Improved the Quality of Experience Assessment with Quality Index Based Frames over IPTV Network

Diego J. Botia[1], Javier F. Botia[1], Natalia. Gaviria G[1],
[1]Research Group in Telecommunications Applied GITA– ARTICA. Engineering Department. Universidad de Antioquia
Medellín, Antioquia, Colombia
Email: {dbotia, javier.botia, nagaviri}@udea.edu.co

José M. Menéndez[2] , David . Jimenez[2]
[2] Research Group in Applications of Visual Telecommunications - G@TV. E.T.S. Ingenieros de Telecomunicación.
Universidad Politécnica de Madrid – Spain
Email: {jmm, djb}@gatv.ssr.upm.es

ABSTRACT

Assessing the Quality of Experience *(QoE*) for multimedia services transmitted over a network has become a major issue for telecom providers (*Telcos*), because it can determine the real user satisfaction regarding the service they hire. For this reason, we present the metrics based on the Reduced Reference (*RR*) methodology called Quality Index Based Frame (*IQBF*) and New Undecodable Frames Index (*NUFI*) that are simple metrics to implement and it requires low computer resources. Also, high correlations are observed with metrics widely used, such as PSNR, VQM and SSIM, and clearly state their advantage.

Key Words— Objective Metrics, Subjective Metrics, IPTV, QoS, QoE

1. INTRODUCTION

Television has become the main vehicle of information and entertainment in the world, and has experienced a sound evolution over the years, ceasing to be passive to get in a new element that allows interaction with the user. It provides easy access to a large set of new services and applications (due to the capacity of access in the last mile to broadband networks for instance) using access with *ADSL2+, VDSL2, FTTx,* and *DOCSIS 3.0,* among others. Nordström [6] described that the telcos may increase their number of users by improving their levels of quality of service, strengthen their networks and infrastructure, and using standards, which are elements that can affect the contracted services. A mechanism is necessary to establish the real user satisfaction on the contracted services with the *Telco* especially in multimedia services, and that mechanism should fulfill the minimum quality levels defined in the *SLA* (Service Level Agreements). The services that generate inelastic traffic, such as voice and video impose certain constraints to the *QoS* metrics, where the delay, packet loss and delay variance (*jitter*) directly influence the quality of image and audio perceived by the user. To assess the quality from the user's perception, there are a set of metrics that determines the quality of experience (*QoE*), which evaluates the service more accurately and it can be measured both objectively (quantitatively) and subjectively (qualitatively) [8]. For hence, some artifacts, such as blurred vision (*Blur*), shaking of the video (jerkiness) or the Frame Loss Rate are perceived by the user as a bad image, generating a low *QoE*. The video quality is used to evaluate the performance of different compression and video processing systems, to control and monitoring of *QoS* through appropriate selection of system parameters. Different visual impairments depend on the video encoding, dynamic network conditions and the amount of motion and spatial details (textures, borders, etc) [18].

The impact of packet loss on video quality is great, mainly because each video sequence is composed of the so-called *GOP* (Group of Pictures), composed of a series of *I/P/B* frames that determine how it must be encoded and compressed before it can be transmitted over a network at different bitrates. The *I* frame in a *GOP* is the most important information and is used by *P* and *B* frames in order to encode and decode the sequence. Therefore, a single frame can be split into several packets, and if the video stream is sent over a network with few guarantees (e.g Besteffort), high congestion or link errors may happen that would cause the discarded or lost packet directly impacting on the *GOP*. It has been shown that the loss of a frame leads to what is known as error propagation due to the hierarchical structure of the *MPEG* encoding system. Error propagation generates visual artifacts that appear for packet errors (e.g slice error, blocking, ghosting, and freeze frame). As shown in Fig.1 if we lost one frame I, it would affect all sequences in images contained in the GOP. The artifacts can only occur in the event of a frame I lost, and the error propagates until the following frame I of the next GOP received. If we lose a frame B, it can only affect one image. When a drop frames B are presented, they are discarded due to their less impact on the video quality. Likewise, an error in the frame P will propagate right through the following frames P or B [1]. Frame P also has information about the motion vectors that it allows reconstructing the frame (temporal redundancy), therefore, frames P encode the changes or differences in the movement and saving bitrate. On the other hand, a stream that contains a high amount of frames B requires less bandwidth compared with a stream built with a high number of frames I or P. This feature is useful when setting the length of the GOP in coding. Therefore, we want to design metrics that may provide an estimation of the *MOS value* based on the set of lost frames, captured by the reduced reference methodology, and establish its correlation with other full reference metrics to establish their degree of accuracy. The main advantage is that the *telco* may easily assess a users' *QoE* and they could making improvements to coding or transmission parameters in event that the *QoE* is very low.

This paper is organized as follows; in Section II the main *QoE* metrics for *iDTV* platforms are explained. In Section III we detail the *IQBF* and *NUFI* metrics. In Section IV the *testbed* and the analysis of the results of the simulations are shown. Finally conclusions and future work are mentioned.

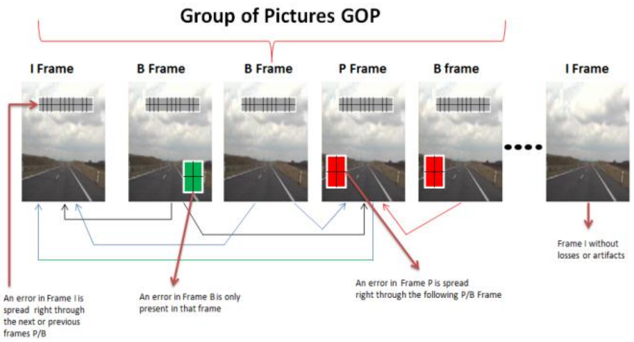

Fig 1. Effect of error propagation on the Frames I / P / B

2. RELATED WORK

Quality of Experience

The quality of experience (*QoE*) arises from the need to determine the degree of user satisfaction with the service. *QoE* has been defined in different ways by different authors. Li-yuan [9], indicates that the *QoE* involves two aspects: first one is to monitor the user experience online and the second one the service control to ensure that the *QoS* can widely know the user requirements. Furthermore, according to Lopez et.al [10], the *QoE* is an extension of the *QoS* in the sense that the former provides information about the delivery services from the viewpoint of the end user. According to Kilkki [11] and Winkler et.al [12], in the *QoE* there are subjective and objective measurements for the video transmitted over the network, where the former depends on the user's expectations and therefore, takes into account the feelings, perceptions and opinions. The objective metrics are computational models that predict the quality of the image perceived by one observer.

There are several parameters that affect significantly the *QoE* and are classified in three categories: encoding and compression parameters, network parameters and others. The encoding and compression parameters are related to mechanisms that control the amount of quality loss during the encoding process. These depend mainly on the encoding algorithm (*H.264*, *MPEG2*, etc.), bitrate, frame rate, temporal relationship with *P* or *B* frames, among others. Network parameters are obtained from the packetization of the video stream, transmission process through the network, such as *PLR*, *delay*, *jitter*, *bandwidth*, among others. The other parameters refer to the nature of the scene, the amount of movement, color, contrast, image size, as well as social and economic factors, user preferences, ages, etc. Due to the complexity of *HVS* (Human Vision System), there have been a lot of proposals for both objective and subjective metrics to try to approximate the real user experience, and the *ITU* has attempted to standardize some of them [19] [20] [21]. One of the main problems is that not always objective metrics can be adequately correlated with human perception.

For *IPTV* systems subjective methods are used to establish the performance of *TV* systems by using measurements that are more directly related to the perception of users. Through a series of video clips are score in a numerical predefined range, that allow to calculate the *MOS* (Mean Opinion Score). According to the *ITU* Ibid [13], *MOS* is a numerical measure used in multimedia traffic that determines the perceived quality of the data received after compression and / or transmission. This measure is in the range from 1 (lowest) to 5 (best). According to Kuipers et al [14], the minimum threshold of quality widely accepted is a *MOS* of 3.5. Although subjective tests (such as *MOS*) are quite useful in measuring user satisfaction, the implementation of them is complex due to time consumption and high costs involved.

Also, if we want to implement traffic management techniques in real time, it is necessary to find a relationship with objective metrics, measurable by network equipment objective quality metrics; according to Winkler Ibid [12], some algorithms are designed to characterize the quality of video and predict the viewer's opinion regardless of user perception. The most commonly used objective metrics are the metrics by the amount of reference information required, which can be classified as *NR* (No Reference), *FR* (Full Reference) and *RR* (Reduced Reference) [15] [16]. *FR* metrics are the most used and measures the degradation in the test video received related to the reference video located at the source. We need access to the full reference video, usually without impairment and compression, also imposing a spatial and temporal alignment, since each pixel in each frame is compared with the received video. The *PSNR* (Peak Signal to Noise Ratio) metric, is the best known, and evaluates the quality of the received video sequence and is mapped on a subjective scale [17] [23].

Even though several studies have used this mapping, it was found that the *PSNR* metric has the disadvantage that the image content is not verified and cannot identify artifacts that may appear for packets loss. In addition not always correlate with the real user perception, because only comparison is made pixel by pixel without performing an analysis of the structural elements of the image (e.g contours or specific distortions introduced either by the encoders or transmission devices in the network and spatial and temporal artifacts), therefore have been proposed some metrics that perform the extraction and analysis of features and artifacts on the video [12].

The *SSIM* (Structural Similarity Index Metric) [22] that calculates the mean, variance and covariance between the videos sent and received. For calculate the *SSIM*, 3 components are measured (luminance similarity, contrast similarity and structural similarity), which are combined into a single value called *SSIM* index, to range between 0 and 1; where 0 indicates zero correlation with original picture and 1 means that it is the same image [16]. Another important metric is *VQM* (Video Quality Metric) [3] [23], which takes as input the original video and processed video, and verifies the quality levels based on human eye perception and subjective aspects. *VQM* divides the image into spatial and temporal blocks–sequences measuring elements such as, blurring, the overall noise, block distortion and color distortion. The result that is close to 0 is regarded as the best value possible.

The *RR* metrics only selects some parameters of the original video, such as motion information. *RR* uses less processing and network resources although it requires an alternate channel for the transmission of these parameters and access to the video reference at some point. *NR* metrics analyze the output test video without the need for access to the video reference. This method is used when the encoding mechanism is known and network monitoring or special diagnostic operations can also be used, but the main problem is its a low correlation with *MOS* and heavy usage of computational resources.

3. QUALITY INDEX BASED FRAMES AND NEW UNDECODABLE FRAMES INDEX

Cruvinel et.al [25] introduced the *UF%* (Undecodable Frames Percentage) metric; where the number of frames from the video sequence that have errors are calculate. Then one or more packets have been lost or discarded by network errors or congestion, and therefore uses a percentage of the frames without decoding. Eq.(1) defines this metric:

$$UF\% = undFrms/totFrms \tag{1}$$

where *undFrms* is the number of frames with error and *totFrms* is the total number of frames from the video sequence. To facilitate the analysis, values of Eq (1) were normalized.

The authors conducted tests using sequences in *CIF* and *QCIF* resolution with multiple packet error rates and they compared the results with the *VQM* metric. Based on this study, we propose the New Undecodable Frames Index Metric (*NUFI*), where we consider the *I/P/B* frames prioritization, defined by their importance level within the *GOP*. For hence, we calculate the number of *I/P/B* frames lost and the total number of *I/P/B* frames of the sequence. An analysis was conducted from data behavior allowing to determine a weight, then it is assigned for each frame as follows: *I* frame with weight of 3, *P* frame with weight of 2 and *B* frame with weight of 1. This guarantees that more balanced according to the importance of each frame. For each *QoS* network type (*Besteffort* or *Diffserv*), the metric is applied in order to determine their behavior on the network and after compares with each FR metric. Eq. (2), defines the *NUFI* metric:

$$NUFI(nt,f) = \sum (UF(f) * weight) \tag{2}$$

where *nt* is the *QoS* network type and *f* is the frame type *(I/P/B*). *UF(x)* corresponds to the uncoded frames of each type. Table II presents the proposed algorithm to facilitate the calculation of the metric.

Serral-Gracia et.al [26] proposed a framework called *PBQAF* (Profile Based QoE Assessment Framework). This framework defines 3 status for frames (correct, disrupted and lost) through an analysis of the payload. Moreover, it performs a mapping to generate an associated quality index and it is captured from the payload of the received packets associated with a particular *PLR* (Packet Loss Rate). Eq.(3) presents the quality function to generate the mapping function *M:*

$$Q(f) = M(PLR(f)) \tag{3}$$

where *PLR(f)* is the packet loss rate of the frame *f*. The mapping function is given by Eq.(4):

$$M(x) = 1- x \tag{4}$$

where *x* determines the rate of frames loss, so when there is high packet loss rate then the quality index will be less and will tend to 0. Bearing this in mind, we propose a mapping between the *IQBF* metric and *MOS* metric considering each frame type (*I/P/B*). Eq.(5), shows the *IQBF* index:

$$IQBF(seq,nt) = \sum ((1-NFL(f)) / 3) - \alpha \tag{5}$$

where *seq* is the current video sequence to be evaluated, *nt* is the *QoS* network type (*BestEffort* or *Diffserv*) and *NFL(f)* is the number of frames *(I/P/B)* lost; α is an adjustment factor that was placed at 0.05 and this factor guarantees that the result of *IQBF* will be a positive value. Table I represents the proposed mapping between quality index and *MOS* metric and it will be evaluated in the *testbed* of the section IV.

TABLE I
MAPPING IQBF VS MOS

MOS	IQBF VALUE
5(Excellent)	>=0.85
4 (Good)	>=0,65 -- < 0,85
3(Acceptable)	>=0,45 -- < 0,65
2 (Poor)	>= 0,25 -- < 0,45
1 (Bad)	< 0,25

where $IQBF(f) \in [0,1]$ and as shown in Table I for a value $IQBF(f) > 0.65$, yields a very high MOS, which means a video with minimal artifacts.

4. TESTBED AND RESULTS

Testbed were implemented in two scenarios that facilitated the simulation. The former is a *Besteffort* network (no *QoS*) using *FIFO* queues (First In First Out). The latter scenario was a network with *QoS* using *Diffserv* scheme, using the congestion avoidance algorithm *WRED* (Weighted Random Early Detection) [4], which employs 3 virtual queues to those they applied *TSW3CM* policy. Scheduler mode that is set up for the priority queue 0 was *PRI* (priority queuing with average rate limitation) [24]. All scenarios were simulated using *NS-2* and *Evalvid* framework [5]. The general simulation process begins with the selection of different video *RAW* sequences uncompressed in format *YUV* with 4:2:0 video color modes, which led to the *ffmpeg* and main concept tool encoder to adapt at different bitrates and *GOP*s. Video traces were generated suitable to send over the network through packet encapsulation. A *MTU* of 1024 bits is applied, using *RTP* (Real Time Transport Protocol) through the *MP4trace* tool, where information is obtained as the number of frame, the frame type *(I*, *B* or *P*), the frame size, the sequence number and timestamp. *MSU VQMT* tool [7] was used to obtain the *Y-PSNR*, *SSIM* and *VQM* metrics, considering, the original reference video and the video distortion received.

TABLE II
ALGORITHM PROPOSED FOR NUFI

```
Input: NT { Network Type  BestEffort ←---- 0 ;  Diffserv ←---- 1}
        NFIL   NFPL   NFBL   {NFIL  Number Frames I Loss; NFPL
                                 Number of Frames P Loss  NFBL
                                 Number of Frames B Loss}
        NFI  NFP  NFB  {NFI  Number Frames I ;  NFP   Number of
                              Frames P   NFL  Number  of Frames B }
Output: NUFI(NT)   {New Undecodable Frames Index}
j← 0
QoE_Index(NT,NFIL,NFPL,NFBL,BR,GOP)  {BR Bitrate, GOP
                                      Group of Pictures}
For all j< = NumFlows(NT)  Do
  If frame = I   Then
    UFI(j) ←  NFIL(j) / NFI   weight ← 3
Else If frame = P then
    UFP(j) ← NFPL(j) / NFP   weight ← 2
Else
    UFB(j) ← NFBL(j) / NFB   weight ← 1
    End IF
   End IF
End  IF
     NUFI(NT,j) ← $\sum$ UF(frame) * weight
   End For
End
```

Fig. 2 shows the scenario formed through a video sender (Server Video on Demand) and 9 sources of cross traffic that consists of *CBR* and *On-Off traffic* sources. The video stream *MPEG-4* comes complete with background traffic flows on-off which it has an exponential distribution with average size packet with 1500 bytes, burst time 50 ms, idle time 0.01 ms, and rate 1 Mbps. The access network represents a video receiver (simulating a last mile *ADSL2*) with a bandwidth link of 20 Mbps and with several backgrounds traffic sinks for each one links for a bandwidth of 10 Mbps.

We tested the traffic behavior and *QoE* metrics with different percentages of errors on the link that will be the bottleneck and was established between the core router and the access router, using a loss model with uniform distribution with rates of 0%, 1%, 5% and 10%, a bandwidth of 10 Mbps and a delay of 5 ms.

Fig. 3 shows some screenshots of the videos evaluated (Spanish Public TV News, Mass, and Highway [27]) set at a resolution of 720x480 standard definition under the *NTSC* standard. Table III presents the encoder parameters. For each video stream several parameters are combined such as *GOP* length (10, 15 and 30), bitrates recommended by the *DSL Forum* [2] (1.5, 2, 2.5 and 3 Mbps and packet loss rates over the access link to the network for both *Besteffort* and *Diffserv* network, which resulted in 289 different data.

Fig. 4 and 5 show the behavior of the *NUFI* and *IQBF* metrics regarding the *QoS* network type, *GOP* and *PLR*. High quality values for the network with *QoS* and low *PLR* (0% and 1%) was observed. The results show that we increasing losses and *GOP* length, the *IQBF* Index decreases and the *NUFI* index increases. In all cases, we found better values of *NUFI* and *IQBF* applying the *QoS* network. For all sequences were derived the correlations for *IQBF*, *NUFI* and *UF%* (normalized values) in relation with *VQM*, *Y-PSNR*, *SSIM* and *MOS* metrics.

Fig 6 presents an example of the correlation with *SSIM*, where value of the correlation coefficient R^2 is higher for *IBQF* and *NUFI*. We also found in the lower right side of the chart that the values are close to 1 for *SSIM*, which derive from video sequences where the *Diffserv* strategy was applied, thus have a better *QoE* values, and it was proved in the different scenarios simulated. This behavior was similar for all video sequences and whole full reference metrics. Fig 7 shows the correlation between the *MOS IQBF* determined by Table I and the *MOS* reference calculated from the *VQM, PSNR* and *SSIM* metrics for sequence highway. In both cases, a correlation value $R^2=0.829$ was calculated which is good for our experiment.

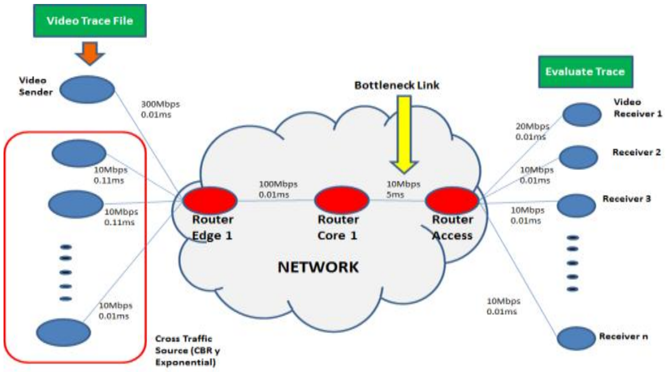

Fig 2. Scenarios Developed with NS-2 (BestEffort / Diffserv).

TABLE III
ENCODER PARAMETERS

Length GOP	10, 15 and 30 Frames
Frame Rate	30 fps
Bitrate	1.5 a 3 Mbps
Frame Type Support	I, B, P
Frame Sequence	IPBBPBBPBBP…..
Resolution	NTSC 720x480 p
Video Color Mode	YUV (4-2-0)
Sequence News TVE (High Level Activity)	2931 Frames Duration 1 min:31 sec
Sequence Mass (Low Level Activity)	1728 Frames Duration 57 sec
Sequence Highway (High Level Activity)	2398 Frames Duration 1 min:17 sec

Fig 3. Screenshots from video sequences assessment.

Table IV represents the correlation coefficients R^2 and standard errors for the sequences obtained according to the relationships for all *QoE* metrics evaluated. High correlation values were obtained in the majority of metrics, showing an approximation to the user´s perception due to that metrics such as *SSIM* and *VQM* are considered elements of the *HVS*. Also better correlation values were achieved for *NUFI* and *IQBF* than for *UF%*.

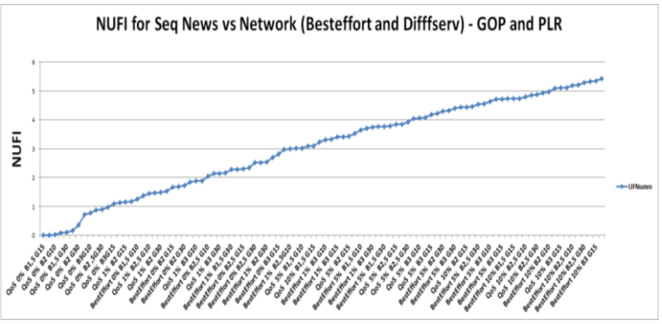

Fig 4. Behavior of NUFI with Network, GOP and PLR for sequence news.

5. CONCLUSIONS AND FUTURE WORK

The simplicity in the implementation of metrics based on frames as *NUFI* and *IQBF* have allowed a good correlation with metric widely used as *Y-PSNR*, *SSIM* and *VQM*. There are higher values of linear correlations *for ICBF* and *NUFI* with regard to the *UF%* metric, because the loss of *I/P/B* frames due to *PLR*, and the each frame weights within the *GOP* was considered. It was observed the negative impact that is presented on the *QoE* for *PLR* greater than 1% and for *GOP*s with length greater than 10 frames, because by extending this length in case of *I* frame lost, the propagations effects of the error will be larger and number of artifacts will be more visible by the user. The proposed metric may be simplifying the assessment strategies for video quality because it requires few computing resources as opposed to full reference metric. The metric *IQBF* and *NUFI* only require capture packets by analyzing its behavior on the network and establishing the amount of lost frames. Also best values for *IQBF* and *NUFI* were obtained to implement the *Diffserv* strategy. Results show that the proposed metrics allow with a simple way approach to the *Telco,* to obtain the *QoE* real values that the users could observe without using intensive computing resources; thus a possible savings of time and money to the *Telco* is generated. As future work we expect to apply he proposed metrics to high definition and 3D video streams as well as test its behavior with other *QoS* strategies.

ACKNOWLEDGMENT

The authors acknowledge the support provided by *ARTICA* (Regional Alliance for ICT applied), which are part of the University of Antioquia, *EAFIT*, *UPB*, *UNAL*, *IPS* Universitaria, University *ICESI* and *UNE* Telecommunications as well as to Colciencias and the Ministry of ICT. We also appreciate the support provided by the Engineering Department of the University of Antioquia and the Superior Technical School of Telecommunications at the Universidad Politécnica de Madrid–Spain.

REFERENCES

[1] J. Greegrass, J. Evans, and A. Begen. “Not All Packets are Equal Par 1. The Impact of Network Packet Loss on Video Quality,” *IEEE Internet Computing*, 2009.

[2] DSL Forum, “*Triple-play Services Quality of Experience Requirements and Mechanism*,” Working text WT-126 version 0.5, Feb. 21, 2006.

[3] F. Xiao, VQM–DCT Based Video Quality Evaluation [Online]. 2000. Available: http://compression.ru/video/quality_measure/vqm.pdf

[4] M. Barbera, A. Lombardo, G. Schembra, and A. Trecarichi, “Improving fairness in a WRED-based DiffServ network: A fluid-flow approach,” ACM Digital Library. *J. Perfor. Eval*. Vol. 65, Issue 10, Oct. 2008.

[5] J. Klaue , N. Rathke, and A. Wolisz, "Eval Vid – A Framework for Video Transmission and Quality Evaluation," *13th Int Conf. on Model. Tech and Tools for Comp Perfor Eval*, pp. 255-272, Urbana, Illinois, USA, Sep. 2003.

[6] E. Nordström. Overview of IPTV Systems. Ernst Consulting & Education [Online]. Sweden. 2009. Available: http://www.du.se/~eno/itv/iptv-system.pdf

[7] D. Vatolin, MSU Video Metric Quality Tool. MSU Graphics and media Lab [Online]. 2012. Russia. Avalilable: http://compression.ru/video/quality_measure/video_measurement_tool_en.html

[8] S. Winkler, *Digital Video Quality–Vision,Models and Metrics*, Ed Jhon Wiley & Sons, Switzerland, 2005.

[9] L. Li–yuan, Z. Wen–an, and S. Jun–de, “The Research of Quality of Experience Evaluation Method in Pervasive Computing Environment,” *1st Int. Symp. on Pervasive Comp. and Appl*, 3–5 Aug. 2006, pp. 178– 182.

[10] D. López, F. González, L. Bellido, and A. Alonso, “Adaptive multimedia streaming over IP based on customer oriented metrics”, *Int. Symp. on Comp. Net.*, 16–18 Jun. 2006, pp. 185–191.

[11] F. Kuipers, R. Kooij, D. Vleeshauwer, and K. Brunnstrom, “Techniques for Measuring Quality of Experience [Online]. Book Series: Lecture Notes in Computer Science, Ed. Springer ISSN0302–9743, pp. 216–227, May 2010. Available: http://www.springerlink.com/content/5028804658914365

[12] S. Winkler, and P. Mohandas, “The Evolution of Video Quality Measurement: from PSNR to Hybrid Metrics,” *IEEE Trans. Broadcasting*, Vol. 54, N. 3, Sep. 2008.

[13] K. Kilkki, K. “Next Generation Internet and QoE,” *Presentation at EuroFGI IA.7.6 Workshop on Socio–Economic Issues of NGI*, Santander, Spain, Jun. 2007.

[14] S. Winkler, “Video Quality Measurement Standars–Current Status and Trends,” *7th Int. Conf, on Inf., Comm. and Signal Processing*, 2009. ICICS 2009, IEEE. ISBN: 978–1–4244–4656–8.

[15] F. Boavida, E. Cerqueira, R. Chodorek, M. Grega, M. Leszczuk, Z. Papir, P. Romaniak, and C. Guerrero, “Benchmarking the Quality of Experience of Video Streaming and Multimedia Search Services: The content Network of Excellence,” *KSTiT 2008 - XXIII Symp of Telecomm and Teleinfor, Inst of Telecomm. and Electrotechn of Univ of Techn and Life Scien.*, Bydgoszcz, Poland, 10–12 Sept. 2008.

[16] E. Cerqueira, L. Veloso, M. Curado, and E. Monteiro, “Quality Level Control for Multi-user Sessions in Future Generation Networks,” University of Coimbra INESC Porto. Portugal A. Khan, L. Sun, E. Ifeachor, Content Classification Based on Objective Video Quality Evaluation for MPEG4 Video Streaming over Wireless Networks. *Proc. of the World Congr. on Eng*. 2009, Vol. I, WCE 2009, London, U.K. Jul. 2009.

[17] K. Chih H, C. Shieh, W. Hwang, A. Ziviani, “An Evaluation Framework for More Realistic Simulations of MPEG Video Transmition,” *J. of Inf. Scien. and Eng.*, N. 24, 2008, pp 425–440.

[18] J. Greegrass, J. Evans, and A. Begen, "Not All Packets are Equal Par 2. The Impact of Network Packet Loss on Video Quality," *IEEE Internet Computing*, On page(s): 74–82 Volume: 13, Issue: 2, Mar–Apr. 2009.

[19] ITU FG–IPTV.DOC–0814, "Quality of Experience Requirements for IPTV Services," Dec. 2007,

[20] ITU–T Recommendation P.910, "Subjective video quality assessment methods for multimedia applications". *International Telecommunication Union*, Geneva, Switzerland, 1999.

[21] ITU–T Recommendation. *QoE requirements in consideration of Service Billing for IPTV Service*, UIT–T FG–IPTV, 2006.

[22] Z . Wang , A. Bovik, H. Sheikh, and E. Simoncelli, "Image Quality Assessment: from Error Visibility to Structural Similarity," *IEEE Trans. on Image Processing*, Vol. 13, no. 4, pp. 600–612, Apr. 2004.

[23] M. Vranješ, S. Rimac–Drlje, and D. Žagar, *Objective Video Quality Metrics*, University of Osijek, Faculty of Electrical Engineering. 2006.

[24] P. Pieda, J. Ethridge, M. Baines, and F. Shallwani. "A Network Simulator Differentiated Services Implementation," *Open IP*, Nortel Networks, 2000.

[25] L. Cruvinel, and T. Vazao. A., "Simple Metric for Predicting Video Quality of Experience," *IEEE 10a Conferência sobre Redes de Computadores CRC2010* [Online]. 11–12 Nov. 2010, Universidade do Minho, Braga. Portugal. Available: http://cnm.tagus.inesc-id.pt/files/Laercio-CRC2010-revisto.pdf.

[26] R. Serral–Garcia, Y. Lu, M. Yannuzzi, X. Masip–Bruin, and F. Kuipers. Packet Loss Based Quality of Experience of Multimedia Video Flows [Online]. Crente de Recerca d'Arquitectures Avancades de Xarxes (Politencic University of Cataluña), Spain, Delft University of Technology, Netherland, 2009. Available: http://personals.ac.upc.edu/rserral/research/techreports/psnr_mos.pdf

[27] ASU. Video YUV trace files [Online]. Arizona State University, 2012. Available: http://trace.eas.asu.edu/yuv/index.html

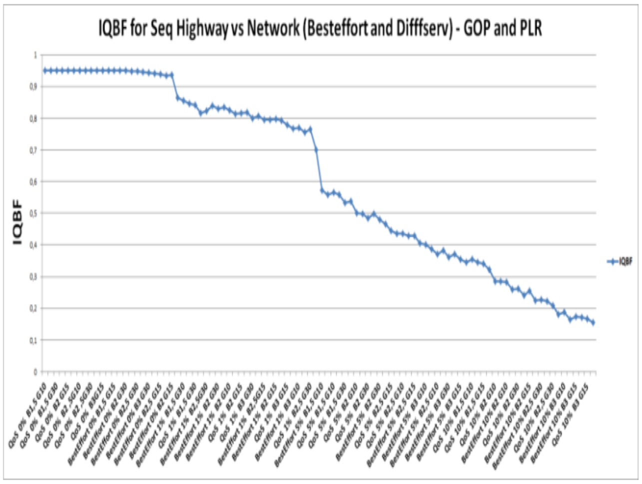

Fig 5. Behavior of IQBF with Network, GOP and PLR for sequence highway.

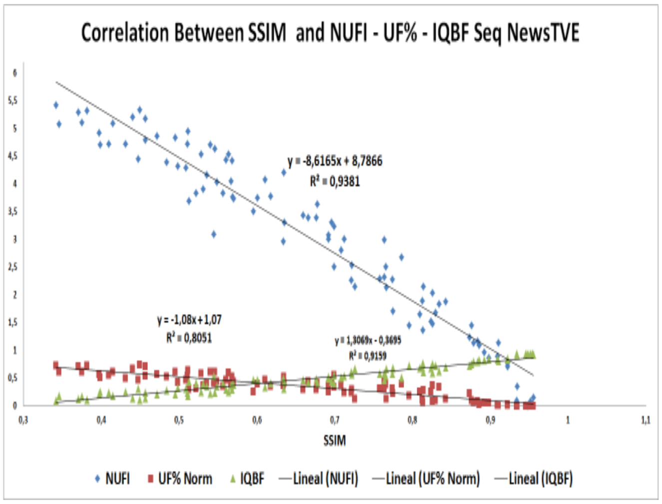

Fig 6. Correlation between SSIM and NUFI, UF%, IQBF for sequence news.

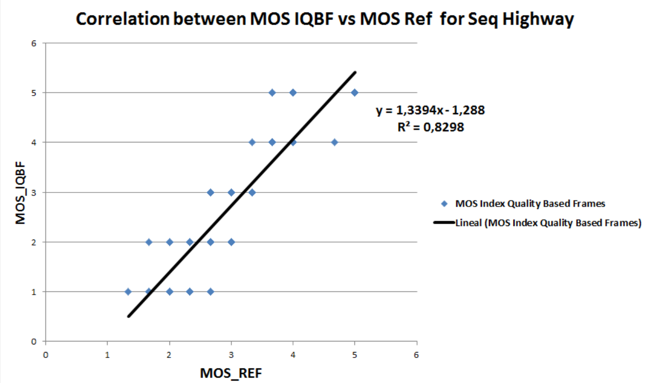

Fig 7. Correlation between MOS IQBF vs MOS Reference for sequence highway.

TABLE IV

RESULTS FOR QOE METRICS–CORRELATION COEFFICIENT R^2

METRIC	NEWS		HIGHWAY		MASS	
	R^2	STD ERROR	R^2	STD ERROR	R^2	STD ERROR
PSNR vs NUFI	0,947	1,94465	0,829	2,81083	0,81451	3,23389
PSNR vs UF%	0,808	3,7075	0,798	3,06231	0,51294	5,24029
PSNR vs IQBF	0,918	2,41011	0,825	2,84730	0,76302	3,65525
MOS vs NUFI	0,906	0,39728	0,860	0,35989	0,79791	0,54596
MOS vs UF%	0,755	0,64123	0,834	0,39153	0,48203	0,87406
MOS vs IQBF	0,870	0,46624	0,858	0,36204	0,73767	0,62203
MOS IQBF vs MOS_Ref	0,844	0,51232	0,829	0,39733	0,67701	0,69021
SSIM vs NUFI	0,938	0,04424	0,896	0,02409	0,67201	0,11793
SSIM vs UF%	0,805	0,07848	0,892	0,02453	0,49774	0,14593
SSIM vs IQBF	0,915	0,05154	0,902	0,02342	0,66499	0,11918
VQM vs NUFI	0,935	1,12254	0,854	0,86563	0,78018	1,87899
VQM vs UF%	0,811	1,91725	0,840	0,90593	0,49749	2,84099
VQM vs IQBF	0,915	1,28580	0,856	0,86026	0,73578	2,06006