Subjective Assessment of the Quality-Oriented Adaptive Scheme

Gabriel-Miro Muntean, *Member, IEEE*, Philip Perry, *Member, IEEE*, and Liam Murphy, *Member, IEEE*

***Abstract*—The Quality-Oriented Adaptation Scheme (QOAS) supports the distribution of high quality multimedia services to a large number of simultaneous customers via given broadband IP infrastructure. This paper presents subjective testing results that augment previously reported objective performance assessment. Clips representing different classes of multimedia sequences in terms of motion content and types were selected and streamed using a QOAS-based prototype system. Congested delivery network conditions were emulated and the effects of the consequent QOAS-driven adaptations were subjectively assessed by end-users. The test subjects have also graded their perceived quality when using a nonadaptive streaming approach. The QOAS-related results were much higher than those obtained for a nonadaptive approach, being above the "good" perceptual level for all multimedia clips and in all tested delivery conditions.**

***Index Terms*—Adaptive multimedia streaming, end-user perceived quality, feedback control, subjective testing.**

I. Introduction

CURRENTLY there is a trend in multimedia presentation [1] toward on-demand-based access to rich media and very high quality multimedia to home residences and business premises via an all-IP infrastructure [2], [3]. The success or failure of this approach depends on widespread market acceptance, which in turn depends on both the end-user quality of service and the price the users must pay. Network operators and service providers aim for high infrastructure utilization and a large number of customers to increase their revenues. At the same time, customers are interested in receiving high quality streamed multimedia, having access to diverse services, and paying a low cost.

The Quality Oriented Adaptation Scheme (QOAS), an application-level adaptive mechanism for multimedia streaming, was designed such that these opposing goals are balanced. QOAS offers high quality multimedia services to an increased number of simultaneous customers via a given wired IP-based infrastructure such as local broadband IP-networks. These best-effort networks would be subject to variable loads and consequently traffic through them is characterized by periods of increased and variable delays and even by loss. These may affect multimedia streaming and adaptation in general and especially QOAS-based adjustments help reduce the negative effects on the end-user perceived quality.

QOAS [4] was modeled, tested through simulations and implemented by a prototype system. QOAS performance, objectively assessed in terms of estimated end-user perceived quality, network utilization, loss rate and number of customers simultaneously served and previously reported in [5], [6], was very close to that of an ideal adaptive scheme. The results were also much better than those obtained for some well-known streaming approaches such as the adaptive solutions TFRCP [7] and LDA+ [8] or a nonadaptive approach under the same delivery conditions.

However this assessment of QOAS performance was done only in terms of some transmission-related parameters, number of customers and estimates of end-user perceived quality using an objective metric. There was a need to carry out subjective perceptual tests on the prototype system on which QOAS was deployed in order to verify the objective end-user quality results.

This paper describes the subjective tests performed in order to assess QOAS performance in terms of end-user perceived quality when streaming different types of multimedia clips and with different motion content in extreme delivery conditions. QOAS-related perceptual test results are analyzed and then compared with those obtained when a nonadaptive approach is used for streaming the same multimedia content in identical delivery conditions.

Next, some other adaptive solutions for multimedia streaming and some proposals made in order to accurately assess the end-user perceived quality are presented before QOAS is briefly outlined. Subjective testing approach, testing conditions, multimedia clips used during testing, the emulated delivery conditions and the method for test results assessment are then presented along with the analysis of tests results. The last section includes some conclusions drawn and some suggestions for further work directions.

II. Related Works

Extensive research has focused on finding solutions that provide a certain level of quality for multimedia-based services delivered over best-effort IP networks. Different solutions were proposed [9], but adaptive streaming-based schemes [10] are among the most used since they adjust the transferred multimedia data to existing delivery conditions, minimizing both the costs of their deployment and their effect on existing traffic.

At the same time significant research focused on assessing the end-user perceived quality, which is highly important when the proposed solutions for multimedia streaming have to be tested. Quantifying end-user perceived quality is also significant in the

Manuscript received May 27, 2004; revised October 28, 2004. This work was supported by the Research Innovation Fund of Enterprise Ireland.

G.-M. Muntean and P. Perry are with School of Electronic Engineering, Dublin City University, Glasnevin, Dublin 9, Ireland (e-mail: munteang@eeng.dcu.ie; perryp@eeng.dcu.ie).

L. Murphy is with the Computer Science Department, University College Dublin, Belfield, Dublin 4, Ireland (e-mail: Liam.Murphy@ucd.ie).

Digital Object Identifier 10.1109/TBC.2005.846187

effort to provide accurate information about the effects that the network conditions have on the quality of delivery to those adaptive streaming solutions such as QOAS that take this into account.

Next, different proposed solutions in these research areas (adaptive multimedia streaming and end-user perceived quality assessment) are presented.

A. Adaptive Multimedia Streaming Solutions

The adaptive streaming schemes were classified in [9] according to where the adaptation takes place: **sender-driven solutions** such as TFRCP [7], LDA+ [8], RAP [11] or LQA [12], **receiver-driven solutions** such as RLM [13] or RLC [14], and **transcoding-based solutions** (e.g. MPEG transcoders [15], filters [16]).

In [13] the authors have proposed an adaptive solution based on multicast groups that allowed the clients to directly select the desired multimedia quality in the absence of feedback. This receiver-driven adaptation scheme was later improved and the result was reported in [14]. On-the-fly transcoding is used in [15] to meet the clients' requirements, whereas [16] presents a more general transcoding-based solution that relies on filters deployed in the distribution network to match the quality level required by clients. Among the sender-driven schemes, the adaptive solution proposed in [17] varies some encoding-related parameters at the server to adjust the bit-rate of transmitted multimedia data according to feedback from clients that monitor some parameters related to multimedia transmission only, while the work in [12] describes a layered encoding-based adaptive solution.

Recently, different rate adjustment sender-driven solutions for adaptively streaming video have been proposed, such as a protocol that manages its window size in a similar manner to TCP, but does not retransmit lost packets [18]. Limitations include its inflexibility and its problems with time sensitive media. The Loss-Delay based Adjustment algorithm (LDA) [19] uses RTCP reports to estimate round trip delays and loss rates, a packet-pair technique to estimate the bottleneck link bandwidth, and some user-initialized parameters. The enhanced Loss-Delay Adaptation algorithm (LDA+) [8] also makes use of RTCP reports to collect loss and delay statistics, and adjusts the transmission rate in a TCP-like manner subject to equal losses and delays. The Rate Adaptation Protocol (RAP) [12] uses TCP-like packet acknowledgment to estimate loss rates and delays. When there is no loss, the rate is additively increased as a function of round trip delay, otherwise the rate is halved as in TCP. In [7] a TCP-Friendly Rate Control Protocol (TFRCP) is presented, based on a TCP model previously proposed in [20]. When there are losses, the rate is limited to that computed according to the TCP model, otherwise the rate is doubled. TFRCP's major problem is that it updates its rate every M time units and changes in traffic that occur on a faster scale may be taken into account too late.

More recent solutions include sender, receiver and **hybrid adaptive based mechanisms** such as [21]–[24].

The most significant criticism in relation to the previously mentioned solutions is that in spite of obtaining some significant results in terms of streaming-related metrics, there were almost no reports related to the assessment of their effect on the end-user perceived quality.

Commercial adaptive streaming solutions like Real Networks' SureStream [25] and Microsoft's Multimedia Multi-bitrate (MBR) solution [26] are proprietary and detailed technical information has never been revealed. However the available information states that they were specially designed to allow for adaptations at very low bit-rates, unlike QOAS, which addresses high quality high bit-rate video streaming.

B. End-User Perceived Quality Assessment Methods

The IETF IP Performance Metrics Working Group[1] has proposed metrics such as connectivity, one-way delay, round-trip delay, delay variation, loss rate, etc. that can be used for quantitative measurement of network performance [27]. However, they are not directly related to how the end-users perceive the quality of the service provided.

The ITU-T defined another set of parameters including interruption duration, reliability performance, maintainability, bit error ratio, etc. that can be related to end-users' QoS expectations, but they are more meaningful to the network or service provider than to the viewers [28].

As a consequence, extensive research has tried to focus on the area of end-user perceived quality assessment and two main directions were explored: **objective** and **subjective testing**.

Objective methods aim at determining the quality of a multimedia sequence in the absence of the human viewer. They are classified in [29] according to their possible usage in conjunction with adaptive streaming solutions as **out-of service methods** (the original sequence is fully available and no time constraints are imposed) and **in-service metrics** (used while streaming is in progress when the original sequence is not available and with strict timing requirements).

A classification based on the existence of the original multimedia stream [30] distinguishes three approaches: **full reference methods** (based on picture comparison), **reduced reference solutions** (that rely on feature extraction) and **no reference methods** (also called single-ended). Only the last category of methods is useful for in-service applications.

Other researchers divide the metrics associated to these objective methods into **mathematical-based** (rely on mathematical formulae or on functions based on intensive psycho-visual experiments) and **model-based** (based on complex models of the human visual system) [31], [32].

The full-reference Peak Signal to Noise Ratio (PSNR) [33] and Weighted Signal to Noise Ratio (WSNR) [33] and the no-reference Picture Appraisal Rating (PAR) [34] proposed by Snell & Wilcox[2] for MPEG-2 videos are among the mathematical metrics. Although they seem appropriate and are very simple, many studies [29], [35] have shown that PSNR and WSNR are poorly correlated to human vision.

Among the model-based metrics are PQR [36], KDD [37], DVQ [38], VQM [39], PVQM [40], PDM [41] and MPQM [42]. The Picture Quality Rating (PQR) [36] is a full reference

[1]IETF IP Performance Metrics (IPPM) Working Group (WG), http://www.ietf.org/html.charters/ippm-charter.html.

[2]Snell & Wilcox, Web Site, http://www.snellwilcox.com.

metric based on Tektronix[3]/Sarnoff[4] Human Vision Model that relies on the proprietary JNDmetrix (Just Noticeable Difference).[5] Kokusai Denshin Denwa (KDD) Research and Development Laboratories[6] proposed a proprietary full-reference model [37] based on mean square error that is weighted by a set of sequential Human Visual Filters applied at pixel, block, frame and sequence levels. The Digital Video Quality (DVQ) metric [38] and the Video Quality Model (VQM) [39], full-reference metrics proposed by NASA[7] and the Institute for Telecommunication Sciences, NTIA[8] USA respectively, are subject to U.S. patents [43], [44]. The Perceptual Video Quality Measure (PVQM) [40] uses the same approach for measuring video quality as used for speech in the Perceptual Speech Quality Measure (PSQM), standardized by the ITU-T [45]. The Perceptual Distortion Metric (PDM) [41], proposed by L'Ecole Polytechnique Federale de Lausanne (EPFL) Switzerland, is based on a spatio-temporal model of the human visual system. Although complex, this full-reference metric's greatest advantage is that details about it are made public. Researchers from EPFL have also proposed the Moving Pictures Quality Metric (MPQM) [42] which is a full reference video quality metric based on a multi-channel human visual model that takes into consideration contrast sensitivity and intra-channel masking and the no-reference MPQM (Q) [35] that describes the joint impact of MPEG rate and data loss on video quality. The latter is highly useful for in-service applications and is also used by QOAS [4].

ITU-T Video Quality Expert Group[9] has extensively studied objective metric proposals for standardization and concluded that no metric outperforms the others in all conditions. In consequence, no objective solution is currently able to fully replace subjective testing [33] which are necessary.

Subjective tests as defined by ITU-R BT.500 [46] have been used for many years in order to assess the quality of television pictures. In the area of telecommunications, five major ITU-T recommendations concern subjective testing: P.910 [47]—one way video test methods, P.911 [48]—quality assessment methods for multimedia applications, P.800 [49]—conditions for audio content testing, P.920 [50]—conversation quality assessment and P.930 [51]—video impairment reference system. Of these, the first two mostly present recommendations about methods, systems, clip contents and environment conditions for subjective testing and scales for assessing the end-user perceived quality while viewing multimedia clips.

The advantages of subjective testing include the fact that the tests can be designed to accurately represent a specific application. Also direct users' opinions are gathered and valid results can be obtained regardless of the system used, the motion content of the clips, their compression, etc. Among the drawbacks are that a wide variety of possible methods and test element parameters must be considered, complex setup and control are required, many observers must be selected and tested, and it is very time consuming and costly.

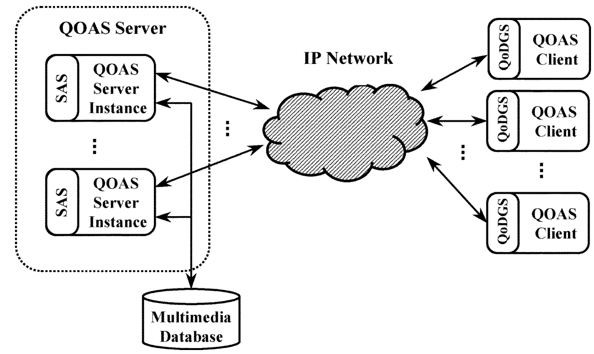


Fig. 1. The QOAS-based adaptive multimedia system architecture.

III. Overview of Quality-Oriented Adaptation Scheme (QOAS)

Many adaptive schemes presented in the previous section have shown good adaptation results in certain scenarios. However they do not involve end-user quality as perceived by the customers during streaming process in their adjustment policies. Unlike them, QOAS bases its adaptation process on estimates of the end-user perceived quality made at the receiver. This perceived quality is estimated in-service using the no-reference moving picture quality metric-Q proposed in [35] that describes the joint impact of MPEG rate and data loss on video quality. More details about Q usage as part of the QOAS are given in [4].

A. QOAS-Based System Architecture

The architecture of the prototype system that implements QOAS is presented in Fig. 1. It includes multiple instances of adaptive client and server applications that bi-directionally exchange video data and control packets through the delivery network [52], [53]. QOAS is distributed and its components are deployed at both the client side (***QOAS Client Application***) and the server side (***QOAS Server Application***). The ***Multimedia Database*** stores multimedia content to be streamed to the customers on request.

The main component of the QOAS Client Application is the ***Quality of Delivery Grading Scheme*** (QoDGS) whereas of major importance in the QOAS Server Application is the ***Server Arbitration Scheme*** (SAS). QoDGS and SAS cooperate in order to implement the QOAS feedback-controlled adaptation mechanism. Next, the roles of QoDGS and SAS are briefly indicated and the QOAS principle is presented.

B. Quality of Delivery Grading Scheme (QODGS)

The QoDGS maps some transmission related parameters values and variations and estimates of end-user perceived quality into application-level scores that describe the quality

[3]Tektronix, http://www.tek.com.

[4]Sarnoff, http://www.sarnoff.com.

[5]Just Noticeable Difference Metrics, http://www.JNDmetrix.com.

[6]Kokusai Denshin Denwa Research and Development Laboratories, KDDI Corporation, http://www.kddilabs.jp/english.

[7]The National Aeronautics and Space Administration (NASA), USA, http://www.nasa.gov.

[8]Institute for Telecommunication Sciences, National Telecommunications and Information Administration (NTIA), USA, http://www.its.bldrdoc.gov.

[9]Video Quality Experts Group (VQEG), http://www.its.bldrdoc.gov/vqeg.

of delivery. It monitors some parameters such as delay, jitter and loss rate, computes estimates of end-user perceived quality using Q and analyzes their short-term and long-term variations. Short-term monitoring is important for learning quickly about transient effects such as sudden traffic changes and for fast reacting to them. The long-term variations are monitored in order to track slow changes in the overall delivery environment such as new users in the system which need a reaction to. These short-term and long-term periods are considered, an order and two orders of magnitude greater than the feedback-reporting interval.

In the first of the QoDGS's three stages, instantaneous values of the monitored parameters are saved in different length sliding windows and their short-term and long-term variations are assessed. At the same time, session-specific lower and higher limits are maintained for each parameter, allowing for corresponding partial scores to be computed in comparison with them. In the second stage, the relative importance of all the monitored parameters in this delivery infrastructure is considered (by weighting their contributions) and the partial scores are used to compute short-term $(\mathrm{QoD_{ST}})$ and long-term $(\mathrm{QoD_{LT}})$ quality of delivery grades. This second stage also takes into account estimates for short-term and long-term end-user perceived quality. In the third stage, $\mathrm{QoD_{ST}}$ and $\mathrm{QoD_{LT}}$ are weighted to account for their relative importance and the overall client score $(\mathrm{QoD_{Score}})$ is computed.

Extensive tests were performed in order to make sure that the design of QODGS ensures that best results will be obtained in terms of adaptiveness, responsiveness to traffic variations, stability, link utilization and end-user perceived quality in local broadband IP-networks. A detailed presentation of QoDGS is made in [4].

C. Server Arbitration Scheme (SAS)

SAS takes adaptive decisions based on the values of a number of recent feedback reports, in order to minimize the effect of noise in the $\mathrm{QoD_{Score}} - \mathrm{s}$. This arbitration process is *asymmetric* requiring fewer feedback reports to trigger a decrease in quality than for a quality increase. This ensures a fast reaction during bad delivery conditions, helping to eliminate their cause and allow for the network conditions to improve before any quality upgrade. These adaptive decisions are such performed that maintain system stability by minimizing the number of quality variations. The late arrival of a number of feedback messages is considered as an indication of network congestion and triggers quality degradations. This permits the functionality of the streaming scheme even if the feedback is not available.

More details about SAS are presented in [4].

D. Principle of QOAS

Streamed multimedia data is received at the client where the QoDGS continuously monitors both some network-related parameters such as loss rate, delay and jitter and the estimated end-user perceived quality using Q. According to their values and variations, QoDGS grades the quality of delivery (QoD) in terms of application-level quality scores $(\mathrm{QoD_{Score}} - \mathrm{s})$ that are sent to the server as feedback. These scores are analyzed by the SAS that may suggest taking adaptive decisions in order

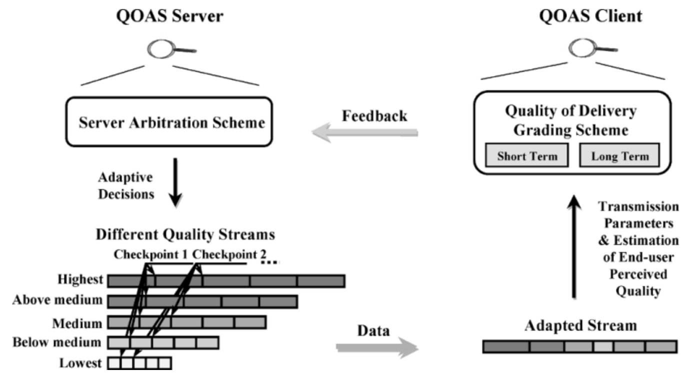


Fig. 2. QOAS adaptation principle, illustrated for QOAS-based pre-recorded multimedia streaming.

to maximize the end-user perceived quality in existing delivery conditions. These decisions affect an internal state defined for the QOAS server component that was associated with the streamed multimedia clip's quality as shown in Fig. 2. The figure presents the five-state quality model used during testing with the following states: excellent, good, average, poor and bad. Any QOAS server state modification affects the multimedia data transmission rate. For example, when increased traffic in the network affects the client-reported quality of delivery, SAS switches to a lower quality state, which determines a reduction in the quantity of data sent, helping to improve the situation. This is performed as research has shown [54] that viewers prefer a controlled reduction in multimedia quality to the effect of random losses on the streamed multimedia data. When the delivery conditions improve, the QOAS server component gradually increases the quality of the transmitted stream and in consequence the transmission rate. In the absence of loss this causes an increase in end-user perceived quality.

IV. Test Setup and Description

A. Limitations of Objective Tests

QOAS was previously tested via extensive simulations involving extremely variable and highly loaded delivery conditions and background traffic commonly expected in IP networks. This simulation-based objective testing used different types of multimedia clips, representing different classes of motion content. QOAS performance was assessed in terms of infrastructure utilization, loss rates, number of simultaneous customers served and end-user perceived quality as estimated by the no-reference Moving Pictures Quality Metric [35]. The results related to QOAS were highly positive stand alone and in comparison to other schemes and were reported in [4]–[6], [52], [53].

Although the simulations allow for fully assessing QOAS performance in terms of network-related parameters, they can only offer estimations of end-users' perceived quality with certain accuracy. Since currently there is not a broad agreement that a certain metric or a group of metrics reflects the viewers' opinions in a wide range of situations, **subjective tests** were performed in order to determine the real end-users' perceptual assessment of the quality of multimedia clips streamed using QOAS. These results were also compared with those obtained when a non-adaptive solution is used.

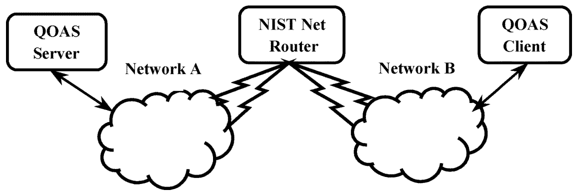


Fig. 3. Setup for subjective testing.

B. Test Setup

The test-bed presented in Fig. 3 consists of a ***QOAS Server*** machine and a ***QOAS Client*** computer, connected to two different networks interconnected by a ***NISTNet Router***. A NISTNet network Emulator [55] was installed on the Router in order to forward the packets between networks after introducing controlled bandwidth and delay constraints. The test-bed uses 100 Mbps network cards and there is no other traffic on the two networks so that the only traffic bottleneck is generated and fully controlled through the NISTNet emulation.

C. Prototype System for Multimedia Streaming

QOAS, as described in Section III, was implemented using Microsoft Visual C++ 6.0. The SAS upgrade period considered was 6 sec and the downgrade timeout 1 sec, whereas the QoDGS short-term period was set to 1 sec, and the long-term period was considered 10 sec. The QOAS Server Application adaptively streams multimedia data according to feedback reports. This feedback is sent by the QOAS Client Applications and is based on the effects the delivery network conditions have on the end-users' perceived quality. For MPEG-2 decoding these clients make use of Canopus[10] Amber MPEG decoder cards and the corresponding SDK. No error control or error concealment methods were employed.

A similar client-server implementation is used for the deployment of the ***nonadaptive solution*** (NoAd). The only difference is that NoAd streams multimedia data at the maximum bitrate available (average rate in these tests is 4 Mbps) and consequently encoded quality, regardless of the delivery conditions or eventual other problems that may affect the streaming process (e.g., loss, increased or variable delays etc.) or the end-user perceived quality.

D. Test Environment

The test environment follows the ITU-T R. P.910 [47] and ITU-T R. P.911 [48] recommendations. The display is a 19′ monitor situated in a room with no natural light. The only source of light that allows for answering the questionnaire is localized and neither reflects in the monitor nor interferes with the subjects' visual path. The monitor parameters (brightness, luminance, hue, etc.) have been set to average values. The viewing distance was within the limits suggested in [47] (5 times the picture height), and remained fixed for the duration of the tests. The audio components of the streamed multimedia clips are played out by two 10 W speakers and are not affected by QOAS adaptation. These audio components were the only source of sound in the testing room.

[10]Canopus United Kingdom, http://www.canopus-uk.com.

TABLE I
CLASSIFICATION OF CLIPS USED

Clip	Motion Content	Type
diehard1	High	Movie
dontsayaword	Average	Movie
familyman	Low	Movie
roadtoeldorado	Average	Cartoons

E. Test Considerations

In order to ensure good results, training was performed prior to starting the testing process. During the training test operators have presented to the subjects the goal of these tests and have explained what is required from them. Candidates were also screened for normal or corrected-to-normal visual acuity and for normal color vision. None of the test subjects had other visual impairments that may affect their assessment of multimedia clips quality.

Subjects' boredom and/or fatigue have an important impact on the multimedia clips quality assessment and on the accuracy of the answers. Therefore the participation to testing was voluntary, the subjects being allowed to leave at any time.

QOAS adaptation does not adjust the audio component of the multimedia stream as it takes only a small fraction of the bandwidth in comparison to the video component. Since the perceived quality of any multimedia sequence is highly influenced by the associated sound (e.g. audio and video must be appropriately synchronized), all clips were streamed with their soundtracks.

F. Multimedia Clips

Sequences from four movies representing classes of multimedia clips with different motion content and type were used for perceptual testing. The sequences, presented in Table I, were also used during simulation-based objective testing [5], [6]. For pre-recorded multimedia streaming, QOAS requires the existence of different quality versions for each sequence. In consequence the original clips were MPEG-2 encoded at five different average bitrates equally distributed between 2 Mbps and 4 Mbps using the same resolution of 320×240, frame rate (25 frames/sec) and IBBP frame pattern (9 frames/GOP) in all cases. Details related to the peak/mean rate ratio for each version of these multimedia sequences are given in Table II. For best result [30], each test lasted 1 minute, allowing the subjects both to get accustomed with the movie content and to notice quality variations.

G. Testing Method and Grading Scale

The chosen testing method is a combination between the Absolute Category Rating (ACR) and the Degradation Category Rating (DCR), presented in detail in ITU-T R P.910 [47]. ACR involves the subjects grading separately each clip but an implicit reference must be well known by all the assessors, which is not expected in this case. DCR involves showing the reference clip

TABLE II
PEAK/MEAN RATE RATIOS FOR ALL MPEG-2 ENCODED QUALITY VERSIONS OF THE CLIPS USED DURING SIMULATIONS

Clip	MPEG2 Average Encoding Rate				
	2.0 Mbps	2.5 Mbps	3.0 Mbps	3.5 Mbps	4.0 Mbps
diehard1	7.48	7.43	6.31	5.65	4.06
roadtoeldorado	6.91	6.51	6.23	6.12	6.05
dontsayaword	5.56	4.51	4.36	4.08	3.56
familyman	3.99	3.67	3.42	3.09	2.93

before each test clip, which would make the total test unfeasibly long in this case. By combining ACR and DCR, the reference clip is shown first and then the multimedia sequences that have to be assessed. The difference between the multimedia sequences is not their content, but the streaming approach used: QOAS and the nonadaptive (NoAd), respectively. The subjects are then asked to grade their subjective quality on the ITU-T R. P.910 [47] five-point scale for grading perceptual quality of video sequences. This scale spans from 1—equivalent to "bad" quality level to 5 that corresponds to "excellent" quality. No fractional grades were accepted.

The participants are also asked to indicate what characteristics related to the quality of the multimedia streaming have liked the most (e.g. continuity, audio/video synchronization, etc.) and mostly disliked respectively (e.g. tiling, jerkiness, de-synchronization, etc.). No time limit was imposed on the perceptual test subjects for this grading process.

H. Test Description

Since QOAS is likely to be deployed in delivery environments where multimedia accounts for the majority of traffic, difficult delivery conditions that involve multimedia-like background traffic were emulated for the subjective testing. In order to allow also for comparison with simulation testing results, two subjective tests were devised such that UDP-CBR background traffic was emulated using NISTNet and varied in a staircase up manner with steps of 0.4 Mbps every 20 s as shown in Fig. 4 and periodically with steps of 0.7 Mbps and on period of 30 s, as presented in Fig. 5, respectively. These steps are smaller and respectively greater than the QOAS adaptation step of 0.5 Mbps [4] (see Table II). This background traffic is emulated on top of a 95.5 Mbps constant bitrate traffic that corresponds to a well-multiplexed high load, which aims at creating high loaded delivery conditions.

The goal of the first test was to determine the effect of QOAS adaptive streaming on the end-users' perceived quality in a situation when no loss was experienced during simulations. The second test, which involved short periods of loss during simulations with QOAS, intends to determine how these expected lossy periods affect the viewers' grading of the quality. These two tests were repeated with a nonadaptive approach for streaming multimedia and their results compared with those obtained when using QOAS.

Table III and Table IV present the average bitrates and respectively the average loss rates during each of these two tests when both QOAS and NoAd streaming solutions were used with all the clips taken into consideration.

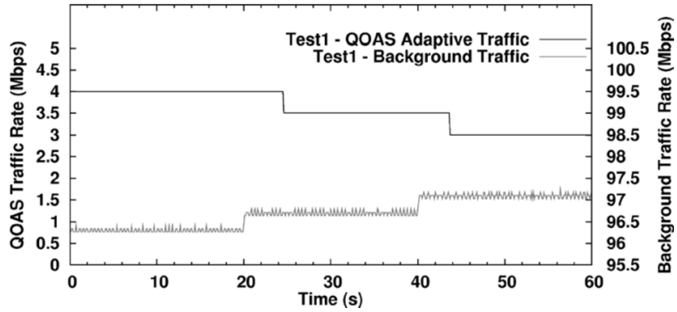


Fig. 4. Test 1: QOAS bit-rate adaptation with background traffic variation when streaming *diehard1*.

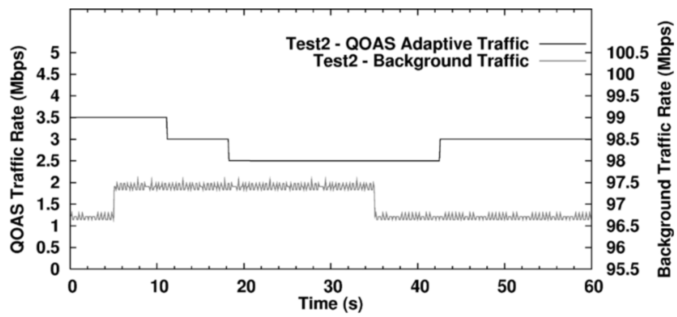


Fig. 5. Test 2: QOAS bit-rate adaptation with background traffic variation when streaming *diehard1*.

TABLE III
TEST 1—AVERAGE BITRATE AND LOSS DURING STREAMING INVOLVING QOAS AND NoAd

Scheme	QOAS		NoAd	
Clip	Avg. Rate (Mbps)	Loss Rate (%)	Avg. Rate (Mbps)	Loss Rate (%)
diehard1	3.27	0.26	4.0	1.51
dontsayaword	3.25	0.21	4.0	1.42
familyman	3.29	0.27	4.0	1.65
roadtoeldorado	3.26	0.23	4.0	1.69

V. TEST RESULTS

A. Test 1—Staircase-Up Multimedia-Like Cross Traffic

Test 1 involved 42 subjects aged between 18 and 48, with various levels of experience related to multimedia streaming (i.e. 22—familiar, 19—not familiar and 1—expert). Nineteen of the subjects were wearing glasses or contact lenses and none had other visual impairments that may affect their perception of video quality.

Each of the four selected multimedia clips was streamed in the presence of the *Test 1* emulated delivery conditions, causing adaptive rate adjustment during QOAS streaming as indicated in Fig. 4. During testing the NoAd solution, the clips were streamed at a constant bitrate of 4 Mbps. Next the subjects graded their perceived quality when using both QOAS and NoAd. The resulting average perceived quality score that corresponds to each multimedia sequence and each streaming

TABLE IV
TEST 2—AVERAGE BITRATE AND LOSS DURING STREAMING INVOLVING QOAS AND NoAd

Scheme	QOAS		NoAd	
Clip	Avg. Rate (Mbps)	Loss Rate (%)	Avg. Rate (Mbps)	Loss Rate (%)
diehard1	2.89	0.33	4.0	2.43
dontsayaword	2.88	0.32	4.0	2.21
familyman	2.91	0.29	4.0	2.52
roadtoeldorado	2.87	0.28	4.0	2.26

TABLE V
TEST 1—SUBJECTIVE TESTING RESULTS FOR QOAS AND NoAd AVERAGE USER PERCEIVED QUALITY ON 1–5 ITU-T R. P.910 SCALE

Scheme	QOAS		NoAd	
Clip	Grade (1-5)	St. Dev.	Grade (1-5)	St. Dev.
diehard1	4.00	0.71	2.02	0.72
dontsayaword	4.18	0.75	2.44	1.00
familyman	4.21	0.83	1.85	0.79
roadtoeldorado	3.74	0.71	1.93	0.72

approach is presented in Table V as well as the standard deviation of the results from the average.

These results indicate that QOAS streaming was very appreciated by the subjects and scored above 4, the ITU-T "good" quality level, for all the movies and very close to 4 for the cartoon sequence. The low standard deviation values indicate that the results obtained are consistent, despite the coarse granularity of the grading process (i.e. fractional grades were not accepted). The QOAS results are then statistically compared to the NoAd results whose average subjective quality scores are below the "poor" level. The t-tests performed on the user perceived quality scores given by the subjects to each multimedia clip have confirmed that there is very significant statistical difference between QOAS and NoAd in favor of the former (significance level $\alpha = 0.01$). This result is valid regardless of the clips motion content and type.

B. Test 2—Periodic Multimedia-Like Background Traffic

Test 2 involved 42 subjects with ages between 21 and 45, with various levels of experience related to multimedia streaming (i.e. 19—familiar, 21—not familiar and 2—experts). This time, 16 subjects were wearing corrective glasses or contact lenses. Table VI presents the results of the subjective grading of the multimedia streaming quality when QOAS and NoAd approaches were used respectively. QOAS has scored much above 4, the "good" perceptual quality level for all the movies and slightly below 4 for the cartoon sequence. Similar with the first test, the results' standard deviations have low values showing their consistency. It is significant to notice that the short lossy periods that have occurred during *Test 2* before the scheme has performed its feedback-based adjustments have not significantly influenced the perceived quality of the subjective

TABLE VI
TEST 2—SUBJECTIVE TESTING RESULTS FOR QOAS AND NoAd AVERAGE USER PERCEIVED QUALITY ON 1–5 ITU-T R. P.910 SCALE

Scheme	QOAS		NoAd	
Clip	Grade (1-5)	St. Dev.	Grade (1-5)	St. Dev.
diehard1	4.22	0.69	1.33	0.67
dontsayaword	3.98	0.64	1.45	0.67
familyman	4.24	0.66	1.31	0.56
roadtoeldorado	3.85	0.69	1.37	0.62

results when using QOAS-based adaptive streaming. In comparison the increased traffic delivery conditions have severely affected the end-user perceived quality for NoAd streaming whose scores are just above the minimum perceptual level. As in *Test 1* the t-tests performed on the scores given by the test subjects for each multimedia clip have confirmed a very significant statistical difference between QOAS and NoAd, with the results in favor of the QOAS (significance level $\alpha = 0.01$).

C. Influence of Video Motion Content and Loss

The subjective testing results of *Test1* for QOAS presented in Table V suggest that the higher the motion complexity of a sequence the lower the subjective appreciation in loaded delivery conditions is. That is, network congestion appears to have a more noticeable effect on clips with high motion content. This observation is supported by an ANOVA test, which indicated that the results are significantly different $(p < 0.05)$ from a statistical point of view. When the NoAd approach was used, high packet losses occurred and their effect on the subjects perceived quality of the streamed multimedia clips was unrelated to the motion content or type of these sequences.

During *Test 2* when the emulated delivery conditions triggered some loss even when QOAS was used for streaming, the viewers' perceived quality was also affected independent from the motion content as shown in Table VI. This finding is supported by an ANOVA test that has found the results obtained for tested clips significantly different $(p < 0.05)$.

In conclusion the motion content of the movie clips directly influences the perceptual quality assessment of the streamed multimedia clips at low loss rates. However, when the loss rate increases and begins to severely affect the streaming process, this relationship between the motion content and the end-user perceived quality weakens.

D. Influence of Increased Loaded Delivery Conditions

Although the results of subjective *Test 2* when using QOAS seem higher than those of *Test 1*, by performing t-tests on the perceptual quality scores obtained for these two tests for each multimedia sequence involved in testing, the null hypothesis that there is no statistical difference between the mean values of the scores obtained for *Test 1* and *Test 2* respectively cannot be rejected. This finding is stated with a very high level of confidence of 99% (significance level $\alpha = 0.01$). This indicates a very significant stability of the QOAS that maintains the same end-user appreciation in relation to the perceived quality in various loaded delivery conditions.

However when similar t-tests were performed for the NoAd approach, there was a very significant statistical difference between the results in favor of those obtained during *Test 1* (significance level $\alpha = 0.01$). Although it was expected in NoAd case that higher loaded delivery conditions would cause higher loss and therefore would affect negatively the end-user perceived quality, this result confirms the QOAS benefit in difficult delivery conditions.

E. Influence of Multimedia Content Type

Analyzing *Test 1* and *Test 2* results when using QOAS, there was a very significant statistical difference between the subjective scores obtained for the clips that contain movie scenes and those associated with the cartoon clip. This result was confirmed by paired t-tests that were performed on results obtained for each movie sequence and the cartoon sequence with a significance level of $\alpha = 0.01$. A potential cause might be the different MPEG-2 encoding output for the cartoon sequences as shown in Table II. Unlike for the movie content, the peak/mean ratio for cartoon content does not significantly increase with the decrease in the average encoding bit-rate. Also the content with many colors and clearly defined edges might be more affected in terms of the end-user subjective quality corrupted during streaming and especially in situations when loss occurs.

However these findings were not confirmed for the case when NoAd approach was used for streaming in these difficult delivery conditions that have triggered high loss. Paired t-tests that were performed on results obtained for each movie sequence and the cartoon sequence could not reject the null hypothesis that there is no difference between the mean values of those sets of results (significance level of $\alpha = 0.05$ for *Test 1* and $\alpha = 0.01$ for *Test 2*).

In conclusion the type of multimedia content significantly influences the end-user perceptual quality of multimedia clips streamed using QOAS. When using the NoAd approach, much higher losses occur that more severely affect the streaming process and thus the end-user perceived quality. In this context no relationship between the type of streamed multimedia clips and end-user quality assessment could be demonstrated.

VI. Perceptual Assessment of Various Multimedia-Related Features

A. Most Appreciated Multimedia Features

During both *Test 1* and *Test 2*, the subjects were also asked to indicate the features related to streamed multimedia clips that they have mostly appreciated and the results are summarized in Table VII and Table VIII. The numbers in the tables are the percentage of subjects that thought that the feature helped their perception of the clip.

When streaming multimedia sequences using QOAS the most appreciated features were the *clarity* of the video content—most appreciated by on average of 70% of subjects during both tests and *continuity* of streaming—mostly liked on average by 60% of subjects. In contrast when using NoAd the *clarity* of the streamed multimedia content was the most appreciated feature and has attracted just over 30% of the votes during *Test1* and less than 10% during *Test2*. At the same time NoAd streaming *continuity* was appreciated by just over 5% of the subjects.

TABLE VII
Test 1—Statistical Results: the Subjects Most Appreciated Characteristics When Streaming: (A) *diehard1*, (B) *dontsayaword*, (C) *familyman*, and (D) *roadtoeldorado*

(%)	QOAS				Non-Adaptive			
Clip	A	B	C	D	A	B	C	D
Continuity	52.4	66.7	69.0	47.6	4.8	16.7	9.5	4.8
Quality Stability	45.2	64.3	52.4	35.7	4.8	11.9	4.8	4.8
Clarity	78.6	78.6	71.4	69.0	35.7	38.1	31.0	28.6
Media Synch.	54.8	52.4	59.5	28.6	14.3	31.0	9.5	9.5

TABLE VIII
Test 2—Statistical Results: the Subjects Most Appreciated Characteristics When Streaming: (A) *diehard1*, (B) *dontsayaword*, (C) *familyman*, and (D) *roadtoeldorado*

(%)	QOAS				Non-Adaptive			
Clip	A	B	C	D	A	B	C	D
Continuity	76.2	61.9	71.4	45.2	7.1	0.0	0.0	0.0
Quality Stability	57.1	47.6	54.8	45.2	7.1	0.0	0.0	2.4
Clarity	66.7	66.7	71.4	66.7	9.5	11.9	7.1	9.5
Media Synch.	66.7	52.4	64.3	38.1	9.5	0.0	0.0	0.0

Media synchronization and *quality stability* have attracted on average the appreciation of more than 50% of subjects in both tests when QOAS was used for streaming. When delivering multimedia using the NoAd approach, *media synchronization* was appreciated by around 10% of the subjects during *Test 1* and around 2% during *Test 2*, whereas the *quality stability* was indicated by less than 2% of viewers.

The results related to quality stability are very significant for QOAS evaluation as they confirm that the quality variations introduced by QOAS adaptation process do not affect negatively the end-users perceived quality. Also the fact that the levels of appreciation for all these features are significantly higher for QOAS than for NoAd indicates the important benefit of QOAS when streaming multimedia in terms of end-user perceived quality.

Analyzing these subjective results, although there is a certain degree of variation between the viewers appreciation of certain features, the results could not be related to the clips' motion content. At the same time the appreciation of all the features in relation to the cartoon sequence was much lower than for any other movie clip, fact that may suggest a dependence on the type of the encoded content.

B. Least Appreciated Multimedia Features

The subjects were asked to indicate the multimedia features they have disliked the most during *Test 1* and *Test 2* respectively and the results related to each streamed multimedia clip are presented in Table IX and Table X.

TABLE IX
TEST 1—STATISTICAL RESULTS: THE CHARACTERISTICS MOST DISLIKED BY TEST SUBJECTS WHEN STREAMING: (A) *diehard1*, (B) *dontsayaword*, (C) *familyman*, AND (D) *roadtoeldorado*

(%)	QOAS				Non-Adaptive			
Clip	A	B	C	D	A	B	C	D
Jerkiness	28.6	33.3	31.0	45.2	69.0	59.5	88.1	78.6
Quality Variation	26.2	14.3	14.3	28.6	71.4	64.3	73.8	71.4
Blurring	2.4	0.0	2.4	9.5	16.7	14.3	16.7	21.4
Tiling	9.5	11.9	23.8	21.4	69.0	61.9	78.6	69.0
Media De-synch.	14.3	19.0	9.5	45.2	59.5	47.6	66.7	59.5

TABLE X
TEST 2—STATISTICAL RESULTS: THE CHARACTERISTICS MOST DISLIKED BY TEST SUBJECTS WHEN STREAMING: (A) *diehard1*, (B) *dontsayaword*, (C) *familyman*, AND (D) *roadtoeldorado*

(%)	QOAS				Non-Adaptive			
Clip	A	B	C	D	A	B	C	D
Jerkiness	14.3	31.0	9.5	28.6	83.3	90.5	78.6	83.3
Quality Variation	16.7	16.7	23.8	21.4	66.7	64.3	57.1	64.3
Blurring	0.0	7.1	2.4	2.4	21.4	19.0	16.7	19.0
Tiling	26.2	28.6	23.8	38.1	88.1	85.7	90.5	78.6
Media De-synch.	9.5	16.7	4.8	31.0	76.2	81.0	83.3	76.2

When delivering multimedia content using QOAS, features such as *blurring* and *tiling* have been seldom indicated as a cause of dissatisfaction with the multimedia service. However on average 27% of subjects have found *jerkiness* the most disturbing aspect, around 20% of them have disliked *quality variations* the most and *media de-synchronization* was indicated by less than 17% of viewers.

Although these are low figures for the extreme loaded delivery conditions emulated during *Test1* and especially during *Test2*, they are due to the natural QOAS-based quality variations, low but nonzero loss rate recorded and some implementation imperfections of the prototype system that deploys QOAS that may contribute toward a degradation of the end-user perceived quality.

However when NoAd was used for streaming in the same loaded delivery conditions, on average more than 80% of the subjects were annoyed by *tiling* and more than 75% by *jerkiness*. At the same time more than 65% of viewers have noticed and have disliked *media de-synchronizations*, whereas *quality variations* were have mostly disturbed more than 60% of the subjects.

The average of 60 points difference between the figures associated with QOAS and NoAd respectively indicate the huge benefit of the adaptive approach in comparison with the non-adaptive solution. It is very significant to notice that QOAS has very significantly reduced the effect of *tiling*—usually associated with loss and most disturbing for viewers and very much lowered the negative effect of other perceptual-related features related to multimedia streaming.

Analyzing the results it could be said that *blurring* had the least effect on the subjects' perceived quality, regardless of the streaming method. At the same time the influence of the motion content of the multimedia sequences that originate from movies on the subjects dissatisfaction with certain features could not be demonstrated. However the type of the multimedia sequence seems to be significant, consistent more negative results being obtained for the cartoon sequence, regardless of the streaming solution used.

VII. CONCLUSIONS

The Quality Oriented Adaptation Scheme (QOAS) was designed in order to ensure both higher quality and increased efficiency while delivering high bit-rate multimedia streams to users via broadband IP networks. QOAS balances the customers' desire for high quality with the network operators' and service providers' aim at achieving high infrastructure utilization and serving more customers.

Extensive simulation-based tests have assessed QOAS under a range of different network conditions and their results have shown very good QOAS performance. The subjective tests presented in this paper, complement those objective tests and aim at assessing the end-users perceived quality when using QOAS for multimedia streaming in most difficult delivery conditions selected from the simulation tests.

Multimedia sequences representing different classes of multimedia clips in terms of motion content and type were selected and streamed using a prototype system that deploys QOAS. Extremely difficult delivery network conditions were emulated and the effect of the consequent adaptations performed by QOAS was subjectively assessed by end-users.

Regardless of the motion content of the streamed clips, when using the QOAS approach the multimedia streaming quality was highly graded by the test subjects, being above the "good" perceptual level defined in the ITU-T R. P.910. Slightly lower results were obtained when streaming cartoon content, in both cases when difficult delivery conditions were emulated. This indicates that multimedia type significantly influences the end-user subjective quality assessment.

The test subjects have also graded their perceived quality when using a nonadaptive streaming approach in identical delivery conditions and as expected the QOAS-related results were much higher than those obtained for the nonadaptive approach.

Although it is generally accepted that packet loss affects the end-user perceived quality, this papers reports results of a study that considers its influence on the subjective quality scores in conjunction with the effect of streaming adaptiveness and the clips' degree of motion content. It is significant to mention the observation that the higher the loss rate, the lower the influence of the motion content on the end-user perceived quality.

The test subjects were also asked to indicate the characteristic related to multimedia streaming that was mostly appreciated and disliked respectively. They have specified these features in relation to each clip and both QOAS and nonadaptive approach.

These subjective testing results, confirm the already existing objective test results and highly recommend QOAS not only as a very efficient solution for delivering multimedia-based services to remote viewers but also as one that ensures very good quality even in most variable and loaded network conditions.

VIII. Further Work

Subjective tests to compare QOAS-based streaming to other adaptive solutions are in progress.

Further tests will assess the QOAS's degree of TCP friendliness against different traffic that consists of various flavors of TCP. Also QOAS will be extended for multicast transmissions, taking into account some multicast specific characteristics such as multiple feedback and arbitration of heterogeneous client reports.

Further work that explores the possibility of using QOAS with MPEG-4 in situations where broadband networks are not available is envisaged. Extensive objective and subjective testing that take into account the possible different effects the coding scheme and narrower bandwidth have on the end-user perceived quality will follow, so that the performance of this QOAS extension is assessed properly.

Acknowledgment

The authors thank Dr. J. Murphy and Mrs. C. Hava Muntean for their valuable support during this research.

References

[1] D. Wu, Y. T. Hou, W. Zhu, Y.-Q. Zhang, and J. M. Peha, "Streaming video over the Internet: approaches and directions," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 3, pp. 282–300, Mar. 2001.

[2] S. Dravida, D. Gupta, S. Nanda, K. Rege, J. Strombosky, and M. Tandon, "Broadband access over cable for next-generation services: a distributed switch architecture," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 116–124, Aug. 2002.

[3] E. W. M. Wong and S. C. H. Chan, "Performance modeling of video-on-demand systems in broadband networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 7, Jul. 2001.

[4] G.-M. Muntean, P. Perry, and L. Murphy, "A new adaptive multimedia streaming system for All-IP multi-service networks," *IEEE Trans. Broadcast.*, vol. 50, no. 1, pp. 1–10, Mar. 2004.

[5] ——, "Performance comparison of local area video streaming systems," *IEEE Commun. Lett.*, vol. 8, no. 5, pp. 326–328, May 2004.

[6] ——, "A quality-orientated adaptation scheme for video-on-demand," *IEE Electron. Lett.*, vol. 39, no. 23, pp. 1689–1690, Nov. 2003.

[7] J. Padhye, J. Kurose, D. Towsley, and R. Koodli, "A model based TCP friendly rate control protocol," in *Proc. ACM NOSSDAV*, Basking Ridge, NJ, USA, Jun. 1999, pp. 137–151.

[8] D. Sisalem and A. Wolisz, "LDA+ TCP-friendly adaptation: a measurement and comparison study," in *Proc. ACM NOSSDAV*, Chapel Hill, NC, USA, Jun. 2000.

[9] D. Wu, Y. T. Hou, W. Zhu, and Y.-Q. Zhang, "Transporting real-time video over the Internet: challenges and approaches," *Proc. IEEE*, vol. 88, no. 12, pp. 1855–1875, Dec. 2000.

[10] X. Wang and H. Schulzrinne, "Comparison of adaptive Internet multimedia applications," *IEICE Trans. Commun.*, vol. E82-B, no. 6, pp. 806–818, Jun. 1999.

[11] R. Rejaie, M. Handley, and D. Estrin, "RAP: an end-to-end rate-based congestion control mechanism for real time streams in the Internet," in *Proc. IEEE INFOCOM*, New York, USA, Mar. 1999, pp. 1337–1345.

[12] ——, "Layered quality adaptation for Internet video streaming," *IEEE J. Select. Areas Commun. (JSAC), Special Issue on Internet QoS*, vol. 18, no. 12, pp. 2530–2543, Dec. 2000.

[13] S. McCanne, V. Jacobson, and M. Vetterli, "Receiver-driven layered multicast," in *Proc. ACM SIGCOMM*, Stanford, California, USA, August 1996, pp. 117–130.

[14] L. Vicisano, J. Crowcroft, and L. Rizzo, "TCP-like congestion control for layered multicast data transfer," in *Proc. IEEE INFOCOM*, vol. 3, San Francisco, California, Mar. 1998, pp. 996–1003.

[15] L. Wang, A. Luthra, and B. Eifrig, "Rate control for MPEG transcoders," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 2, pp. 222–234, Feb. 2001.

[16] N. Yeadon, F. Garcia, D. Hutchison, and D. Shepherd, "Filters: QoS support mechanisms for multipeer communications," *IEEE J. Select. Areas Commun. (JSAC)*, vol. 14, no. 7, pp. 1245–1262, Sept. 1996.

[17] J. C. Bolot and T. Turletti, "Adaptive error control for packet video in the Internet," in *Proc. Int. Conf. Image Processing*, Lausanne, Switzerland, Sep. 1996, pp. 25–28.

[18] S. Jacobs and A. Eleftheriadis, "Providing video services over networks without quality of service guarantees," in *WWW Consortium Workshop on Real-Time Multimedia and Web*, Sophia Antipolis, France, Oct. 1996.

[19] D. Sisalem and H. Schulzrinne, "The loss-delay based adjustment algorithm: a TCP-friendly adaptation scheme," in *Proc. NOSSDAV*, Cambridge, UK, Jul. 1998, pp. 215–226.

[20] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP throughput: a simple model and its empirical validation," in *Proc. ACM SIGCOMM*, Vancouver, Canada, Oct. 1998, pp. 303–314.

[21] S. Cho, H. Woo, and J. Lee, "ATFRC: adaptive TCP friendly rate control protocol," in *Lecture Notes in Computer Science*. Germany: Springer-Verlag Heidelberg, 2003.

[22] C. Krasic, J. Walpole, and W.-C. Feng, "Quality-adaptive media streaming by priority drop," in *Proc. ACM NOSSDAV*, Monterey, California, USA, Jun. 2003, pp. 112–121.

[23] J. Byers, M. Frumin, G. Horn, M. Luby, M. Mitzenmacher, A. Roetter, and W. Shaver, "FLID-DL: congestion control for layered multicast," in *Proc. Int. Workshop on Networked Group Communication*, Palo Alto, California, USA, Nov. 2000, pp. 71–82.

[24] I. Rhee, V. Ozdemir, and Y. Yi. (2000, Apr.) TEAR: TCP Emulation at Receivers—Flow Control for Multimedia Streaming. Tech. Rep., Department of Computer Science, North Carolina State University, USA. [Online] Available: http://www.csc.ncsu.edu/faculty/rhee/export/tear_page/

[25] RealNetworks. SureStream. [Online] Available: http://www.realnetworks.com

[26] Microsoft, "Windows Media, MBR ,", Available: http://www.microsoft.com.

[27] V. Paxson, G. Almes, J. Mahdavi, and M. Mathis. (1998, May) Framework for IP Performance Metrics. *RFC 2330* [Online] Available: http://www.faqs.org/rfcs/rfc2330.html

[28] "Terms and Definitions Related to Quality of Service and Network Performance Including Dependability,", ITU-T Recommendation E.800, Aug. 1994.

[29] S. Winkler, A. Sharma, and D. McNally, "Perceptual video quality and blockiness metrics for multimedia streaming applications," in *Proc. Int. Symp. Wireless Personal Multimedia Communications*, Aalborg, Denmark, Sep. 2001, pp. 553–556.

[30] D. Miras, Network QoS needs of advanced Internet applications—a survey, in Internet2 QoS Working Group, 2002.

[31] C. J. van den Branden Lambrecht, "Perceptual Models and Architectures for Video Coding Applications," Ph.D. Thesis, L'Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland, 1996.

[32] P. Frossard, "Robust and Multiresolution Video Delivery: From H.26x to Matching Pursuit Based Technologies," Ph.D. Thesis, L'Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland, 2001.

[33] Video Quality Experts Group (VQEG). (2000, Apr.) Final Report. [Online] Available: http://www.its.bldrdoc.gov/vqeg/pdf/final_report_april00.pdf

[34] M. Knee. (2000, Jan.) The Picture Appraisal Rating (PAR)—a single-ended picture quality measure for MPEG-2, White Paper. *Snell & Wilcox* [Online] Available: http://www.snellwilcox.com/products/mosalina/content/downloads/parpaper.pdf

[35] O. Verscheure, P. Frossard, and M. Hamdi, "User-oriented QoS analysis in MPEG-2 video delivery," *J. Real-Time Imaging*, vol. 5, no. 5, pp. 305–314, Oct. 1999.

[36] (2000) A guide to maintaining video quality of service for digital television programs, White Paper. *Tektronix* [Online] Available: http://www.broadcastpapers.com/tvtran/25W_14 000_0.pdf

[37] Pixelmetrix and KDD media to jointly market VP series picture quality analyzer. Pixelmetrix Press Release. [Online] Available: http://www.pixelmetrix.com/rel/press%20release/1kdd.pdf

[38] A. B. Watson, J. Hu, and J. F.J. F. McGowan III, "Digital video quality metric based on human vision," *J. Electronic Imaging*, vol. 10, no. 1, pp. 20–29, 2001.

[39] A. A. Webster *et al.*, "An objective video quality assessment system based on human perception," in *SPIE Human Vision, Visual Processing, and Digital Display IV*, vol. 1913, San Jose, USA, February 1993, pp. 15–26.

[40] A. P. Hekstra *et al.*, "PVQM—a perceptual video quality measure," *J. Signal Processing: Image Communication*, vol. 17, no. 10, pp. 781–798, 2002.

[41] S. Winkler, "A perceptual distortion metric for digital color video," in *Proc. Human Vision and Electronic Imaging SPIE*, vol. 3644, San Jose, USA, Jan. 1999.

[42] C. J. van den Branden Lambrecht and O. Verscheure, "Perceptual quality measure using a spatio-temporal model of the human visual system," in *Proc. SPIE*, vol. 2668, San Jose, USA, Feb. 1996, pp. 450–461.

[43] A. B. Watson, "Method and Apparatus for Evaluating the Visual Quality of Processed Digital Video Sequences," U.S. Patent no. 6 493 023, Dec. 2002.

[44] S. Wolf and M. H. Pinson, "In-Service Video Quality Measurement System Utilizing an Arbitrary Bandwidth Ancillary Data Channel," U.S. Patent no. 6 496 221, Dec. 2002.

[45] "Objective Quality Measurement of Telephone-band (300–3400 Hz) Speech Codecs,", ITU-T Recommendation P.861, Feb. 1996.

[46] "Methodology for the Subjective Assessment of the Quality of Television Pictures,", ITU-R BT.500, 1974–2002.

[47] "Subjective Video Quality Assessment Methods for Multimedia Applications,", ITU-T Recommendation P.910, Sep. 1999.

[48] "Subjective Audiovisual Quality Assessment Methods for Multimedia Applications,", ITU-T Recommendation P.911, 1998.

[49] "Methods for Subjective Determination of Transmission Quality,", ITU-T Recommendation P.800, Aug. 1996.

[50] "Interactive Test Methods for Audiovisual Communications,", ITU-T Recommendation P.920, Aug. 1996.

[51] "Principles of a Reference Impairment System for Video,", ITU-T Recommendation P.930, Aug. 1996.

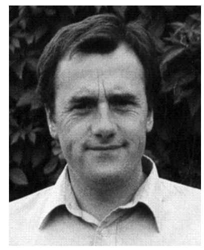
[52] G.-M. Muntean and L. Murphy, "Adaptive pre-recorded multimedia streaming," in *Proc. IEEE GLOBECOM'2002*, Taipei, Taiwan, Nov. 2002, pp. 1728–1732.

[53] ——, "An adaptive mechanism for pre-recorded multimedia streaming based on traffic conditions," in *Proc. 11th W3C World Wide Web Conf.*, Honolulu, Hawaii, USA, May 2002.

[54] G. Ghinea and J. P. Thomas, "QoS impact on user perception and understanding of multimedia video clips," in *Proc. ACM Multimedia*, Bristol, U.K., 1998, pp. 49–54.

[55] NIST Net. [Online] Available: http://snad.ncsl.nist.gov/itg/nistnet

Gabriel-Miro Muntean is a Lecturer with the School of Electronic Engineering, Dublin City University, Ireland, where he obtained the Ph.D. degree in 2003 for research on quality-oriented adaptation schemes for multimedia streaming. He was awarded the B.Eng. and M.Sc. degrees in Software Engineering from the Computer Science Department, "Politehnica" University of Timisoara, Romania in 1996 and 1997, respectively. Dr. Muntean's research interests include QoS and performance-related issues of adaptive solutions for multimedia delivery over wired and wireless networks. He is a Member of the IEEE.



Philip Perry (M'92) is a Senior Research Fellow in the Performance Engineering Laboratory, with responsibilities in both the Department of Computer Science at University College Dublin, Dublin, Ireland, and the School of Electronic Engineering at Dublin City University, Dublin, Ireland. He obtained the Ph.D. degree in microwave engineering from the Department of Electronics and Electrical Engineering at University College Dublin, Dublin, Ireland in 1998. He studied for his Master's degree at the University of Bradford, Yorkshire, England (1989) while his primary degree is from the University of Strathclyde, Glasgow, Scotland (1987). His current research interests are focused on time constrained applications and enabling technologies for mobile systems.

Liam Murphy obtained the B.E. degree in electrical engineering from University College Dublin in 1985, and the Masters and Ph.D. degrees in electrical engineering and computer sciences from the University of California, Berkeley, in 1988 and 1992, respectively. He is currently a Senior Lecturer in Computer Science at University College Dublin, where he is the Director of UCD's Performance Engineering Laboratory. His current research interests include performance issues in multimedia transmission, Voice over IP, and component oriented software systems. Dr. Murphy is a Member of the IEEE.