# OBJECTIVE PERFORMANCE EVALUATION FOR MPEG-4 VIDEO STREAMING APPLICATIONS

M. IVANOVICI*

*Utilizând sistemul de monitorizare a traficului de rețea pe care l-am conceput şi implementat, am evaluat performanțele unei aplicații de transmisie video MPEG-4. Pentru a măsura într-un mod obiectiv calitatea percepută de către utilizator, am definit următoarele două metrici: numărul de cadre video pierdute şi numărul de cadre video alterate. Experimental, am determinat dependența dintre calitatea percepută de către utilizator şi degradarea calității la nivelul rețelei prin pierdere de pachete.*

*Using the network traffic monitoring system we designed and implemented, we evaluated the performance of an MPEG-4 video streaming application. In order to objectively assess the user-perceived quality, we defined two metrics: the number of dropped video frames and the number of altered video frames. We determined experimentally the dependence of the user-perceived quality for an MPEG-4 video streaming application on the quality degradation–through packet loss–at network level.*

**Keywords:** performance evaluation of network applications, user-perceived quality, MPEG-4, video streaming.

## Introduction

Video streaming, or video over IP, is a widely-used application in today's Internet. Its real-time characteristics make it a very demanding application, in terms of network services. The amount of bandwidth it requires depends on the compression rate, and implicitly on the quality of the video signal. Video streaming also requires high reliability for the data transfer between the streaming server and the client.

The MPEG standards [1] describe the most frequently used compression algorithms for video sequences. The purpose of the MPEG-4 standard is to establish a compression system for extremely low bit rate transmission. The high compression rate it achieves reduces the required bandwidth for a video streaming application, with respect to that of previous standards. However, this also makes the application vulnerable to any packet loss at network level.

---

* Research Assistant, Applied Electronics and Information Technology Chair, "POLITEHNICA" University Bucharest, Romania.

Network applications usually require a minimum Quality of Service (QoS) level in order to run according to user expectations [2], [3]. Knowing the application requirements makes it possible to predict whether a certain connection is valid for a certain application and what will be the user-perceived quality (UPQ) for that application.

A survey on QoS application needs was published by the Internet2 QoS Working Group [4], but their approach is not objective and the conclusions are vague. TF-STREAM reported on best-practice guidelines for deploying real-time multimedia applications [5]. ITU-T defined network performance objectives for IP-based services in [6]. HEAnet reviewed several aspects of perceived quantitative quality of applications [7]. Most of these approaches are qualitative, whereas we aim at creating a quantitative representation of UPQ that can be related to QoS parameters.

There are two types of metrics used to quantify the quality of a video sequence: *subjective* and *objective*. Subjective video quality measurements are time consuming and must meet complex requirements. ITU recommendations [8], [9], [10], [11] formalise the procedure for subjective measurements, by specifying the conditions of the experiments—like viewing distance and room lighting—as well as data analysis methods. The objective metrics can be implemented as algorithms and are human error free. They are either based on the human vision system or on distance measures, like root mean squared error (RMSE) or peak signal noise ration (PSNR). These simple measures do not capture the user-perceived degradation in the video signal—image attributes like sharpness and colorfulness should be taken into account [12], [13]. The metrics can also be *with reference*—the sequence at receiver is compared to the original sequence at transmitter—or *without reference*, when only the sequence at receiver is analysed.

The Video Quality Experts Group (VQEG) [14] reported on the perceptual video quality measurement algorithms [15]. A survey of video quality metrics based on models of the human vision system can be found in [16]. Several no-reference blockiness metrics are studied and compared in [17]. Most of the existing metrics for the video quality quantify the degradation introduced by the compression algorithm itself or by the frame rate that is used. There are no metrics or studies that objectively assess the degradation in video quality caused by the packet loss at network level.

## 1 Experimental setup

The test setup we used is depicted in Figure 1. A detailed description of the monitoring system is available in [1]. The application traffic flowing from the server to the client is analysed; based on the traffic traces, we quantify the quality degradation induced in the network emulator: one-way delay, jitter and packet loss, as instantaneous or average values, as well as histograms.

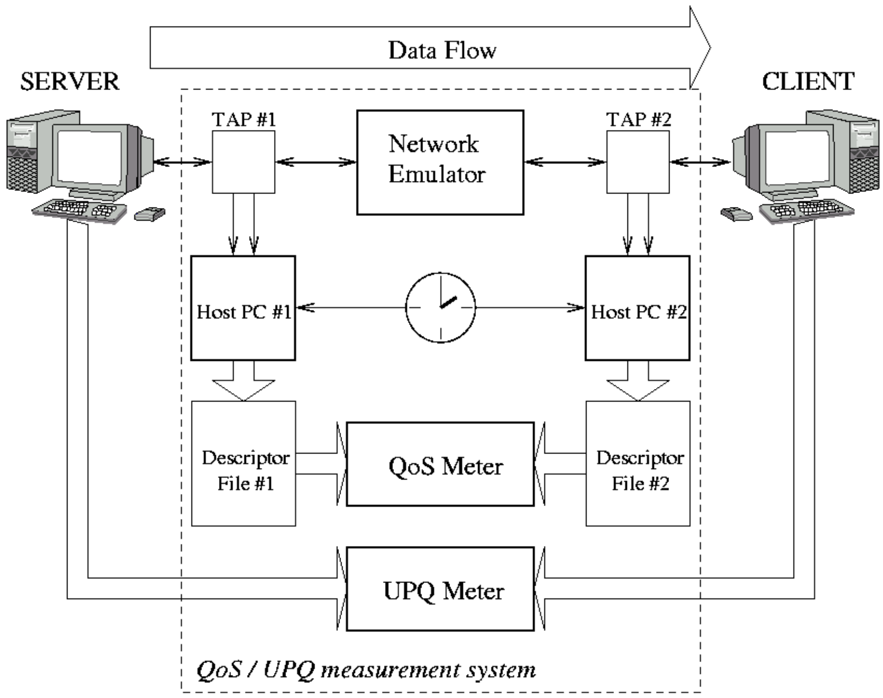


Figure 1. Test setup.

In case of an MPEG video-streaming application, the server packetizes the MPEG stream in order to send it to the client that requested it. The protocol used for the control of the streaming is RTSP (Real-time Streaming Protocol) or RTP (Real-time Protocol). Details on the RTP payload for MPEG-4 streams can be found in [18] and [19]. The transmission protocol used by the application under study is User Datagram Protocol (UDP). UDP traffic is an *inelastic* traffic, i.e., the application doesn't adjust its transmission rate to network conditions. In addition, lost packets are not retransmitted. Therefore, packet loss at network level will cause gaps in the MPEG video stream.

The MPEG-4 streaming server we used was the Helix streaming server from Real Networks [20]. The MPEG-4 client "mpeg4ip" [21] was modified so that every video frame that is rendered on screen is saved on disk as an individual bitmap file. Based on these files, one can recreate the video sequence at reception, as it would have been seen by a potential user. Using the metrics described in Section 2, we then compute the user-perceived quality for those video sequences. What follows is the correlation of the measured network quality degradation with the calculated UPQ.

## 2 Proposed UPQ metrics for video streaming applications

We propose two objective reference-based metrics for the assessment of the user-perceived quality for video streaming applications: the *number of dropped video frames* (NDF) and the *number of altered video frames* (NAF).

NDF is computed as the difference between the number of frames in the original video sequence at transmitter (server) and the number of video frames that are effectively rendered at receiver (client). This number indicates how many frames were skipped because of the missing bits in the MPEG video stream.

NAF indicates how many frames—from the ones received and rendered—are affected by impairments. Figure 5 (c), (d) shows two such altered frames.

## 3 Experimental Results

In the experiments we performed, we introduced artificial packet loss using the NIST Net network emulator [24]. Packet loss was introduced in the server-client direction, with values between 0 and 1%.

We ran the tests using two different MPEG-4 video sequences: "football" and "train" (Figure 5 (a) and (b), respectively). The video sequences are 10 seconds long, with 250 frames, each of 320 x 240 size. The average transmission rate was approximately 1 Mb/s.

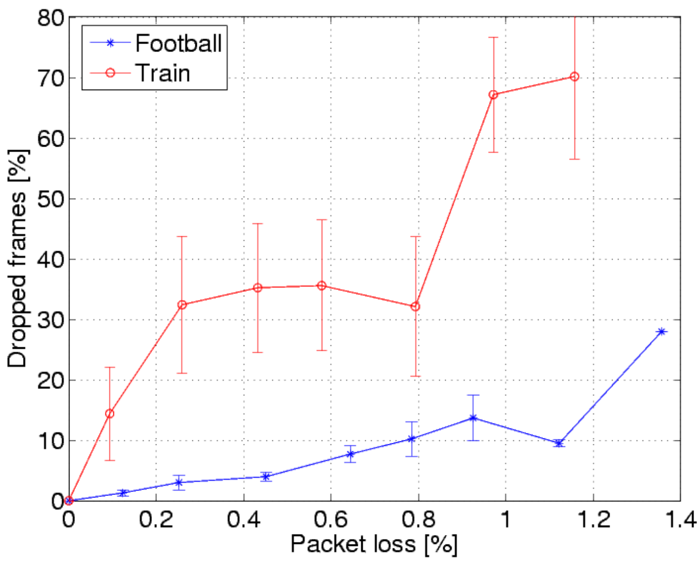


Figure 2. The percentage of dropped frames as function of packet loss.

The decreased number of altered frames, for loss rates exceeding 0.8% (see Figure 3), is a consequence of the increased number of dropped video frames. This shows that severe degraded frames are no longer rendered.

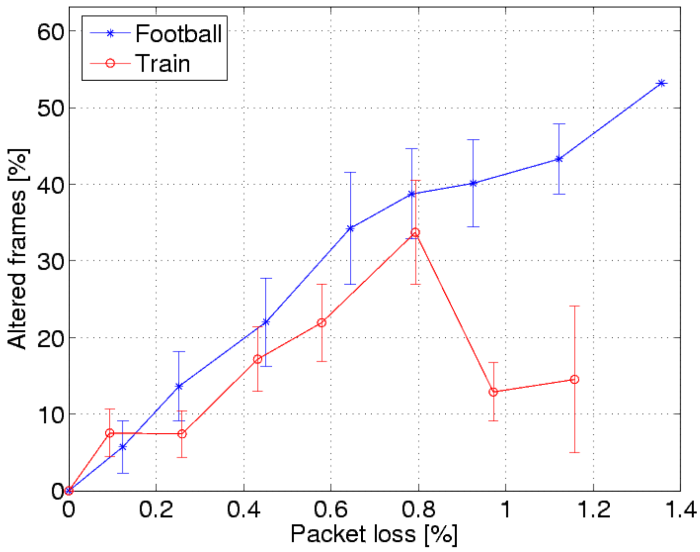


Figure 3. The percentage of altered frames as function of packet loss.

By putting together the two metrics, one can plot the *total number of affected frames* (TNAF)—both dropped and altered—as a function of packet loss at network level. The monotonic increase of TNAF can be observed in Figure 4.

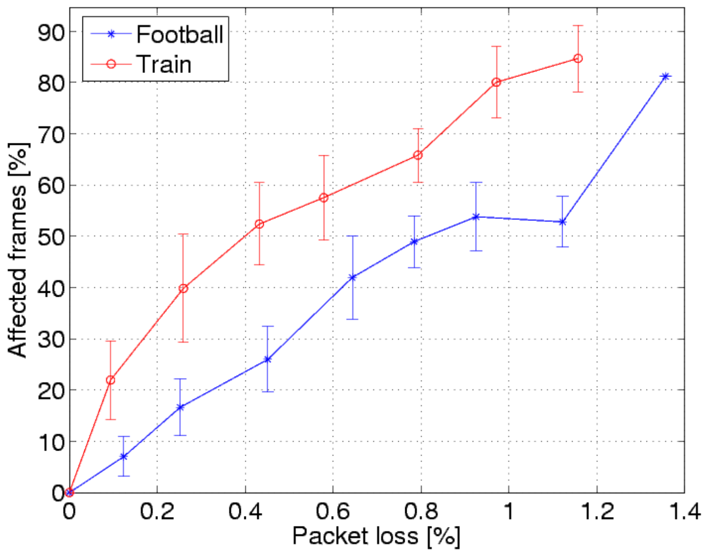


Figure 4. The total percentage of affected frames (dropped and altered) as function of packet loss.

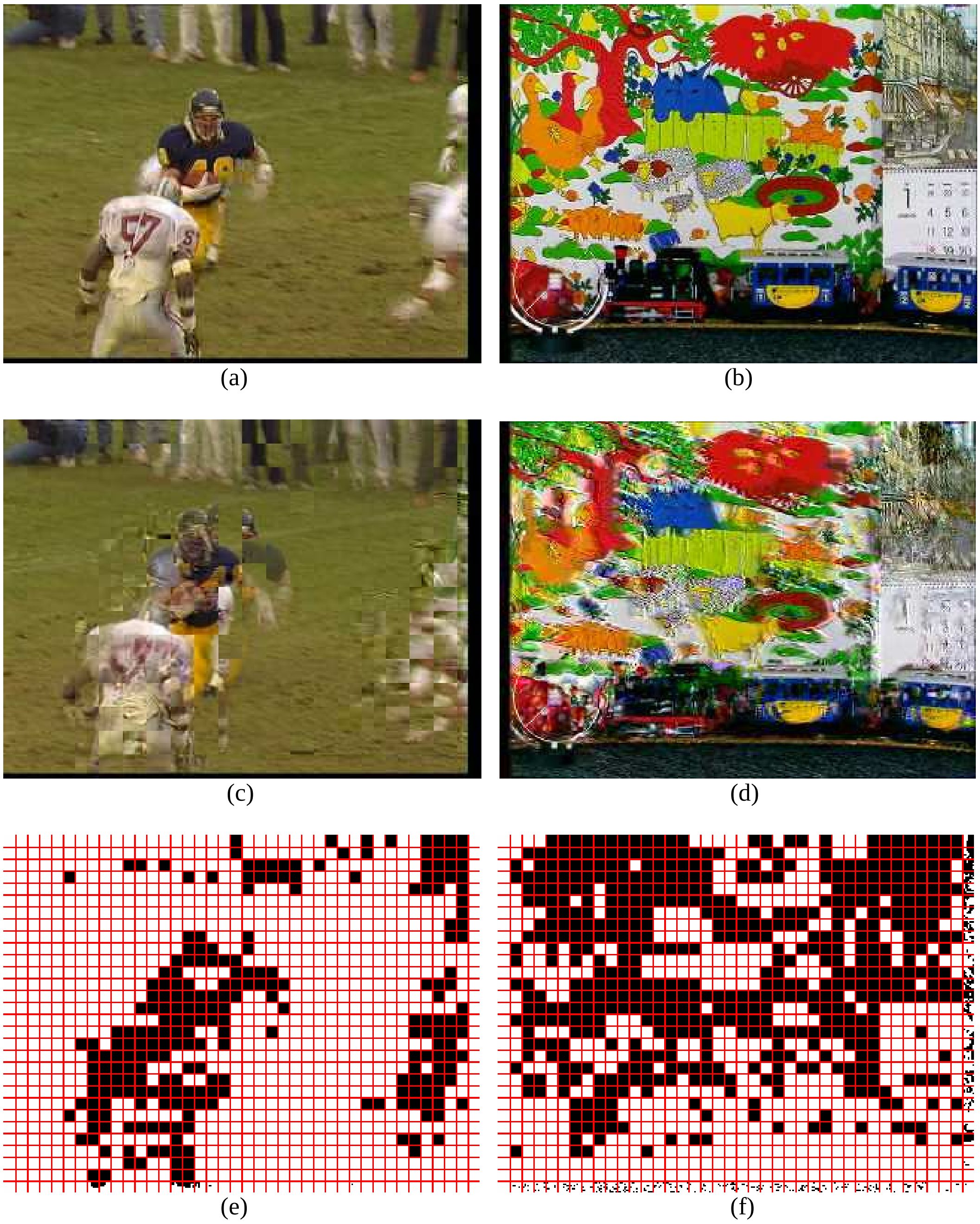


Figure 5.Original (a), (b) and degraded frames (c), (d) from the video sequences “football” and “train”, respectively. Difference images on 8x8 blocks (e), (f) between original and degraded frames.

In Figure 4 one can observe that, for example, a 0.6% packet loss causes 60% of the frames to be affected. This shows how vulnerable to packet loss an MPEG-4 video streaming application may be. Losing one packet containing the information of a I (intra) frame from the MPEG-4 stream implies the degradation of all the following P (predictive) or B (bi-directional predictive) frames. Altering the information of a P frame implies only the degradation of another adjacent frame. Altered B frames do not cause the degradation of other video frames.

Figure 5 shows two original frames (a) and (b) from the video sequences "football" and "train", respectively. Figure 5 (c) shows a frame with 21.38% of the pixels (20.58% of the 8x8 blocks) being affected by impairments, while the frame in Figure 5 (d) has 44.50% of its pixels (46.25% of its 8x8 blocks) affected by impairments. Figures 5 (e) and (f) represent the difference images on 8x8 blocks, for the original and degraded frames.

## Conclusions

The novelty of our work relies in the fact we accurately measure network quality degradation and objectively assess application UPQ in parallel. This allows us to experimentally determine the dependence of UPQ on network quality degradation for any particular application.

In this paper we focused on an MPEG-4 video streaming application. We proposed two objective reference-based metrics for the user-perceived quality of the video sequence.


## Acknowledgements

The author would like to thank Marius Malciu (CERN, Geneva) for the fruitful discussions and suggestions regarding the processing of the video data. The comments made by Razvan Beuran (CERN, Geneva) and Marius Preda (INT, Evry) helped improve the final version of this article.


## R E F E R E N C E S


[1] Motion Picture Experts Group, http://www.mpeg.org

[2] *R. Beuran, M. Ivanovici, B. Dobinson, N. Davies, P. Thompson*, "Network Quality of Service Measurement System for Application Requirements Evaluation", International Symposium on Performance Evaluation of Computer and Telecommunication Systems, SPECTS'03, Montreal, Canada, July 20-24, 2003, pp. 380-387.

[3] *R. Beuran, M. Ivanovici*, "User-Perceived Quality Assessment for VoIP Applications", technical report, CERN-OPEN-2004-007, January 2004.

[4] *D. Miras*, "A Survey on Network QoS Needs of Advanced Internet Applications", working document, Internet2 QoS Working Group, December 2002.

[5] *V. Cavalli, E. Verharen*, "TF-STREAM Real Time Multimedia Applications", TERENA Technical Report, March 2002.

[6] ITU-T Recommendation Y.1541. “Network Performance Objectives for IP-Based Services”, ITU, draft, October 2001.
[7] *V. Reijs*, “Perceived Quantitative Quality of Applications”, http://www.heanet.ie/Heanet/projects/nat_infrastruct/perceived.html.
[8] ITU-R Recommendation BT.500, “Subjective quality assessment methods of television pictures”, ITU, 1998.
[9] ITU-T Recommendation P.910, “Subjective Video Quality Assessment Methods for Multimedia Applications”, ITU, 1996.
[10] ITU-R Recommendation J.140, “Subjective assessment of picture quality in digital cable television systems”, ITU, 1998.
[11] ITU-T Recommendation J.143, “User requirements for objective perceptual video quality measurements in digital cable television”, ITU, 2000.
[12] *S. Winkler*, “Visual Fidelity and Perceived Quality: Towards Comprehensive Metrics”, Proc. SPIE Human Vision and Electronic Imaging, vol. 4299, pp. 114-125, San Jose, California, January, 2001.
[13] *S. Winkler*, “Issues in Vision Modelling for Perceptual Video Quality Assessment”, Signal Processing, vol. 78, no. 2, pp. 231-252, October, 1999.
[14] Video Quality Experts Group, http://www.vqeg.org
[15] VQEG, “Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment”.
[16] *C. J. van den Branden Lambrecht*, “Survey of Image and Video Quality Metrics based on Vision Models”, presentation, August, 1997.
[17] *S. Winkler, A. Sharma, D. McNally*, “Perceptual Video Quality and Blockiness Metrics for Multimedia Streaming Applications”, Proc. $4^{th}$ International Symposium on Wireless Personal Multimedia Communications, pp. 553-556, Aalbord, Denmark, September, 2001.
[18] *D. Curet, E. Gouleau, S. Relier, C. Roux, P. Clement, G. Cherry*, “RTP Payload Format for MPEG-4 FlexMultiplexed Streams”, IETF draft, July, 2002
[19] *J. van der Meer, D. Mackie, V. Swaminathan, D. Singer, P. Gentric*, “Transport of MPEG-4 Elementary Streams”, IETF draft, November, 2002.
[20] Real Networks, http://www.realnetworks.com
[21] MPEG4IP, http://mpeg4ip.sourceforge.net
[22] National Institute of Standards and Technology, NIST Net, http://snad.ncsl.nist.gov/itg/nistnet/.