Objective Assessment of Region of Interest-Aware Adaptive Multimedia Streaming Quality

Bogdan Ciubotaru, *Student Member, IEEE*, Gabriel-Miro Muntean, *Member, IEEE*, and Gheorghita Ghinea, *Member, IEEE*

Abstract—**Adaptive multimedia streaming relies on controlled adjustment of content bitrate and consequent video quality variation in order to meet the bandwidth constraints of the communication link used for content delivery to the end-user. The values of the easy to measure network-related Quality of Service metrics have no direct relationship with the way moving images are perceived by the human viewer. Consequently variations in the video stream bitrate are not clearly linked to similar variation in the user perceived quality. This is especially true if some human visual system-based adaptation techniques are employed. As research has shown, there are certain image regions in each frame of a video sequence on which the users are more interested than in the others. This paper presents the Region of Interest-based Adaptive Scheme (ROIAS) which adjusts differently the regions within each frame of the streamed multimedia content based on the user interest in them. ROIAS is presented and discussed in terms of the adjustment algorithms employed and their impact on the human perceived video quality. Comparisons with existing approaches, including a constant quality adaptation scheme across the whole frame area, are performed employing two objective metrics which estimate user perceived video quality.**

Index Terms—**Content adaptation, multimedia streaming, region of interest, user perceived quality.**

Fig. 1. Typical architecture for adaptive multimedia streaming.

I. INTRODUCTION

LATELY, multimedia applications including IP television (IP TV), voice-over-IP (VoIP), video-on-demand, distance learning and teleconferencing, have become increasingly popular among the Internet users. These applications require medium to high bandwidth for high quality streaming and are very sensitive to variation in Quality of Service (QoS) parameters such as loss, delay and delay jitter [1]. As the current best-effort Internet cannot guarantee any levels for these QoS parameters [2], which often experience extreme values and variations depending on network conditions, most multimedia streaming applications are negatively affected in their viewer perceived quality.

At the same time there is an increasing popularity in the usage of mobile device, such as laptops, PDAs, or smart phones, equipped with various wireless communication solutions (e.g. WiFi, WiMax, cellular) [3]. Although wireless technologies aim to support similar bandwidth and QoS levels as the wired communication solutions [4], currently this is far from being achieved. Consequently, along with the problems associated with wired solutions, delivering multimedia content over wireless networks (and often to mobile users) involves several additional performance issues. These issues are mainly related to device and user mobility and the loss-prone nature of the wireless medium [5].

However, in order to become commercially viable, multimedia applications have to maximize their users' quality of experience (QoE). Consequently various solutions to increase user QoE when streaming high bitrate multimedia content are required.

Research [6] has shown that one of the most efficient solutions to optimize user QoE when streaming multimedia over heterogeneous networks with variable load is to adapt the multimedia content bitrate based on networks' current conditions. Adaptive multimedia streaming also allows for using efficiently the available bandwidth of the communication channel and for avoiding network overload.

There are many multimedia streaming solutions [6] which use a typical architecture such as the one illustrated in Fig. 1. They base their adaptation on the values of some network level QoS metrics which are poorly correlated with the human quality perception. The success of some recent adaptive schemes such as QOAS [7] which uses an estimation of user perceived quality in the adaptation process suggests that the results of the research in the area of user QoE which consider the characteristics of the human visual system [8] should be taken into account.

Existing adaptive solutions affect equally the whole viewing area of the multimedia frames in the content bitrate adjustment process. However, research has shown that there are certain regions in each frame of any video sequence on which the users are more interested than in the others [9].

This paper presents the Region Of Interest-based Adaptive Scheme (ROIAS) [10] as a novel bitrate adaptation technique based on certain characteristics of human visual system, which adjusts differently the regions within each frame of the streamed multimedia content based on the user interest in them. Two versions of ROIAS are presented and discussed in terms of the adjustment algorithms employed and their impact on the human perceived video quality. Objective testing results show how both ROIAS versions provide higher user perceived quality for the video stream than the solutions that reduce the bitrate by degrading equally the whole image area.

The structure of the paper is as follows: Section II presents some existing adaptation schemes for multimedia streaming along with results of research in relation to region of interest encoding techniques and video quality assessment. ROIAS is presented in Section III which also details its architecture, the feedback-based adaptive multimedia streaming technique it uses and two proposed alternative versions of the scheme. Section IV details the simulation environment and scenario and presents objective testing results and result analysis. At the end conclusions and possibilities for further work are described.

II. RELATED WORK

A. Adaptive Multimedia Streaming

There are many performance issues when delivering multimedia over variously loaded best-effort networks to heterogeneous users in terms of connectivity, device characteristics and expectations. Among the most significant causes of degradation of the performance when streaming multimedia are the low bandwidth available, lossy connections, highly variable background traffic and highly loaded delivery conditions. The combined effect of these factors ultimately affects end-user Quality of Experience (QoE).

As QoE is difficult to assess, research has focused on proposing techniques to increase Quality of Service (QoS) level. Various solutions were proposed from bandwidth over-provisioning to traffic engineering, but they were very expensive, difficult to deploy or provided limited scalability and flexibility. Among the most successful solutions are the adaptive multimedia streaming schemes which adjust the bandwidth used by the applications according to the existing network conditions, increasing or decreasing both the transmission and content encoding rates.

Among the approaches proposed, network or transport level adaptive solutions such as TFRCP [11], LDA+ [12] and RAP [13] provided certain level of QoS when streaming multimedia over wired networks, but were poorly linked to end-user perceived quality. Application layer solutions such as LQA [14] and cross layer methods as presented in [15] get closer to the users and try to achieve higher end-user perceived quality of multimedia streaming. However only the Quality Oriented Adaptation Scheme (QOAS) [7] involves a user perceived quality estimation in the feedback-based multimedia adaptation process. Consequently QOAS shows significant improvements in end-user QoE when used for adaptively streaming multimedia in both wired and wireless environments.

More recently, diverse solutions were proposed for adaptive multimedia transmissions over wireless access networks [16] or wireless ad-hoc networks [17]. Among the proposed solutions are adaptation mechanisms at the level of layers [16] or objects [18], fine-granular scalability schemes [19] and perception-based approaches [20]. Complementing these approaches the IEEE 802.11e standard [21] provides QoS features that may help improving users QoE allowing for multiple-priority-based distribution of multimedia content.

However all these solutions involve content adjustments which affect equally the whole area of the video frames, regardless of different user interest in various frame regions as research on regions on interest (ROI) has demonstrated [9].

B. Region of Interest

There has been considerable interest in Region of Interest research, primarily based on the premise that where a user's gaze rests corresponds to the location of the symbol currently being processed in working memory. Consequently, the idea is to allocate screen real estate preferentially, with more resources being earmarked for the portion corresponding to the Region of Interest.

Accordingly, Reingold and Loschky found that when they adapted a high-resolution window at the point-of-gaze and degraded resolution in peripheral areas, participants had longer initial saccadic latencies in peripheral areas (the time taken to identify a visual target), than when a low resolution was uniformly displayed across the whole display window [22]. Loschky and McConkie found, in support of earlier studies [23], [24], that if degradation is increased in peripheral areas, then the size of the adapted high-resolution window at the point of gaze also needs to be increased, if the users level of performance is to be maintained [25]. In related work, Osberger and Maeder presented a method of automatically determining the perceptual importance of different regions of an image [26]. Based around the human visual system, using gray scale images, Osberger and Maeder merged five factors that were known to influence attention: contrast with region background; region size, shape and location; determination of foreground and background areas. These factors were combined into an overall "Importance map" (IM), which was used to classify the importance of image regions. Based on the IMs, the authors demonstrated a technique for controlling adaptive quantization processes in an MPEG encoder [27].
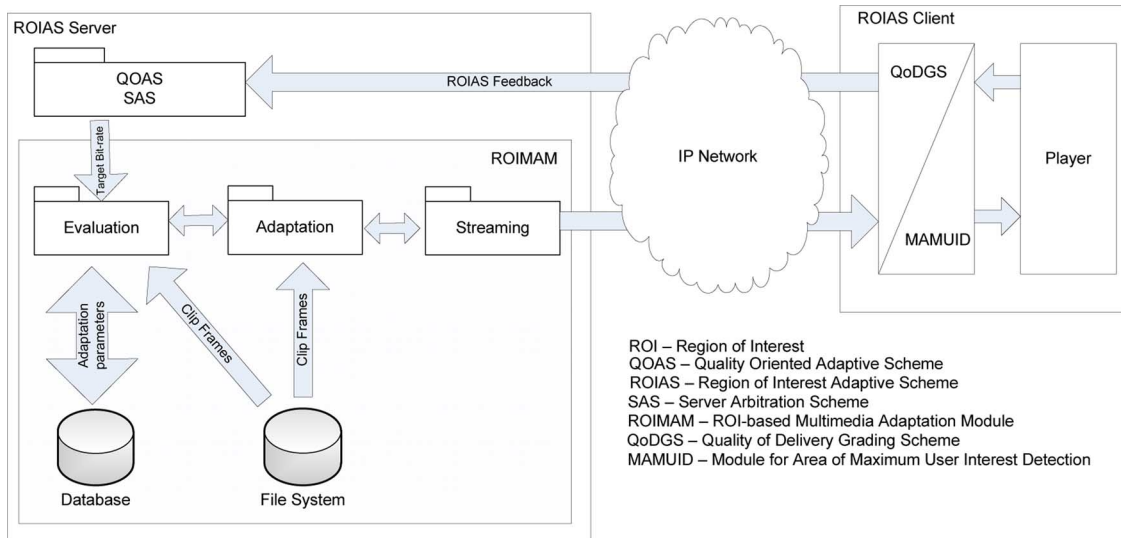
Fig. 2. ROIAS architecture.

In more recent work, Agrafiotis *et al.* [28] present a framework for model-based, context-dependent video coding, based on exploitation of characteristics of the human visual system. The system utilizes variable-quality coding, based on priority maps which are created using mostly context-dependent rules.

Loschky and Wolverton, on the other hand, in their work tackle the interesting issue of perceptual disruptions in GCDs, specifically examining perceptually acceptable update delays in multi-resolutional displays, showing that an update of 60 ms is ample enough to be perceptually acceptable [29].

C. Video Quality Assessment

Video quality assessment methods and metrics are very important for testing adaptive multimedia applications in general and especially for their quality-related evaluation. They are particularly useful to assessment of the effects variable network conditions have on user perceived quality.

Video quality assessment methods can be classified in two categories from the point of view of user involvement in the assessment process: subjective methods and objective metrics [30]. Subjective testing is performed using human observers involved in video perceptual quality assessment [31] and follows methodologies and recommendations such as those from ITU-R BT.500 [32], ITU-T R. P.910 (one way video test methods) [33], and ITU-T R. P.911 (quality assessment methods for multimedia applications) [34].

Objective methods are classified in [35] from the point of view of usability in conjunction with adaptive streaming solutions as out-of service methods (the original sequence is available and no time constraints are imposed) and in-service methods (performed during streaming without original sequence and with strict time constraints). In relation to the existence of the original multimedia stream during the quality assessment [36] the objective methods can be classified into full reference methods (use comparisons with the reference stream), reduced reference solutions (employ feature extraction) and no reference methods (no original stream is required for quality assessment).

Among the most important and widely used objective video quality metrics are the full-reference Pick Signal-to-Noise Ratio (PSNR) [37], Video Quality Measurement (VQM) [38] and Moving Pictures Quality Metric (MPQM) [39]. PSNR is easy to use, has low computational complexity, but was criticized for poor correlation with human perceived quality [8]. VQM measures the perceptual effects of different kind of video impairments such as blurring, jerky motion, blockiness, etc. and provides a higher correlation with subjective quality assessment. MPQM is an objective metric especially designed to consider some human visual system characteristics such as contrast sensitivity and visual masking. It also has a no-reference version defined for MPEG video streams [40].

III. REGION OF INTEREST-BASED ADAPTIVE SCHEME FOR MULTIMEDIA STREAMING

A. ROIAS Principle

Existing adaptive streaming solutions treat the video frame area as a whole and consequently the adaptation process affects the quality of all its regions in the same way, regardless of their position within the frame. However research has shown that there are certain regions in each video frame that the viewers are more interested in then on other areas [9]. Additionally user interest for a specific frame area decreases with the increase in its distance to the region of highest user interest.

Consequently this paper proposes a Region of Interest Adaptive Scheme (ROIAS) for multimedia streaming which differentiates the content adjustment process within the frame based on user interest on certain regions. During the network delivery-based content adaptation, ROIAS defines various *Regions of Interest* (ROI), concentric around the *Area of Maximum User Interest* (AMUI). During the adjustment process, ROIAS decreases ROI's multimedia encoding quality gradually as its distance from AMUI increases. In this way ROIAS achieves higher end-user QoE in comparison to the case when content quality is decreased equally across the whole frame area.

B. *ROIAS-Based Multimedia Streaming System Architecture*

Fig. 2 presents the architecture of the ROIAS feedback-based multimedia streaming system which involves a ROIAS Server and a ROIAS Client which exchange data and control information over an IP network. ROIAS system architecture extends the classic QOAS system architecture [7] which involves a QOAS client module - *Quality of Delivery Grading Scheme* (QoDGS) and a QOAS server unit - *Server Arbitration Scheme* (SAS) which monitor and report on multimedia quality of delivery at the client and suggest content adjustment in terms of target bitrates in existing network conditions.

Unlike QOAS which met those multimedia bitrate targets by adjusting the quality of the whole frame viewing area equally, ROIAS employs two additional modules in order to further improve end-user perceived quality by selectively adjusting the quality of some frame regions depending on user interest.

The server side *ROI-based Multimedia Adaptation Module* consists of three main sub-modules in charge with evaluation, adaptation and streaming, respectively.

The *Evaluation* unit receives the target bitrate from QOAS SAS and the AMUI from the ROIAS client and calculates a set of multimedia adaptation parameters required to achieve the target stream bitrate. These parameters are related to the position of the AMUI, the number of concentric ROIs to be considered, the level of compression to be achieved in each of them, compression-related and encoding-scheme dependent parameters such as quantization factor, etc.

The *Adaptation* unit receives from the evaluation unit the adaptation parameters, processes the clip frame accordingly and obtains the ROIAS adapted multimedia stream.

The *Streaming* unit streams the adapted multimedia content over the IP-network to the ROIAS Client.

ROIAS *Module for Area of Maximum User Interest Detection* is located at ROIAS client and determines the AMUI for each user regarding currently streamed multimedia sequence. Feedback informs ROIAS server about AMUI, in order to take best content adjustment decisions. If the client lacks hardware and/or software support to determine the user interest, default values will be used by the server.

Although currently unicast ROIAS-based multimedia streaming is envisaged only, this architecture is so designed as to allow for extension to multicast. An arbitration scheme will be part of ROIAS server-side Multimedia Adaptation Module's Evaluation unit to aggregate users various areas of maximum interest.

C. *QOAS Multimedia Streaming Adaptation Scheme*

QOAS [7], [30] is based on the fact that random losses have a greater impact on the end-user perceived quality than a controlled reduction in quality [9]. Therefore QOAS employs an end-to-end sender-driven feedback-based adaptation mechanism which adjusts both the quality of the streamed multimedia content and the transmission rate so that it minimizes losses and maximizes end-user QoE in existing delivery conditions.

The QOAS client-side QoDGS continuously monitors some transmission parameters and estimates the end-user perceived quality. It regularly computes Quality of Delivery scores (QoD_{scores}) that reflect the multimedia streaming quality in
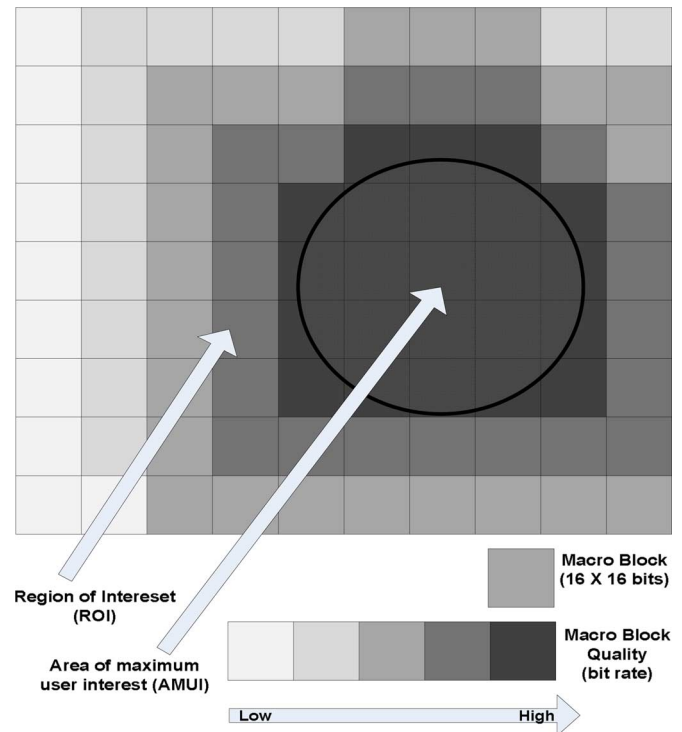

Fig. 3. ROI-based bitrate adaptation scheme.

current delivery conditions. These scores are sent as feedback to the QOAS server-side which analyzes them and proposes adjustment decisions to be taken in terms of target bitrates in order to increase the user QoE in existing delivery conditions.

D. *ROIAS Multimedia Encoding Scheme*

Video encoding techniques exploit visual information redundancy and human visual system sensibility to certain characteristics of the image and lack of sensibility for others to achieve compression and enable streaming. Compression algorithms used in video streaming are usually lossy, some information being lost during the compression-decompression process. This leads to a saving in bandwidth required for streaming, but also to a decrease in quality of the decompressed video stream compared to the original video sequence.

The bitrate of the compressed video stream depends on the compression algorithm employed and also on the value of some specific parameters of the algorithm.

As most multimedia streaming applications use MPEG-1, MPEG-2 and MPEG-4-based encoding, ROIAS uses MPEG's Discrete Cosine Transform, quantization and variable length coding mechanisms to enable ROI-based differentiated quality encoding. MPEG compression is performed at block level (8×8 pixels) and macro block level (16×16 pixels). An important compression parameter is the quantization coefficient. The higher the coefficient's value the lower the bitrate and also the video quality [41]. ROIAS uses the quantization coefficient to vary the quality and consequently the number of code bits for each macro block in the image, depending on its distance to the AMUI.

Fig. 3 presents ROI-based bitrate adaptation at frame level. The macro blocks that are within the AMUI, which is described

AMUI - Area of maximum user interest
ROI – Region of Interest
R – AMUI Radius
D – Macro Block's Distance to the centre of the ROI
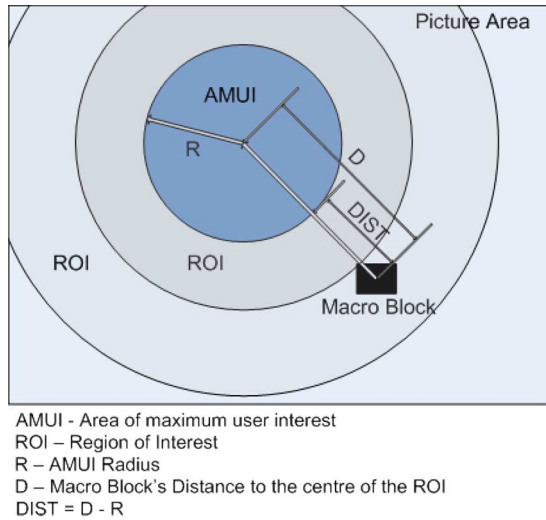DIST = D - R

Fig. 4.   ROI and macro-block quality dependence on the distance to AMUI.

as a circle with a specified radius, are compressed at the highest quality (highest bitrate). The blocks that are situated within a ROI outside the AMUI are encoded at lower qualities dependent on how far their positions are from the AMUI. Consequently the quantization coefficient is determined for each macro block depending on its distance from AMUI. The quantization coefficient is constant for all the blocks within the same ROI resulting in similar quality.

Different solutions can be proposed in terms of the relation between the quantization coefficient (and consequently the quality of the macro block) and its distance from AMUI, given certain target bitrate. ROIAS uses two different approaches: linear and logarithmic quality variation, respectively. These determine two ROIAS versions: Linear quality-distance adaptation ROIAS (Linear-ROIAS) and Logarithmic quality-distance adaptation ROIAS (Logarithmic-ROIAS).

E. Linear-ROIAS - Linear Quality-Distance Adaptation

Equation (1) formalizes the linear dependence between the Quantization Coefficient (QC) and the distance (DIST) of each macro-block from the AMUI. In (1) QC_{max} is the quantization coefficient associated with the highest video quality in this sequence and AC is the ROI-dependent Adaptation Coefficient, which is varied during adaptive multimedia streaming in order to meet the target bitrate. The higher AC, the faster QC is rising, leading to a greater reduction in the resulting multimedia stream bitrate, but also to higher quality degradation.

$$QC(DIST) = QC_{max} + AC * DIST \qquad (1)$$

Fig. 4 presents an illustration of the DIST parameter for a particular macro-block within the video frame area.

The main advantage of employing linear quality variation for ROIAS is the fact that quality decreases smoothly with the increase in distance from ROI to the AMUI. The main drawback of this degradation technique is the low quality of the macro-blocks positioned furthest from the AMUI, which leads to a poor local user perceived quality.

F. Logarithmic-ROIAS - Logarithmic Quality-Distance Adaptation

In a similar fashion with Linear-ROIAS, Logarithmic-ROIAS employs (2) to determine the macro-block's Quantization Coefficient (QC)'s value function of the distance (DIST) of the macro-block from the AMUI.

$$QC(DIST) = QC_{max} + AC * \log(DIST) \qquad (2)$$

The logarithmic dependency is more effective from the point of view of user perceived quality, mainly because the quality degradation starts to be perceived by the user only after a specific threshold is reached.

The main advantage of Logarithmic-ROIAS is the fact that it can distribute video quality in a similar manner the human visual system acts, improving overall user QoE. Its main disadvantage is that a sharp decrease in quality is performed as the distance to AMUI increases, running into the risk of quality degradation to be observed by the human viewers.

G. ROIAS Bitrate Adaptation Process

ROI-based bitrate adaptation process consists of four stages and involves both ROIAS client and server modules. In Stage 1 ROIAS client assesses multimedia streaming process quality and informs ROIAS server via feedback. ROIAS server-located QOAS SAS estimates best multimedia stream target bitrate in order to optimize user QoE. The target bitrate is calculated according to the network current conditions and by aiming at an optimal bandwidth usage and user QoE maximization. In this stage AMUI position is also received by ROIAS server from the client.

In Stage 2 adaptation parameters are computed. The adaptation parameters consist of position of the center of ROI in terms of macro-block level coordinates, ROI's radius in terms of macro-block level coordinates and adaptation coefficients (AC as presented in equations (1) and (2)). The evaluation is made at Group of Picture (GOP) level and can be achieved off-line for Video on Demand or on-line for live streaming.

Each GOP is encoded with different adaptation parameters values and the achieved bitrate is computed. The adaptation parameters that led to the closest bitrate to the target bitrate are chosen and used to encode the GOP in the video stream.

In the off-line mode, the GOP is encoded with adaptation parameters taking all the possible values and the achieved GOP bitrate is stored in a database along with the corresponding adaptation parameters values. This database will be queried during Stage 2 to retrieve the adaptation parameters values corresponding to the target bitrate.

In the on-line mode, evaluation is made during the adaptation process leading to real-time constraints. This involves high processing power requirements for evaluation-encoding and optimization techniques for choosing the adaptation parameters values. Using the evaluation adaptation parameters values that are most likely to result in the bitrate closest to the target bitrate minimize the time required to compute the adaptation parameters values.
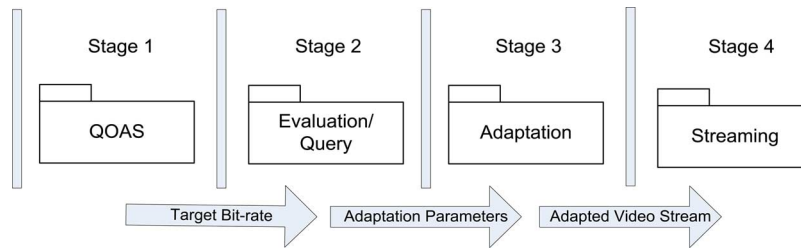
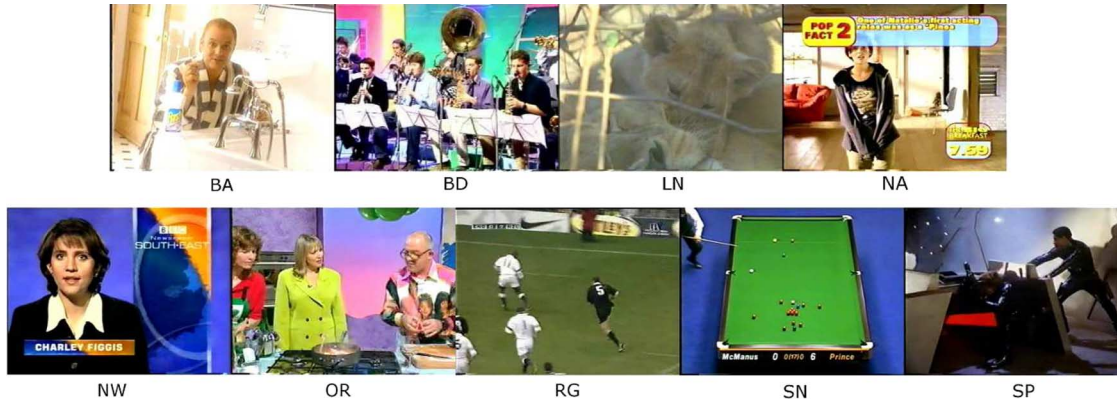Fig. 5. Bitrate adaptation process.

Fig. 6. Movie sequences with different spatial and temporal complexity used during testing.

In Stage 3 the multimedia stream bitrate adaptation is performed based on the adaptation parameters determined in the previous stage and the AMUI region received in Stage 2.

In Stage 4 the ROIAS-adapted multimedia stream is sent over the IP network to the user side ROIAS Client.

Fig. 5 presents the four stages of the ROIAS adaptation process. The data-flow between each stage is also specified. The input of this process is represented by the feedback received by the QOAS server-side module from the ROIAS Client and the output is represented by the adapted multimedia stream.

IV. TESTING AND RESULT ANALYSIS

A. Multimedia Clips

Nine multimedia clips were used for the objective multimedia quality assessment tests. They were selected in order to cover a large spectrum in terms of content type, spatial and temporal complexity. The clips are between 25 s and 40 s long, with a frame rate of 25 fps and a resolution of 640×480 pixels. Their detailed description is presented in [1].

Fig. 6 presents the nine clips used for ROIAS performance evaluation. Each clip contains different types of content with different degree of movement and AMUI predictability:

- BA—commercial clip with average motion and average user interest predictability;
- BD—TV show with average/low motion and average user interest predictability;
- LN—documentary with low motion and low user interest predictability;
- NA—music clip with average motion and average user interest predictability;

- NW—news clip with low motion and high user interest predictability;
- OR—TV show with average motion and average/high user interest predictability;
- RG—football game clip with high motion and low user interest predictability;
- SN—snooker clip with low motion and low user interest predictability;
- SP—movie clip with high motion and high user interest predictability.

B. Acquiring Area of Maximum User Interest

In order to extract the user interest in relation to the multimedia clips, eye-tracking research was employed. An experimental testbed was built from an Arrington Research, Power Mac G3 (9.2) infrared camera-based pupil tracking, the View-Point EyeTracker and the QuickClamp hardware. The Quick-Clamp system was designed to limit head movement and included chin, nose and forehead rests. Consequently, the position of nose and forehead rests remained constant throughout the experiments (45 cm from the screen). However, the position of the chin rest and camera were changed depending on the specific facial features of the participant. To avoid audio and visual distraction, a dedicated, uncluttered room was used for the experiments. To limit physical constraints, except from those imposed by the QuickClamp hardware, multimedia speakers were used instead of headphone speakers. A consistent audio level (70 dB) was used for all participants.

Empirical AMUIs were determined using data obtained from the infrared camera. Eye-tracking data samples contained: X and Y values and timing data. X and Y coordinate values were

TABLE I
AVERAGE MULTIMEDIA QUALITY: PSNR RESULTS

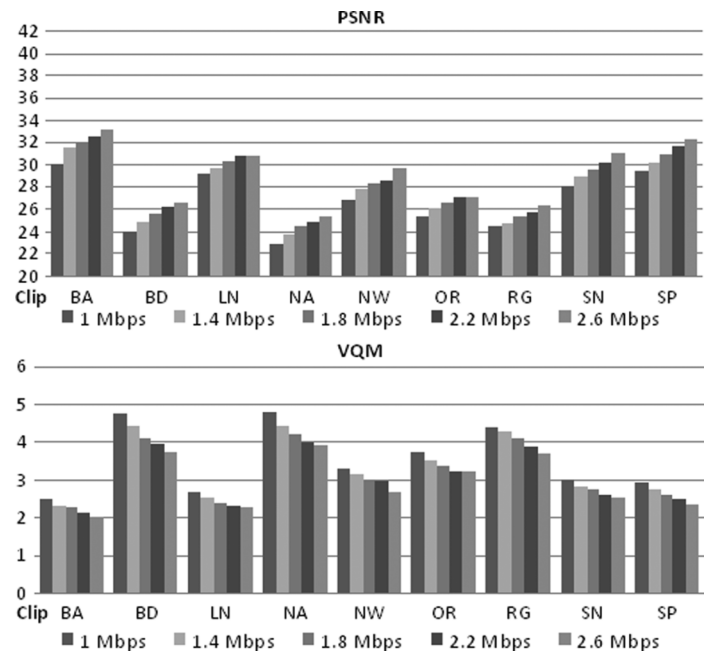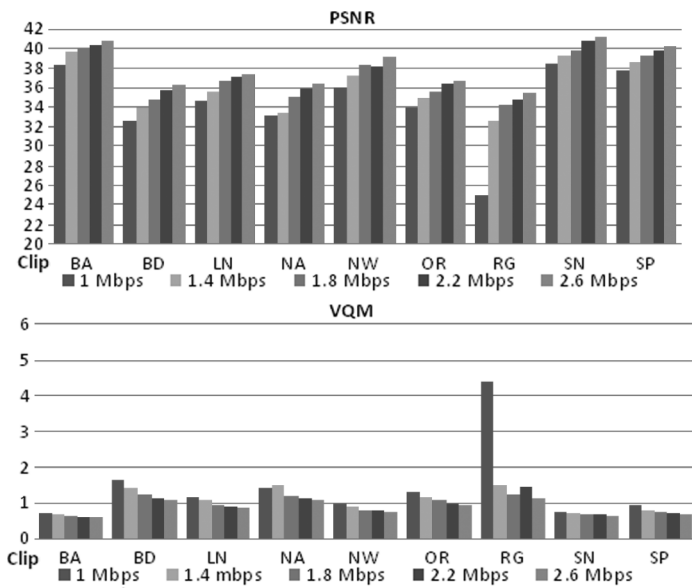|		PSNR								
Mbps		**1.0**		**1.4**		**1.8**		**2.2**		**2.6**
		Score	Gain	Score	Gain	Score	Gain	Score	Gain	Score	Gain
QOAS	Image	33.70	-	35.90	-	37.57	-	38.41	-	38.85	-
	AMUI	44.24	-	46.61	-	48.26	-	49.28	-	49.76	-
LIN-ROIAS	Image	26.64	-20%	27.47	-23%	28.07	-25%	28.57	-25%	29.08	-25%
	AMUI	45.82	4%	48.73	5%	54.31	13%	58.38	18%	63.71	28%
LOG-ROIAS	Image	34.40	2%	36.10	0.58%	37.04	-1%	37.63	-2%	38.14	-2%
	AMUI	48.65	10%	50.20	8%	52.90	10%	55.16	12%	57.41	15%

Fig. 9. Video quality assessment on the whole frame area for Logarithmic-ROIAS in terms of PSNR and VQM.
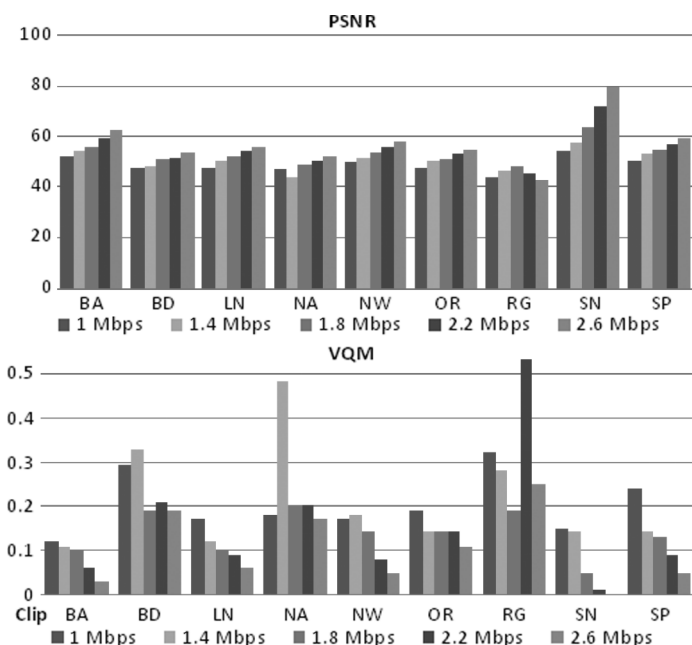

Fig. 10. Video quality assessment on AMUI for Logarithmic-ROIAS in terms of PSNR and VQM.

At the same time, the video quality estimated by both PSNR and VQM is much higher for AMUI than for the whole clip regardless of the clip content. This illustrates how users benefit by using Linear-ROIAS and have higher perceived quality in the area they have the highest interest in.

Fig. 9 presents PSNR and VQM scores for the whole image area when Logarithmic-ROIAS is used. The same trend in quality assessment in relation to bitrate can be observed for Logarithmic-ROIAS as it was noticed for Linear-ROIAS. However a slightly higher user perceived quality is suggested in this case. For example the PSNR score for clip BA encoded at 2.6 Mbps is over 40 for Logarithmic-ROIAS, higher then that achieved when Linear-ROIAS was employed of only 33. A similar situation occurs when VQM is used, confirming the trend.

Fig. 10 presents PSNR and VQM-based quality assessment on the AMUI of the video clips when using Logarithmic-ROIAS. Due to the fact that this logarithmic technique degrades less the ROI situated at the periphery of the image (the areas that are further from the AMUI) the quality provided to AMUI is lower then the quality obtained with the Linear-ROIAS technique. This can be observed in Fig. 10 in comparison with Fig. 8 and is most convincing for clip NA where average VQM score obtained using Logarithmic-ROIAS is 0.25 while for the Linear-ROIAS encoded version is 0.22, which denotes higher user QoE.

Tables I and II summarize the average results of video quality assessment in terms of both PSNR and VQM for the whole image area and AMUI when Linear-ROIAS and Logarithmic-ROIAS are used to deliver the multimedia clips. Improvements in terms of quality perceptual gain (expressed as percentage) when using the two different ROIAS techniques in comparison when employing QOAS are also presented.

Looking for example at Table I Linear-ROIAS achieves a high quality improvement in the AMUI where it recorded a quality gain of up to 28% in comparison with QOAS for the 2.6 Mbps streams. However, due to the fact that Linear-ROIAS trades the quality of the areas furthest from the AMUI for higher quality of AMUI, this technique encounters an important degradation in quality when considering the whole image area in comparison to QOAS.

Logarithmic-ROIAS degrades less the areas furthest from AMUI in comparison with Linear-ROIAS which determines a quality of the whole image area similar to QOAS (small quality improvements for low bitrate streams and little degradations

TABLE II
AVERAGE MULTIMEDIA QUALITY: VQM RESULTS

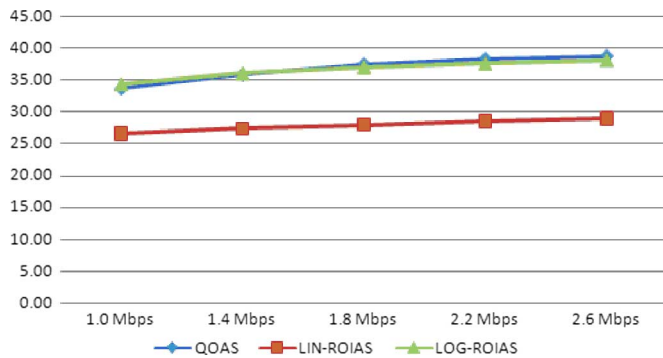| VQM											
Mbps		**1.0**		**1.4**		**1.8**		**2.2**		**2.6**	
		Score	Gain	Score	Gain	Score	Gain	Score	Gain	Score	Gain
QOAS	Image	1.44	-	1.08	-	0.91	-	0.84	-	0.80	-
	AMUI	0.40	-	0.26	-	0.19	-	0.18	-	0.18	-
LIN-ROIAS	Image	3.55	-147%	3.35	-211%	3.19	-251%	3.06	-263%	2.92	-263
	AMUI	0.41	-4%	0.20	22%	0.11	41%	0.08	57%	0.05	72%
LOG-ROIAS	Image	1.47	-2%	1.07	0.21%	0.94	-4%	0.92	-8%	0.85	-5%
	AMUI	0.20	48%	0.21	17%	0.14	27%	0.16	11%	0.10	42%

Fig. 11.   Average PSNR assessed on the entire image.
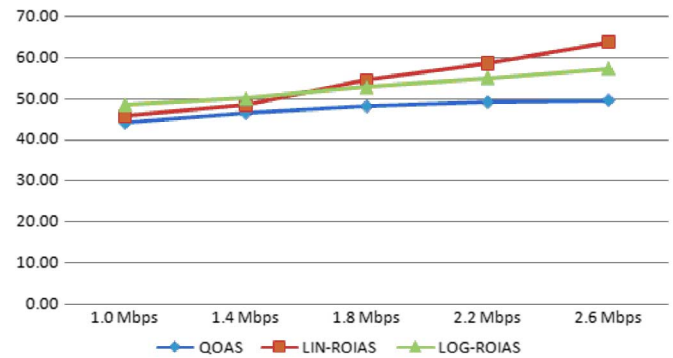

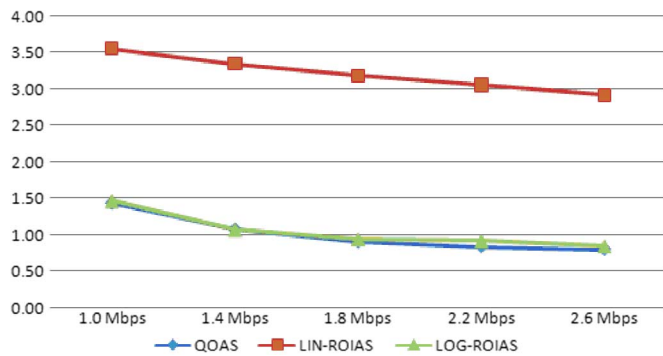Fig. 13.   Average PSNR assessed only on AMUI.

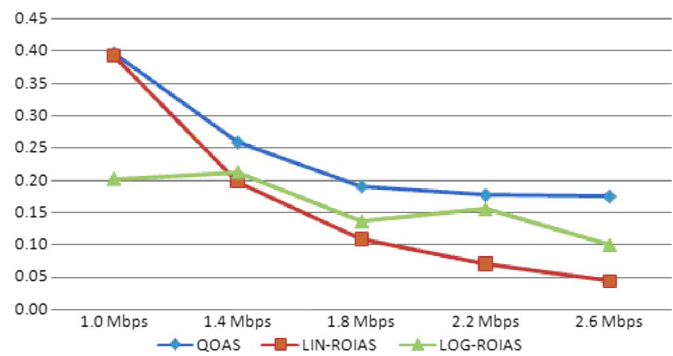Fig. 12.   Average VQM assessed on the entire image.

Fig. 14.   Average VQM assessed only on AMUI.

for higher bitrate clips are recorded). However the AMUI quality gain in comparison with QOAS is less than that of Linear-ROIAS recoding values of 10% for 1 Mbps streams up to 15% for 2.6 Mbps clips.

VQM results presented in Table II confirm the quality assessment trend indicated by the PSNR scores.

E. Comparison-Based Assessment

Figs. 11 and 12 compare Linear-ROIAS, Logarithmic-ROIAS and QOAS video quality assessment results for different target bitrates on the whole multimedia frame area in terms of PSNR and VQM scores, respectively. These results represent average quality scores across all nine multimedia clips. It can be clearly seen from Fig. 11 how the quality of the whole frame area of the video remains the same when using Linear-ROIAS (which employs a ROI-based content adaptation) as it was when QOAS was employed (QOAS adjusts the quality on the whole frame area uniformly). However when

Linear-ROIAS was used, the quality decreased on average with roughly 20%. These results are consistent with those presented in Fig. 12 when VQM was used as video quality metric.

Figs. 13 and 14 compare the same three adaptive techniques on the AMUI only. As Fig. 13 shows, there is a significant benefit of around 15% on average in the estimated perceived quality on the area of user maximum interest when using Linear-ROIAS in comparison with QOAS and of roughly 10% when using Logarithmic-ROIAS instead of QOAS. These results are confirmed by those presented in Fig. 14 concerning the same schemes' quality assessment, but performed in terms of VQM scores.

These results suggest that if the viewer is highly interested in a certain area and presents very little interest in the rest of the frame area, Linear-ROIAS is indicated to be used as this version best enhances image quality in AMUI. However if the user has a relative balanced interest in various areas of the image, but with an obvious peak in AMUI, Logarithmic-ROIAS is best to be employed for streaming as it maintains the same overall user

perceived quality as QOAS, while still improving slightly the quality in AMUI.

V. CONCLUSIONS

As research has shown that there are certain image regions in each frame of a video sequence on which the users are more interested than in the others, this paper described the Region Of Interest-based Adaptive Scheme (ROIAS) which adjusts differently the regions within each frame of the streamed multimedia content based on the user interest in them. Two versions of ROIAS are presented and discussed in terms of the adjustment algorithms employed and their impact on the human perceived video quality. Linear-ROIAS employs a linear quality variation with the distance from the area of maximum user interest within the frame, whereas Logarithmic-ROIAS uses a logarithmic variation of quality with the same distance.

Comparisons with an existing approach which uses a uniform adjustment of quality across the whole area of the image are performed involving a number of different content multimedia clips encoded at various bitrates. The results are presented in terms of two objective metrics which estimate user perceived video quality: PSNR and VQM show how by using ROIAS there is a clear benefit in terms of quality in the areas of highest user interest.

As the objective video quality assessment techniques are often criticized for poor correlation with the human visual system's perception of video quality, future work will assess ROIAS end-user perceived quality subjectively by involving human subjects in perceptual testing. The efficiency and performance of ROIAS will also be evaluated from the point of view of the impact of display size and resolution.

The effect of background traffic and consequent variability of the ROIAS-based adaptive streaming will also be studied.

ACKNOWLEDGMENT

The support of Enterprise Ireland's International Collaboration program and Irish Research Council for Science Engineering and Technology and Microsoft Research is gratefully acknowledged, as is the contribution of Dr. Stephen Gulliver in the collection of eye tracking data.

REFERENCES

[1] S. R. Gulliver and G. Ghinea, "The perceptual and attentive impact of delay and jitter in multimedia delivery," *IEEE Trans. Broadcasting*, vol. 53, no. 2, pp. 449–458, Jun. 2007.
[2] J. Liu, B. Li, and Y.-Q. Zhang, "Adaptive video multicast over the Internet," *IEEE Multimedia*, vol. 10, no. 1, pp. 22–33, Jan.–Mar. 2003.
[3] M. Baldauf, S. Dustdar, and F. Rosenberg, "A survey on context-aware systems," *International Journal on Ad Hoc and Ubiquitous Computing*, vol. 2, no. 4, 2007.
[4] M. S. Kuran and T. Tugcu, "A survey on emerging broadband wireless access technologies," *Computer Networks*, 2007.
[5] G.-M. Muntean and N. Cranley, "Resource efficient quality-oriented wireless broadcasting of adaptive multimedia content," *IEEE Trans. Broadcasting*, vol. 53, no. 1, pp. 362–368, Mar. 2007.
[6] D. Wu, Y. T. Hou, W. Zhu, Y.-Q. Zhang, and J. M. Peha, "Streaming video over the Internet: approaches and directions," *IEEE Trans. Circ. and Sys. for Video Tech.*, vol. 11, no. 3, pp. 282–300, 2001.
[7] G.-M. Muntean, P. Perry, and L. Murphy, "A new adaptive multimedia streaming system for All-IP multi-service networks," *IEEE Trans. Broadcasting*, vol. 50, no. 1, pp. 1–10, Mar. 2004.
[8] W. Zhou, A. C. Bovik, and L. Lu, "Why is image quality assessment so difficult?," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '02)*, 2002, vol. 4, pp. 3313–3316.
[9] R. Gulliver and G. Ghinea, "Stars in their eyes: what eye-tracking reveals about multimedia perceptual quality," *IEEE Trans. Systems, Man and Cybernetics,*, vol. 34, no. 4, pt. A, pp. 472–482, Jul. 2004.
[10] G.-M. Muntean, G. Ghinea, and T. N. Sheehan, "Region of Interest-based Adaptive Multimedia Streaming Scheme," *IEEE Trans. Broadcasting*, vol. 54, no. 2, Jun. 2008.
[11] J. Padhye, J. Kurose, D. Towsley, and R. Koodli, "A model based TCP friendly rate control protocol," in *Proc. NOSSDAV*, 1999.
[12] D. Sisalem and A. Wolisz, "LDA+ TCP-friendly adaptation: A measurement and comparison study," in *Proc. NOSSDAV*, 2000.
[13] R. Rejaie, M. Handley, and D. Estrin, "RAP: An end-to-end rate-based congestion control mechanism for realtime streams in the Internet," in *Proc. INFOCOM*, New York, USA, Mar. 1999, vol. 3, pp. 1337–1345.
[14] R. Rejaie, M. Handley, and D. Estrin, "Layered quality adaptation for Internet video streaming," *IEEE Journal on Selected Areas of Communications (JSAC)*, vol. 18, no. 12, pp. 2530–2543, Dec. 2000.
[15] M. van Der Schaar and N. Sai Shanka, "Cross-layer wireless multimedia transmission: Challenges, principles, and new paradigms," *IEEE Wireless Communications*, vol. 12, no. 4, pp. 50–58, Aug. 2005.
[16] L. Qiong and M. van der Schaar, "Providing adaptive QoS to layered video over wireless local area networks through real-time retry limit adaptation," *IEEE Trans. Multimedia*, vol. 6, no. 2, Apr. 2004.
[17] S. H. Shah, K. Chen, and K. Nahrstedt, "Dynamic Bandwidth management for single-hop Ad Hoc wireless networks," in *Proc. IEEE Int'l. Conf. Pervasive Comp. and Commun.*, Dallas-Fort Worth, USA, Mar. 2003.
[18] K.-A. Cha, "Content complexity adaptation for MPEG-4 audio-visual scene," *IEEE Trans. Cons. Elec.*, vol. 50, no. 2, pp. 760–765, May 2004.
[19] M. van der Schaar and H. Radha, "Adaptive motion-compensation fine-granular-scalability (AMC-FGS) for wireless video," *IEEE Trans. Circuits Syst. Video Technol*, vol. 12, no. 6, pp. 360–371, Jun. 2002.
[20] N. Cranley, L. Murphy, and P. Perry, "Optimum adaptation trajectories for streamed multimedia," *ACM Multimedia Systems Journal*, 2005, Springer-Verlag.
[21] *IEEE Standards for Local and Metropolitan Area Networks: Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 8: MAC Quality of Service Enhancements*, IEEE Std 802.11e, 2005.
[22] E. M. Reingold and L. C. Loschky, "Reduced saliency of peripheral target in gaze-contingent multi-resolutional displays: Blended versus sharp boundary windows," in *Proc. Eye Tracking Research and Applications Symposium*, New Orleans, Louisiana, USA, 2000, pp. 89–93.
[23] S. Shioiri and M. Ikeda, "Useful resolution for picture perception as a function of eccentricity," *Perception*, vol. 18, pp. 347–361, 1989.
[24] B. A. Watson, N. Walker, L. F Hodges, and A. Warden, "Managing level of detail through peripheral degradation: Effects on search performance with a head-mounted display," *ACM Trans. Computer-Human Interaction*, vol. 4, no. 4, pp. 323–346, 1997.
[25] L. C. Loschky and G. W. McConkie, "User performance with gaze-contingent multiresolutional displays," in *Proc. of ACM Eye Tracking Research & Applications Symposium (ETRA) 2000*, Palm Beach Gardens, FL, 2000, pp. 97–103.
[26] W. Osberger and A. J. Maeder, "Automatic identification of perceptually important regions in an image using a model of the human visual system," in *Proc. 14th International Conference on Pattern Recognition*, Brisbane, Australia, Aug. 1998.
[27] W. Osberger, A. J. Maeder, and N. Bergmann, "A perceptually based quantization technique for MPEG encoding," *SPIE Human Vision and Electronic Imaging III*, Jan. 1998, 3299.
[28] D. Agrafiotis, S. J. Davies, N. Canagarajah, and D. R. Bull, "Towards efficient context-specific video coding based on gaze-tracking analysis," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 3, no. 4, 2007.
[29] L. C. Loschky and G. S. Wolverton, "How late can you update gaze-contingent multiresolutional displays without detection?," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 3, no. 4, Dec. 2007.
[30] G.-M. Muntean, P. Perry, and L. Murphy, "Subjective assessment of the quality-oriented adaptive scheme," *IEEE Trans. Broadcasting*, vol. 51, no. 3, pp. 276–286, Sep. 2005.
[31] N. Montard and P. Bretillon, "Objective quality monitoring issues in digital broadcasting networks," *IEEE Trans. Broadcasting*, vol. 51, no. 3, pp. 269–275, Sep. 2005.

[32] "Methodology for the Subjective Assessment of the Quality of the Television Picture," Recommendation ITU-R BT 500-10.

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

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

[35] 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.

[36] D. Miras, "Network QoS needs of advanced Internet applications—a Survey," in *Internet2 QoS Working Group*, 2002.

[37] VQEG, "Objective Perceptual Assessment of Video Quality: Full Reference Television," Mar. 2005 [Online]. Available: ftp://vqeg.its.bldrdoc.gov/Documents/VideoQualityTutorial/TutorialJ144-1149March05.doc

[38] M. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Trans. Broadcasting*, vol. 50, no. 3, pp. 312–322, Sep. 2004.

[39] C. J. Branden Lambrecht and O. Verscheure, "Perceptual quality measure using a spatio-temporal model of the human visual system," *Proc. SPIE*, vol. 2668, pp. 450–461, Mar. 1996.

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

[41] O. Verscheure, X. Garcia, G. Karlsson, and J.-P. Hubaux, "User-oriented QoS in packet video delivery," *IEEE Network*, vol. 12, no. 6, pp. 12–21, Nov./Dec. 1998.

[42] Network Simulator-2 [Online]. Available: http://www.isi.edu/nsnam/ns

[43] MSU Video Quality Measure [Online]. Available: http://www.compression.ru

Bogdan Ciubotaru is a Ph.D. researcher with the Performance Engineering Laboratory, School of Electronic Engineering, Dublin City University, Ireland. He was awarded the B.Eng. and M.Sc. degrees in System Engineering from the Computer Science Department, "Politehnica" University of Timisoara, Romania in 2004 and 2005 respectively. His research interests include wireless mobile networks, multimedia streaming over wireless access networks as well as wireless sensor networks and embedded systems. He is student member of IEEE and Research Institute for Networks and Communications Engineering (RINCE) Ireland.

Gabriel-Miro Muntean is a Lecturer with the School of Electronic Engineering and co-Director of the Performance Engineering Laboratory at Dublin City University, Ireland. He obtained his Ph.D. degree from Dublin City University, Ireland for research in quality-oriented adaptive multimedia streaming over wired networks in 2003. 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 quality and performance-related issues of adaptive multimedia streaming, and personalized eLearning over wired and wireless networks and with various devices. Dr. Muntean has published over 80 papers in top-level international conferences and journals and has authored a book and five book chapters and edited two books. Dr. Muntean is Associate Editor with the IEEE TRANSACTIONS ON BROADCASTING and reviewer for important international journals, conferences and funding agencies. He is member of IEEE and Research Institute for Networks and Communications Engineering (RINCE) Ireland.

Gheorghita Ghinea received the B.Sc. and B.Sc. (Hons) degrees in computer science and mathematics, in 1993 and 1994, respectively, and the M.Sc. degree in computer science, in 1996, from the University of the Witwatersrand, Johannesburg, South Africa; he then received the Ph.D. degree in Computer Science from the University of Reading, United Kingdom, in 2000. He is a Reader in the School of Information Systems and Computing at Brunel University, United Kingdom. His research interests span perpetual aspects of multimedia, Quality of Service and multimedia resource allocation, as well as computer networking and security issues.