Paper

Analysis of Content Quality Evaluation within 3DTV Service Distribution Systems

Grzegorz Wilczewski

Faculty of Electronics and Information Technology, Warsaw University of Technology, Warsaw, Poland

Abstract—In the following paper, evaluation of quality of content distributed within 3DTV service is assessed. The performed analysis reveals a multiplicity of parameters as a compound factor defining the contemporary approach towards quality assessment. Taking into consideration a successful telecommunication service creation, an examination of content chain identifies crucial stages founding the path of modern services of stereoscopic content distribution quality evaluation.

Keywords—3D content perception, quality assessment, service creation, stereoscopic imagery.

1. Introduction

The current state of technology leads directly to a high integration of multiple services dedicated for multiple users. Concerning the fast and constantly growing needs of heterogeneity of customers, more effort is put into newer, more efficient and effective, highly developed services. Contemporary example of such a service might be stereoscopic 3DTV, where multimedia networks became prime objects of improvements and undergo constant updates. Since the information exchange is performed in an upgraded manner, the main goal of multi-service multimedia systems is to keep an appropriate quality level of service.

The main propelling power here is a motivation of transferring a certain focal point, from the service itself to the end user. An evolutionary leap from the contemporary 2D television creates a new range of business opportunities, however it is done by the cost of redesigning and remodeling of the already deployed IPTV systems. Likewise, the 3DTV service enhances the telco-branch [1]–[5], joining both content and service providers with facility manufacturers to create a separate layer of service-dedicated products.

However, in order to successfully introduce the 3DTV service to the market, quality evaluation must be performed. An identification of certain elements constructing the service is necessary to prepare a set of parameters varying the overall outcome of the deployed architecture.

2. Multidimensionality

In order to begin the analysis, one shall investigate the area of 3DTV video quality. The essential stage is to evaluate components contributing to the quality perception of the 3DTV content distribution service. Following the crude evaluation of 3DTV system, the overview of accessible model of quality evaluation is presented in the Fig. 1. By a brief inspection the multidimensionality of the problem concerning description and definition of quality within the 3DTV multimedia systems can be determined.

Starting the analysis, let's consider the guidance throughout the quality plane for the dedicated 3DTV system. Having divided the case into vertical and horizontal components of the content track, which generic contemporary telecommunication services are composed of, the quality model initiates the base layer for further conceptual model analysis. Let's follow the horizontally oriented walkthrough of the quality plane in order to determine possible anchor points for effective service quality definition.

<table>
<tr><td colspan="3">Multidimensional 3DTV quality model components & parameters</td></tr>
<tr><td>Back Office</td><td>Network</td><td>Front Office</td></tr>
<tr><td>Content Creation (techniques, coding parameters ...)</td><td>Bandwidth, packet loss</td><td rowspan="2">Viewing experience (image quality, depth quantity, naturalness, comfort, ...)</td></tr>
<tr><td>Source processing & Format (delivery-ready)</td><td>Delay, jitter synchronization</td></tr>
</table>

Fig. 1. 3DTV quality model, plane view.

From the scratch, the metrics used in quality evaluation supply the interested parties in essential information, towards the differentiation of what is perceived as good, mediocre or unacceptable quality. So, as the topic concerns services of stereoscopic data distribution, one might point out the necessity of including fundamental scheme of telecommunication services as the basic tool to model parameters determining Quality of Service (QoS). Led by research analysis [6] as well as intuitive approach, the aim of the work is to create a quality model representing insights of successful service creation. That is to satisfy the server-client interaction with apprehensible results. To verify from the point of view of the customer all the influential layers beneath the service, it is crucial to follow the path the viewed content crosses until its validity expires.

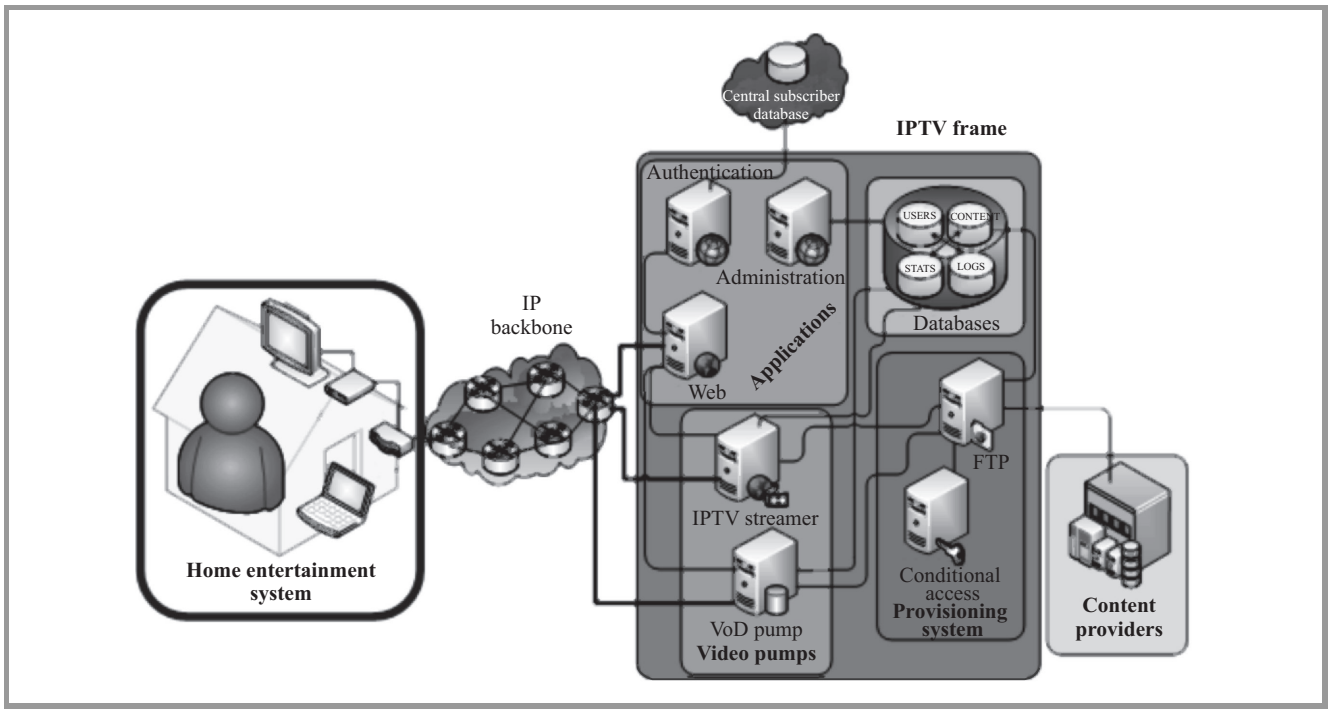

Fig. 2. Content chain inside IP based TV system.

3. Service Creation

The 3DTV service depicted on the scheme in Fig. 1 is composed of the three base vertical layers: Back Office, Network and Front Office. Core features of the aforementioned parts determine behavior of the service, varying parameters influencing final customer's experience (what is presented on the architectural layout from Fig. 2).

3.1. Back Office

The first component of the above mentioned quality plane, the Back Office module, is the origin of the service. In this unit, the video content is created, being at first acquired (in the case of real-life scenes) or synthetically built utilizing software/hardware accelerated 3D object generation. In the case of 3DTV scenario, the majority of applications is based on a video footage, constructed by using various acquisition techniques. One of the most important methods is the practice of utilizing the stereoscopic setup of cameras, camcorders receiving distinguished images. On this level of content creation a basic question rises – if the captured footage is properly composed.

Whenever the acquired video presents appropriate composition of images, thus supporting the Human Visual System (HVS) mechanics, the Back Office characteristics defines the manner in which the Content Provider delivers its product. Amongst several approaches, the ones being selected are the first order transmission of singular, elementary video streams (e.g. dual stream approach, raw data) or the utilization of high bitrate Multi-view Video Coding (MVC) expansion of a firmly established H.264 video codec [7]. Depending on the selected style, a particular stereoscopic pair of images might be given in a variety of setups. For instance, basic schemes allow the utilization of a Side-by-Side video frame positioning or Top-Bottom composition of the visual data. In some particular cases, a reduction of the vertical resolution of video frame might be observed (mainly due to throughput limitation [6], [8]).

As the content reaches the IPTV Frame (Fig. 2), on which the contemporary 3DTV service is based, asset management is performed. From the point of view of the perceived quality evaluation (QoE), this stage is of the upmost importance. Aside from ISP's content manipulation procedures, i.e., DRM or customer provisioning, the crucial point is the transcoding stage, where appropriate parts of content are transformed towards a network-native ingestion or network encapsulation, where the final video stream is formed.

3.2. Network

The area of 3DTV video quality assessment is also parameterized by a networking facility technological layout. Despite the utilized scheme of transportation (for instance Content Delivery Networks), the crude elements include the behavior of the transmission means by its resistance towards the peaks of workload. Especially in narrowband networks (i.e. wireless media) or in improperly managed fixed networks, despite the unicast paradigms of an independent content streaming, overloaded IP cores introduce phenomena which one can list as follows:

– a bandwidth limitation – significantly recognized by jerkiness or pixellate images, whenever high throughput is unavailable,

– a packet (departure/arrival) delay – in case of video sequences the inevitable timing is necessary, as a result of discarding information of no essential usage (arriving after the expected time of presentation).

– jitter – floating timing within packet ordering,

– stream synchronization – both audio and multidimensional video streams shall be delivered and decoded respectfully towards an encoder order scheme.

Network impairments take part while evaluation of the video quality is achieved. There exist several techniques utilizing network specific parameters to relate appropriate phenomena with a loss of the image quality observed on the user side [9]. From the point of view of a service developer, understanding of how the video stream is being encoded might lead to appropriate modeling and shaping of the network resources. Presented on the scheme in Fig. 3, is the generic approach of relating the visual data compression mechanisms for the 3DTV dual stream approach against the transmission manner.

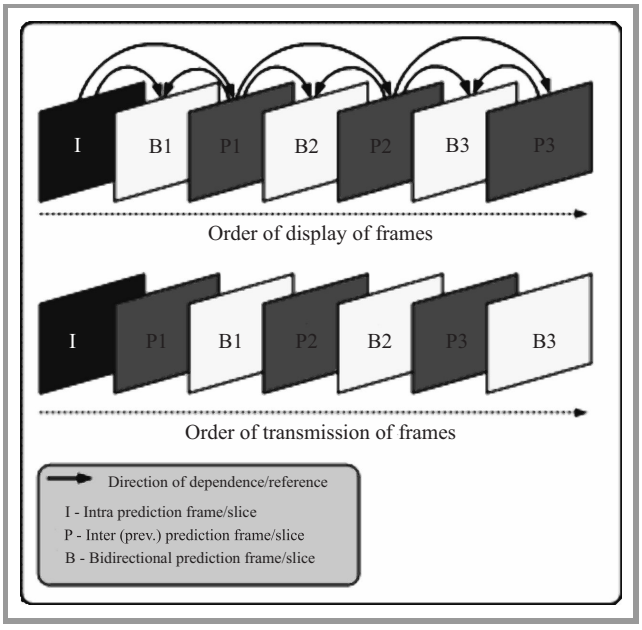

Fig. 3. Frames ordering for transmission and display.

Mentioned in [10], the prediction of inter type might be based on several frames including even those that are following to the current frame. In this point, differentiation between coded and displayed frames must be stated. The order of displaying and transmission (encoding) frames is different. Shown on the Fig. 3, the portion of 7 frames is a part of a structure called the Group of Pictures (GOP). The structure contains of diverse frames, including I, P and B types, what respectively means: intra predicted frame, inter predicted frame with reference to previous frame and bidirectional, inter predicted frame with the possible references towards the previous or further pictures. Referring to the GOP structure, it is an independent part of a motion picture containing, for instance 12 frames. The unit is recognizable by its completeness, and is always delimited by I frames (GOP consists of single "key" – I frame). Starting from the beginning, I frame is coded explicitly by the intra prediction manner, and to decode it there is no need for any previous frame data. The I picture is standalone, decodable frame and the one that has to be flawlessly transmitted. Intra prediction used for encoding process is based on the evaluation of a neighborhood of every single macroblock partition from within the frame. The I frame is thus the initial element of the Group of Pictures. Following, the P frame can be based only on intra coded frame or also the P frame. As depicted on the Fig. 3, the order of encoding determines the dependence the consecutive frames might (or not) refer to. For instance, since bidirectional frame B1 takes information about the position of certain macroblocks from both P1 and I frame, it has to be encoded after those two are completed. Therefore, the clarity of reconstruction and display of the visual data after the transmission is achievable whenever the complete sequence is not damaged. Thus, it is crucial to consider the mentioned network layer impairments while designing the policy for the 3DTV High Definition content and its requirements.

3.3. Front Office

The final stage of the content lifecycle is the Front Office layer of the quality plane. The stressed factor of the Front Office facility is the user's viewing experience. Depending on previously discussed parameters and content chain nodes, their contribution reflects the final score of perception (QoE).

The complexity of the customer side determines the layered approach to evaluate the quality of the perceived video stream. After the user's ordered stream is received and logically coherent (passing through gateway towards Set-Top-Box elements on Fig. 2), the decoding stage begins. Home Entertainment System, equipped with appropriate hardware solution, depicts 3DTV content utilizing certain compositions of the display technology and filtering stage. Market availability of Liquid Crystal Displays (LCDs), Plasma Display Panels (PDPs) and Digital Light Processing projectors (DLPs) supporting the required refresh rates extends the space of the quality evaluation problem. Moreover, sophisticated techniques of achieving stereoscopic imagery (i.e., active shutter glasses or polarization approach) add another dimension towards the definition of quality within 3DTV service, as stated in [5].

Thus, in case of a specific composition of the display technology and filtering technique, yet another case one has to consider, namely the environmental clutter of the reception area. Followed by hardware advancements, some of them might be neutralized, e.g. brightness nonuniformity caused by external reflections or manufacturing impairments. However, in explicit viewing conditions the crosstalk phenomenon will still exist [11].

Table 1
Perception of 3DTV content

Stereoscopic image perception – issues classification				
Display	Viewer	Captioning	Multi-view	Depth
Crosstalk	Condition (age, gender, previous experience)	Depth collision	Aliasing	Vergence – accommodation conflict
Geometry		Misalignment	Monocular occlusion	Depth mismatch
			View mismatch	Depth bracket
				Depth quality
				Depth discontinuity
				Interaxial distance
				Parallax

Summing up, one can identify the Front Office unit of the service creation scheme as having the heaviest impact on the perceived video quality within 3DTV content distribution systems. Further analysis shall emphasize the investigation and research of this area as crucial.

4. Viewer Experience

Referring to the content distribution chain from Fig. 2, final content perception is achieved on the Front Office side. Among various aspects of 3D content quality [12], the existence of the following areas of interest can be enumerated: image quality (assessed with the means of known metrics), depth quantity, naturalness of the scene or viewing comfort. As the 3DTV carries stereoscopic images/footage, the upmost importance is the creation of depth experience so as to perceive 3D scene. To achieve this phenomenon it is necessary to utilize consistent, coherent hardware to reproduce the feeling of the surround view. A basic classification of issues concerning 3D perception might be divided into following sections: display, user's comfort, subtitling/captioning, multi-view, depth of scene. One can distinguish multiple aspects of generation of stereoscopic vision in 3DTV systems, presented in Table 1, 3D-connected issues address the viewer's experience stage.
In order to clarify the positioned aspects of misperception observed while viewing certain content in a variable environment, the following description can be retrieved:

Display section

Crosstalk – phenomenon of image leaking from one channel to another, i.e., portion of a left-eye image seen in right-eye image, might be also observed as content dependent or not [13].

Geometry – distortions originated from irregularities of display, aside of the central point of the display surface.

Viewer section

Condition – as referred to the particular physical condition of the viewer, e.g., age, gender, the extent of the previous experience with 3DTV perception [14].

Captioning section

Depth collision – occurring while subtitles of 2D nature have to be overlaid on the 3D scene, especially of dynamic behavior.

Misalignment – discomfort of geometric gender, when left and right channel captions are not in a parallel alignment.

Multi-view section

Aliasing – depicted as high frequency components of 3D scene are improperly rendered on the display.

Monocular occlusion – occurring mainly in poorly converted 2D images into a stereoscopic view, when a part of the scene is seen only by one channel of HVS.

View mismatch – generated by impairments related to physical irregularities in acquisition hardware, e.g. lenses, sensors and inaccurate video post-processing, e.g. aperture, shutter speeds, white balance, encoding.

Depth section

Vergence – accommodation conflict – a distance dependent viewing issue: an appropriate perception of a virtual/projected 3D scene as opposed to focal comfort [5].

Depth mismatch – a phenomenon of inconsistency of HVS to reconstruct stereoscopy based on a single information cue (mismatch while natural motion of the viewer is noticed).

Depth bracket – depending on focus/vergence trade-off, the effective distance between the closest and the furthest element in the perceived scene.

Depth quality – observed in a 2D+Depth format issue of occlusive depth estimation [15].

Depth discontinuity – occurs when HVS is unable to track and follow depth variations within the fast motion scenes.

Interaxial distance – an issue of acquiring stereoscopic scene when the viewer's inter-ocular distance is smaller than (or varying) the disparity between the axes of the recording hardware lenses.

Parallax – oversaturated placement of the object due to divergence of HVS channels, when reconstructed scene tries to depict object in infinity.

The above mentioned positioning of the issues concerning the viewing experience depicts the overall complexity of the approach towards a coherent quality evaluation of 3DTV service. Some of the contemporary (2D) schemes of QoE assessment might be of use however, certain enhancements supporting the presented issues shall be obtained in order to create reliable quality metric.

5. Conclusions

The presented analysis of content quality evaluation within 3DTV service distribution systems depicts the complexity and multidimensionality of the discussed issue. The considered examination follows the crucial stages of service creation that further imply certain phenomena observed during stereoscopic content perception. The layered structure of the contemporary visual data distribution service determines the multiplicity of parameters being fundamental in conceptual and theoretical approaches towards the key point in 3DTV environment – Front Office section. The complexity of the considered issue rises rapidly, thus modeling of the quality evaluation tool becomes more compound. Resulting from research in the field of 3DTV content quality measurement, some of the mentioned parameters (especially those contributing to the content creation and processing) become plausible to determine. As for the final summary, the presented paper reveals how the contemporary IPTV, 2D TV service differs from 3DTV in terms of complexity of achievable and successful implementation of the enhanced telecommunication merchandise.

References

[1] S. Steinberg, "3DTV: Is the World Really Ready to Upgrade?", *Digital Trends*, Online Magazine, Jan. 7, 2010 [Online]. Available: http://www.digitaltrends.com/home-theater/3d-tv-is-the-world-really-ready-to-upgrade/

[2] H. Wallop, "CES 2010: 3D TVs on sale in UK by April", *The Telegraph*, Jan. 7, 2010 [Online]. Available: http://www.telegraph.co.uk/

[3] P. Otellini, Keynote Speech, Int. Consumer Electronics Show, Las Vegas (NV), Jan. 7, 2010.

[4] T. Spangler, "DirecTV to Launch 'n3D' Channel July 1", *Multichannel News*, Jun 28, 2010 [Online]. Available: http://www.multichannel.com

[5] D. Minoli, *3DTV Content Capture, Encoding and Transmission*. Chichester, UK: Wiley, 2010.

[6] B. Wonseok, L. SangJin, K. JeeGyun, and P. Sang-Il, "Korean terrestrial 3DTV broadcasting system: Current status", in *Proc. IEEE Int. Conf. Consumer Elec. ICCE 2011*, Las Vegas, NV, USA, 2011, pp. 889–890.

[7] I. E. G. Richardson, *H.264 and MPEG-4 Video Compression: Video Coding for Next-generation Multimedia*. Chichester, UK: Wiley, 2003.

[8] K. Byungsun *et al.*, "World's first hybrid 3DTV broadcasting experiment", in *Proc. IEEE Int. Symp. Broadb. Multim. Syst. Broadcast. BMSB 2012*, Seoul, South Korea, 2012, pp. 1–5.

[9] K. Wang *et al.*, "Perceived 3D TV transmission quality assessment: multi-laboratory results using absolute category rating on quality of experience scale", *IEEE Trans. Broadcast.*, vol. 58, no. 4, pp. 544–557, 2012.

[10] G. Wilczewski, "Directions of development of Video on Demand service", M.Sc. Thesis, Warsaw University of Technology, 2011.

[11] J. Bułat and L. Janowski, "Crosstalk Measurement in AGH Kraków", *VQEG Meeting*, Seoul, South Korea, June 2011.

[12] S. J. Daly, R. T. Held, and D. M. Hoffman, "Perceptual issues in stereoscopic signal processing", *IEEE Trans. Broadcast*., vol. 57, no. 2, pp. 347–361, 2011.

[13] L. Wang *et al.*, "Crosstalk evaluation in stereoscopic displays", *J. Display Technol.*, vol. 7, no. 4, pp. 208–214, 2011.

[14] S. Laframboise, D. D. Guise, and J. Faubert, "Effect of aging on stereoscopic interocular correlation", *Optom. Vision Sci*., vol. 83, no. 8, pp. 589–593, 2006.

[15] H. Wang *et al.*, "2D-to-3D Conversion Based on Depth-from-Motion", in *Proc. Int. Conf. Mechatr. Sci., Elec. Engin. Comp.*, Jilin, China, 2011, pp. 205–208.

Grzegorz Wilczewski received his B.Sc. and M.Sc. degrees in Electrical and Computer Engineering from Warsaw University of Technology, Poland, in 2009 and 2011, respectively. He is currently a Ph.D. candidate at Warsaw University of Technology. His research interests include 3DTV service quality monitoring, 3D imagery and digital signal processing.

g.wilczewski@tele.pw.edu.pl
Institute of Telecommunications
Faculty of Electronics and Information Technology
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland