview of FCC Measurements," *IEEE Transactions on Broadcasting* , vol. 56, no. 4, pp. 441-451, Dec. 2010.

[13] Tiago S Bonfim and Marcelo M. Carvalho. VIDEO QUALITY EVALUATION FOR A DIGITAL TELEVISION BROADCASTING SCENARIO. [Online]. http://enpub.fulton.asu.edu/resp/vpqm/vpqm10/proceedings_vpqm2010/vpqm_p41.pdf

[14] G. Bedicks Jr. et al, "Field Measurements for ISDB-Tb in the VHF Band," *Broadband Multimedia Systems and Broadcasting (BMSB), 2011 IEEE International Symposium*, 2011.

[15] C. Akamine et. al., "Field Trial for Brazilian DTV Using Space Diversity," *Broadband Multimedia Systems and Broadcasting (BMSB), 2010 IEEE International Symposium*, 2010.

[16] D. Z. Rodríguez and G. Bressan, "Video Quality Assessments on Digital TV and Video Streaming Services Using Objective Metrics," *IEEE LATIN AMERICA TRANSACTIONS*, January 2012.

[17] DVB, "ETSI TR 101 290,Measurement guidelines for DVB systems," *Digital Video Broadcasting*.

[18] Middleware Ginga. [Online]. http://www.ginga.org.br/es

[19] J. Joskowicz and J. Lopez Ardao, "A Parametric Model for Perceptual Video Quality Estimation," *Telecommunication Systems; Springer*, vol. 49, January 2012.

[20] The Consumer Digital Video Library. [Online]. http://www.cdvl.org/

[21] ITU, "Recommendation ITU-R BT.500-13, Methodology for the subjective assessment of the quality of television pictures," Jan. 2012.

[22] ITU, "Recommendation ITU-T P.910 Subjective video quality assessment methods," April 2008.

[23] Bruno Gardlo, Michal Ries, and Tobias Hossfeld, "Impact of screening technique on crowdsourcing QoE assessments," *Radioelektronika (RADIOELEKTRONIKA), 22nd IEEE International Conference*, 2012.

[24] Tobias, et al. Hoßfeld, "Quantification of YouTube QoE via crowdsourcing," *IEEE International Symposium on Multimedia (ISM)*, 2011.