Paper

The Design of an Objective Metric and Construction of a Prototype System for Monitoring Perceived Quality (QoE) of Video Sequences

Lucjan Janowski, Mikołaj Leszczuk, Zdzisław Papir, and Piotr Romaniak

Department of Telecommunication, AGH University of Science and Technology, Kraków, Poland

Abstract—The paper presents different no reference (NR) objective metrics addressing the most important artefacts for raw (source) video sequences (noise, blur, exposure) and those introduced by compression (blocking, flickering) which can be used for assessing quality of experience. The validity of all metrics was verified under subjective tests.

Keywords—mean opinion score, no reference metric, objective metric, quality of experience, video artefacts.

1. Introduction and the General Prototype Concept

The importance of "live" streaming, which operates on the basis of wireless networks, has been verified in recent years by the emergence of numerous applications such as mobile TV and IP video monitoring systems in urban areas. Unlike traditional applications such as web browsing, multimedia applications require real-time content transmission mechanisms with a low negative impact on user-perceived quality of video communication [1]. To meet this requirement, increase user satisfaction and, consequently, increase the profits for service suppliers, a system of evaluation/verification of video artefacts must be developed and implemented. This solution should be designed for wireless transmission infrastructures in order to control the pseudo-subjective quality of "live" transmitted video sequences [1]. The term "pseudo-subjective" means control with the use of objective metrics, verified on the basis of subjective assessments.

Limitations of traditional solutions based on the notion of quality of service (QoS) require arrangements such as described in [2], i.e., taking into account the characteristics of the transmission media, human vision (human visual system, HVS), and the level of quality as perceived by the user (quality of experience, QoE). However, most of the currently available QoE assessment systems are designed either for one specific type of visual content and application, or for one specific scenario of a wireless service. In recent years, models without a reference (also known as no reference, NR, models) have gained particular focus. To evaluate the QoE, they do not require access to the reference (undistorted) sequence.

It should be noted that the development of new QoE models working in the NR scenario is still a challenging area of research because of the limitations of current metrics (which must be applicable in a non-laboratory environment), diversification of the evaluation based on the content and user profile, resistance to variety of emerging distortions, and the need to meet the requirements of low computational complexity.

This paper highlights the need for assessment of imaging artefacts for "live" streaming applications in a wireless environment and describes the models implemented in the NR scenario assessment. The proposed solution is verified using the results of psycho–physical experiments. The results obtained show the usefulness of the proposed mechanisms for assessing the quality of streaming applications in a wireless environment, and confirm the high correlation with the feelings of users.

The concept presented in this article is to create techniques and tools that can be implemented by service providers with a view to continuing the monitoring of overall video sequence streaming service quality. The results (technology and tools) are expected to be used (mainly) in the wireless service.

The most innovative and distinctive functionality of the system is the introduction of NR metrics to assess and monitor the QoE. It should be noted that the proposed credible assessment and control of the perceived quality of video sequences, based on the QoE numerical estimations and the accuracy calculation of video reconstruction in the context of specific parameters and playback rate conditions, play a fundamental role in ensuring QoE for services based on video sequence streaming.

As already mentioned, the quality estimation solution, allowing to assess video sequences when there is a lack of available references, remains a challenge. In contrast to all methods based on reference solutions (full reference, FR, and reduced reference, RR), which are limited by shortcomings in the quality of the source video sequence, the NR approach evaluates absolute quality, as seen from the perspective of the user. NR does not require the additional, ideal channel to transmit data to be used as a reference. In addition, the NR solution allows for "live" transmitted session tracking, allowing the delivery of estimated results in real-time.

For real applications the authors are interested in absolute quality throughout the supply chain of media (known as end-to-end); in other words, from the beginning (the impact of focus, noise, exposure), through the transition stages (the impact of stream bandwidth scaling), to the end (the impact of the presentation and application). NR-type methods of assessing quality are therefore a natural response to the needs of real video sequence streaming scenarios.

It is particularly important to assess the impact of scalable stream bandwidth. Gaining increasing popularity, video sequence streaming services are still faced with the problem of limited access links bandwidth. Although for wired connections bandwidth is generally available in the order of megabits, higher bit rates are not as common for wireless links. Users of wireless links cannot expect a stable high-bandwidth connection.

Therefore the solution to run video sequence streaming services for such access lines is transcoding video streams "on the fly". The transcoding result is bit rate (and quality) scaling to personalize the stream sent to the current parameters of the access link. Scaling the quality of the video sequences is usually in the (often inseparable) domains of compression, space and time. Scaling in the compression domain usually boils down to operating the codec quantization parameter. Scaling in the spatial domain means reducing the effective image resolution, resulting in increased granularity when one tries to restore the original content to screen size. Scaling in the temporal domain amounts to the rejection of frames, i.e., reducing the number of frames transmitted per second (FPS).

The abovementioned scaling methods inevitably lead to a lower perceived quality of end-user experience. Therefore the scaling process should be monitored for QoE levels. This gives the opportunity to not only control but also to maximize QoE levels in real time, depending on the prevailing transmission conditions. In case of failure to achieve a satisfactory level of QoE, an operator may intentionally interrupt the service, which may help in preserving and allocating network resources to other users.

Unfortunately, determining the level of QoE in any case cannot be reduced to a simple maximization of quantitative parameters in each of the three domains. Consumer perception based on HVS, is highly non–linear and depends on many variables (such as visual content). Therefore attempts are made to create models for automatically determining QoE levels through the analysis of visual content as seen by the user [3].

Attempts to determine the impact of scaling in the compression domain on the perceived QoE quality are particularly difficult. The same compression ratio is not a sufficient indicator of perceived quality. In the NR model it is necessary to identify the impact of this manipulation on the effects shown in the image. The most important effects associated with lossy compression are block artefacts and block flickering. To determine the QoE it is necessary to accurately and quantitatively assess the severity of these effects. Numerous models given in literature [3], [4], [5] usually do not achieve a sufficient correlation with actual user ratings.

It is far easier to model the impact of scaling in the temporal domain, because here at least the FPS value is openly available. Attempts to model the impact of scaling in the domain of perceived quality have been made in several studies, including [6]. It is relatively less complicated to evaluate the effect of a reduction in motion picture effective resolution (i.e., increase in granularity) on the visual effects. These effects were studied in [3], [7], although the former work used other applications.

The methodology of studies presented in the section covering the evaluation of scalable video sequences of this paper is based on subjective quality tests on independent influences of the three abovementioned scaling methods. In addition, the study was carried out on developing metrics, evaluating each quality parameter and presenting the results of statistical analysis of results.

The first value-added feature of this research is the provision of an identical environment for the psycho-physical experiment for all test artefacts and for scaling all three domains of quality using an eleven-point quality scale. This provides an opportunity to compare the results obtained for all considered scaling methods, and to build an integrated model (still being refined) that takes into account the simultaneous combination of methods. Another innovative element is the measure of evaluation of lower qualities due to the large value of QP – a measure with a very high correlation with subjective assessments. Another added value is a detailed statistical analysis of the results obtained for correlation with the mean opinion score (MOS) and statistical reliability. It is an often overlooked element in work on QoE modeling. Furthermore, different video sequences have been used and considered in subjective tests as an additional independent variable, in some cases allowing for the statistical analysis of the impact of a given sequence on the accuracy of the resulting measurement.

In summary, the authors present a concept that involves creating and implementing QoE metrics based on user preferences, assessments of subjective and observer characteristics, and the feedback loop formed by the iterative verification of metrics, modifying their parameters on the basis of these subjective assessments.

The remainder of the article is structured as follows. Section 2 deals with the measurement of quality and artefacts (based on video parameters – Subsection 2.1 and network parameters – Subsection 2.2). Section 3 presents the psycho-physical experiment verification environment. Section 4 presents a statistical analysis of the results in terms of measurement and scaling artefacts in the compression, spatial and temporal domain, and information on how to implement the prototype, while Section 5 contains conclusions and plans for further research.

2. Quality Measures

This section includes a description of video quality metrics dedicated to the assessment of certain video artefacts in a no reference mode. The metrics are used to build a QoE monitoring prototype. This section provides a summary of the work on video quality metrics developed in recent years. For detailed descriptions of the metrics please refer to [1], [8], and [9].

2.1. Measures of Quality Based on QoV Parameters

Exposure. An exposure distortion is understood as the overall quality degradation caused by incorrect exposure time. The metric was inspired by the shape of histograms of images taken for different exposure times. A histogram of a correctly exposed image spreads over the whole luminance range. Histograms of over- and under-exposed images are shifted to the bright and the dark side respectively. The higher the exposure distortion, the more significant the shift. In other words, there are no completely black and white regions on over- and under-exposed images respectively. Consequently, the exposure metric is based on histogram range inspection.

The metric is calculated locally for each video frame. In the first step mean luminance is calculated for each macro-block of a given video frame. The average of the three macro-blocks with the lowest and highest luminance represent luminance (histogram) bounds. The exposure metric for a single frame is calculated as:

$$ex = \frac{L_b + L_d}{2}, \tag{1}$$

where L_b and L_d are bright and dark luminance bounds.

The video level metric is calculated by averaging frame metrics over one scene. The proposed methodology assumes that each natural video sequence has at least some bright and dark regions. It is a significantly more accurate approach than a simple histogram average luminance calculation. For instance, it eliminates the problem when images showing black objects with very few bright regions would be classified as under-exposed.

Blur. The most common approach of image blur estimation utilizes the fact that blur makes image edges less sharp. Recent work representing this approach is described in [10] and [11]. The proposed metric is based on an average width of sharp edges only. Sharp edge selection is critical in terms of predicting accuracy since it eliminates strong content dependency. In the first step, strong edges are detected using the Sobel operator. In the second phase, edge width is measured as a number of neighbouring pixels (localized on the left and right in the same horizontal line) that fulfils the following criteria:

- the right-localized pixel intensity values systematically increase/decrease for rising/falling edges,
- ditto for left-localized pixels,
- the edge slope value does not fall below a certain level, defined by the standard deviation of surrounding pixel intensity.

Noise. The idea behind the proposed noise metric is based on research presented by Lee in [12]. According to this work, the most convenient method of estimating noise in remotely sensed images is to identify homogenous image areas and use them to calculate noise statistics. More recent work utilizing this approach is presented by Farias in [10] and Dosselmann in [11].

The presented approach of identifying homogenous regions guarantees the selection of a comparable number of blocks for images ranging from low to high spatial complexity. It outperforms approaches based on a fixed threshold in terms of a visual nuisance prediction performance. In order to eliminate moderate content dependence (a drawback of existing metrics), the spatial masking phenomenon is addressed by weighting frame-level noise values according to the spatial activity of a given frame. This compensates the well-known property that images with low spatial complexity are more exposed to visual distortion caused by noise.

Block artefact. Blockiness artefact measurement is based on the well-known discrete cosine transform (DCT)-based coding. Each blockiness artefact has at least one visible corner. Recent research utilizing this fact is described in [10] and [11]. In the proposed approach the blockiness artefact is calculated locally for each coding block. The absolute difference of pixel luminance is calculated separately for intra-pairs, represented by neighbouring pixels from one coding block, and inter-pairs, represented by pixels from neighbouring blocks. A ratio between the total value of intra- and inter-difference is considered as the blockiness level.

Block flickering. The flickering metric described here was inspired by the work presented by Pandel in [13]. The implementation task was threefold. The first aim was to define the threshold used to decide whether a given macro-block remains in state of no-update. In [13] the threshold was defined as the mean squared difference between the pixels of the current and corresponding macro-blocks, although the exact value was not revealed. The authors calculated the threshold as an average of absolute differences in pixel luminance for each 16×16 macro-block. Second, a different method for spatial pooling was proposed by calculating the frame-level flickering measure as a mean value over a small number of macro-blocks with the highest values (number of transitions between states). Third, the two previous parameters were adjusted in order to optimize prediction performance defined as a correlation with subjective scores. Similarly to the blockiness metric, averaging over a time window is required; the window size was equal to the sequence length for the purposes of the experiment.

In order to maximize the correlation of the flickering metric F with MOS, the authors considered several threshold values (between 0.5% and 2% of luminance change) and several numbers of macro-blocks with the highest number of transitions between states (between 0.5% and 10% of macro-blocks). The highest correlation with MOS was achieved for the threshold equal to 1% and frame-level flickering averaging over 3% of the total number of macroblocks.

2.2. Measurement of Video Content Characteristics

For the purpose of the subjective experiment the authors were interested in choosing a good representation of videos to be included in the sequence pool; this means video sequences which would obtain different MOSs for the same compression parameters. On the other hand, sequences which are similar in terms of scene complexity should be avoided because they would not provide any additional information to the experiment.

The key parameters describing any video sequence characteristics are spatial and temporal information, i.e., the number of details and the movement dynamics respectively. In order to make the selection task easier, the authors use a scene complexity measure [14]. Scene complexity is a function of both spatial and temporal information which provides information on how difficult a given video sequence is to encode. It should be noted that it is represented by a single value, therefore the task of sequence selection becomes significantly simpler than for selection based on spatial and temporal information. It is easier to decide which scenes are close to each other and which are not.

The question how to measure all these content characteristics remains open. This paper utilizes a method presented in [14] where *scene complexity o* is defined as

$$o = \log_{10}\left(\underset{n}{\mathrm{mean}}\left[SA(n) \cdot TA(n)\right]\right), \tag{2}$$

where $SA(n)$ is spatial activity computed for the nth frame and given by

$$SA(n) = \underset{space}{\mathrm{rms}}\left[\mathrm{Sobel}\left(F(n)\right)\right] \tag{3}$$

and $TA(n)$ is temporal activity computed on the base of nth and $n-1$th frames given by

$$TA(n) = \underset{space}{\mathrm{rms}}\left[F(n) - F(n-1)\right]. \tag{4}$$

In both Eqs. (3) and (4) $F(n)$ denotes the nth video frame luminance channel. Sobel is the Sobel filter [14] and rms_{space} is the root mean square function over an entire video frame.

3. Verification of Measurement by Subjective Psycho–Physical Experiments

In order to properly model the image quality parameters to the assessment of subjects, an appropriate environment for the psycho-physical experiment was created. The experiments were performed at the AGH University of Science and Technology in Kraków. They were attended by over 100 students. Very similar conditions (LCD monitors and lighting) were provided for all test positions (see Fig. 1), and the experiments themselves followed the Video Quality Experts Group (VQEG) methodologies [15] wherever possible.

Fig. 1. Psycho–physical experiment environment.

"Barcelona" "Harp" "Canoa Valsesia"

"Fries" "Rugby" "Mobile & Calendar"

"Baloon-pops" "New York 2" "Betes pas betes"

"Autumn leaves" "Football" "Sailboat"

"Susie"

Fig. 2. Thirteen VQEG test sequences.

The experiment used thirteen VQEG test sequences (see Fig. 2) [15], [16], [17]: “Barcelona” (#2), “Harp” (#3), “Canoa Valsesia” (#5), “Fries” (#7), “Rugby” (#9), “Mobile

& Calendar" (#10), "Baloon–pops" (#13), "New York 2" (#14), "Betes pas betes" (#16), "Autumn leaves" (#18), "Football" (#19), "Saitboat" (#20) and "Susie" (#21). These sequences reflect the broad spectrum of two different characteristics of the content (temporal video activity and spatial video activity). Video sequences were encoded using the H.264 codec (X264 implementation), main-profile (Level 40). In accordance with the VQEG recommendations, QP was selected to obtain the order of the average bit rate streams of 5000 kbit/s (Compression Ratio, $CR = 50.38848$), 1000 kbit/s ($CR = 251.9424$), 500 kbit/s ($CR = 503.8848$), 300 kbit/s ($CR = 839.808$), 200 kbit/s ($CR = 1259.712$) and 100 kbit/s ($CR = 2519.424$).

The initial FPS rate was 30 with FPS rates values of 15, 10, 7.5, 6 and 5 also examined.

The effective initial resolution was the SD/D-1 NTSC resolution (720×486). Additionally the HHR 525 (352×480), SIF (352×240), QCIF (176×144) and SQCIF (128×96) resolutions were examined.

The authors used the ACR methodology, as described in ITU-T Recommendation P.910 [18]. This methodology represents the single-stimulus (SS) approach, which means that all video sequences in the test set are presented one after another without the option of comparison with the reference. Reference sequences are included in the test set and evaluated on the same basis as the others. This approach is known as ACR with hidden reference (ACR-HR). An eleven-step, numerical quality scale was used [18].

4. Statistical Analysis of Results of the Evaluation – Implementation of Prototype

This section contains a description of the methodology used to build models which are the components of the prototype used to evaluate the perceived QoE of streaming video sequences. The prototype includes the following components:

- single metrics to evaluate the quality of the source material,
- metric scaling in the time domain,
- metric scaling in the space domain,
- integrated metrics for the evaluation of H.264 compression (scaling in the domain of compression).

Descriptions of individual quality metrics and sequence characteristics of the video sequences are presented in Section 2. For scaling in the time and space domains, the values are taken directly from the sequence parameters, and the FPS number and resolution do not require specific metric algorithms.

The prototype was implemented in MATLAB, using standard libraries for processing images and video sequences. The system is able to analyze video sequences stored in files on a local disk. The parameters passed when the starting script is called allow any individual metrics, integrated metrics, and metrics that will be counted for the analyzed video sequence selection. In addition, there is the option of script setting, which is able to analyze multiple sequences and record the obtained results into the database.

4.1. Methodology of Model Building

Ratings obtained for en eleven-point scale are a significantly better approximation of normal distribution than results obtained for a five-point scale. This is because the eleven-point scale has two extreme answers, which should not be popular choices (responses 10 and 0). This allows to obtain less skewed distribution than for a five-point scale answer distribution. It should be noted that a skewed distribution is distinctly different from the Gaussian distribution. Therefore in order to model the obtained results the authors assumed a Gaussian distribution of results, making it possible to use the classical regression model.

In addition, all sequences were divided into test and learning sets. All models alongside the presented coefficients were obtained for the learning sets. Only after the final acceptance of the model was it confronted with the test set. Such methodology guarantees correct checking of whether the resulting model has the ability to predict subjective quality, and generalizes the results obtained for the learning sequences.

4.2. Scaling in the Time Domain

Metric scaling in the time domain appears to be very simple, because the information on the number of frames displayed each second is known. However, the constructed metric is not able to correctly model the quality perceived by the user. The reason is the lack of information about the sequence content. The model presented here also takes into account another factor, which is the amount of image detail. In addition, statistical analysis showed that the natural logarithm of the FPS number is a better predictor than the FPS number itself.

For the entire collection of analyzed films the obtained R^2 coefficient is lower than that obtained for the test sequences. However, the coefficient $R^2 = 0.88$ is a very good result and testifies the accuracy of the resulting model.

4.3. Scaling in the Spatial Domain

As is the case for the time domain, scaling in the space domain is easy to spot because the resolution of the present sequence is precisely known. Similarly to the time domain scaling the information about the image resolution is found to be inadequate because the content of the presented sequence affects the quality change.

The resolution change model takes into account both the amount of detail (SA) and the dynamics of the sequence (TA). In addition, using the logarithm of the resolution provided better results than the resolution value itself. In this case, both the coefficients of R^2 obtained for the test sequences and all sequences are equal.

4.4. Scaling in the Compression Domain

Creating a quality model for scaling in the domain of compression was a significantly more difficult task. The first and most important reason is compression complexity. Each compression scheme has numerous different parameters that define the encoding method. Therefore there is no obvious parameter which affects the QoE. Nevertheless, the test sequences obtained by the authors had relatively high single R^2 values, ranging from 0.74 to 0.89.

4.5. Packet Loss

Packet losses have a significant impact on the quality perceived by users. It is obvious that for larger losses the obtained quality is worse, but it is not true that a particular packet loss level indicates a particular quality of the sequence.

Detailed analysis shows that it is important to identify the location of the losses, both in the GOP structure and within the frame. In order to take into account these relationships it will be necessary to build a model based on additional information. In further studies the authors aim to rely on two possible scenarios. The first is image analysis similar to that used in the construction of a metric scale model in the fields of time, space and compression. The second solution is far more accurate inspection and detection of packets which form part of the picture, and/or the lost GOP. Research is being carried out on this analyzer as part of AGH's activities in the Joint Effort Group (JEG) forum.

5. Conclusions and Further Research

The paper presents a system for assessing the QoE based on measurements of artefacts present in video sequences. The validity of objective metrics has been verified under subjective tests. Statistical analysis of results demonstrates relatively high correlation coefficients as far as a no reference scenario is concerned.

Experiments reveal that the validity of quality measures is influenced by video content. Future subjective experiments will focus on a diversity of video sequences in terms of their spatial (number of details) and temporal (motion level) activity.

Co-operation with the VQEG JEG project will provide an opportunity to enhance the derived metrics by packet loss in the near future.

The proposed metrics, which have been coded in the MATLAB environment, will be moved to optimized, fast C/C++ libraries. Preliminary tests of blocking and flickering artefacts confirm the acceleration of metric computation which is important for a real time deployment.

Acknowledgement

Experiments have been performed under a grant of Ministry of Higher Education "Next Generation Services and Networks – Multimedia Services" and a EU FP7 Collaborative Project 218086 INDECT ("Intelligent information system supporting observation, searching and detection for security of citizens in urban environment").

References

[1] P. Romaniak, "Towards realization of a framework for integrated video quality of experience assessment", in *Proc. 28th IEEE Int. Conf. Comp. Commun. Worksh. INFOCOM'09*, Piscataway, USA, 2009, pp. 417–418.

[2] H. Derbel, N. Agoulmine, and M. Salaun, "ANEMA: autonomic network management architecture to support self-configuration and self-optimization in IP networks", *Comp. Netw.*, vol. 53, no. 3, pp. 418–430, 2009.

[3] L. Janowski, M. Leszczuk, Z. Papir, and P. Romaniak, "Ocena jakości sekwencji wizyjnych dla aplikacji strumieniowania na żywo w środowisku mobilnym", *Przegląd Telekomunikacyjny + Wiadomości Telekomunikacyjne*, vol. 82, no. 8–9, pp. 800–804, 2009 (in Polish).

[4] A. Leontaris and A. R. Reibman, "Comparison of blocking and blurring metrics for video compression", in *Proc. IEEE Int. Conf. Acoust. Speech, Sig. Process. ICASSP 2005*, Philadelphia, USA, 2005, vol. 2, pp. 585–588.

[5] S. Tourancheau, P. Le Callet, and D. Barba, "Impact of the resolution on the difference of perceptual video quality between CRT and LCD", in *Proc. IEEE Int. Conf. Image Process. ICIP 2007*, San Antonio, USA, 2007, vol. 3, pp. 441–444.

[6] M. Ries, O. Nemethova, and M. Rupp, "Performance evaluation of mobile video quality estimators", in *Proc. 15th Eur. Signal. Proces. Conf. EUSIPCO*, Poznań, Poland, 2007.

[7] H. Knoche, J. D. McCarthy, and M. A. Sasse, "Can small be beautiful?: assessing image resolution requirements for mobile tv", in *Proc. 13th Ann. ACM Int. Conf. Multim. MULTIMEDIA '05*, New York, USA, 2005, pp. 829–838.

[8] P. Romaniak, M. Mu, A. Mauthe, S. D'Antonio, and M. Leszczuk, "A framework for integrated video quality assessment", *18th ITC Specialist Seminar on Quality of Experience*, May 2008.

[9] P. Romaniak, L. Janowski, M. Leszczuk, and Z. Papir, "A no reference metric for the quality assessment of videos affected by exposure distortion", in *Proc. IEEE Int. Conf. Multime. and Expo*, Barcelona, Spain, 2011.

[10] M. C. Q. Farias and S. K. Mitra, "No-reference video quality metric based on artifact measurements", in *Proc. IEEE Int. Conf. Image Process. ICIP 2005*, Genoa, Italy, 2005.

[11] R. Dosselmann and X. Dong Yang, "A prototype no-reference video quality system", *Proc. 4th Canadian Conf. Comp. Robot Vision CRV 2007*, Montreal, Canada, 2007, pp. 411–417.

[12] J. S. Lee and K. Hoppel, "Noise modeling and estimation of remotely-sensed images", in *Proc. Int. Geosci.Remote Sensin*, Vancouver, Canada, 1989, vo. 2, pp. 1005–1008.

[13] J. Pandel, "Measuring of flickering artifacts in predictive coded video sequences", in *Proc. Ninth Int. Worksh. Image Analys. Multim. Interact. Serv. WIAMIS 2008*, Washington, IEEE Computer Society, pp. 231–234, 2008.

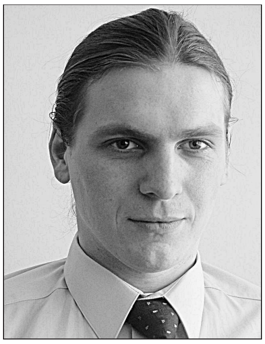
[14] C. Fenimore, J. Libert, and S. Wolf, "Perceptual effects of noise in digital video compression", in *140th SMPTE Techn. Conf.*, Pasadena, USA, 1998, pp. 28–31.

[15] VQEG. *The Video Quality Experts Group* [Online]. Available: http://www.vqeg.org/

[16] VQEG. *Index VQEG Test Sequences*, 2008 [Online]. Available: http://media.xiph.org/vqeg/TestSeqences/ThumbNails/

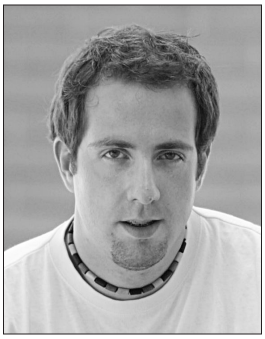
[17] A. Webster, *Objective Perceptual Assessment of Video Quality: Full Reference Television*, 2004 [Online]. Available: http://www.itu.int/itu-t/, http://www.itu.int/ITU-T/studygroups/com09/docs/tutorial_opavc.pdf

[18] *Subjective Video Quality Assessment Methods for Multimedia Applications*, 1999, ITU-T.



Lucjan Janowski is an assistant professor at the Department of Telecommunications (AGH University of Science and Technology). He received his M.Sc. degree in Telecommunications in 2002 and Ph.D. degree in Telecommunications in 2006 both from the AGH University of Science and Technology. During 2007 he worked on a post-doc position in Laboratory for Analysis and Architecture of Systems of CNRS in France where prepared both malicious traffic analysis and anomaly detection algorithms. In 2010/2011 he spent half a year on a post-doc position in University of Geneva working on QoE for health applications. His main interests are statistics and probabilistic modeling of subjects and subjective rates used in QoE evaluation. He has participated in several commercial and scientific projects. He is an author of several research papers.

E-mail: janowski@kt.agh.edu.pl
Department of Telecommunication
AGH University of Science and Technology
Mickiewicza Av. 30
30-059 Kraków, Poland



Mikołaj Leszczuk is an assistant professor at the Department of Telecommunications, AGH University of Science and Technology (AGH-UST), Kraków, Poland). He received his M.Sc. in Electronics and Telecommunications in 2000 and Ph.D. degree in Telecommunications in 2006, both from AGH-UST. He is currently teaching Digital Video Libraries, Information Technology and Basics of Telecommunications. In 2000 he visited Universidad Carlos III de Madrid (Madrid, Spain) for a scientific scholarship. During 1997–1999 he was employed by several comarch holding companies as a Manager of Research and Development Department, President of the Management and Manager of the Multimedia Technologies Department. He has participated actively as a steering committee member or researcher in several national and European projects, including: INDECT, BRONCHOVID, GAMA, e-Health ERA, PRO-ACCESS, Krakow Centre of Telemedicine, CONTENT, E-NEXT, OASIS Archive, and BTI. He is a member of the Video Quality Experts Group Board and a co-chair of Quality Assessment for Recognition Tasks Group. His current activities are focused on e-health, multimedia for security purposes, P2P, image/video processing (for general purposes as well as for medicine), the development of digital video libraries, particularly video summarization, indexing, compression and streaming subsystems. He has been a chairman of several conference sessions. He is a member of IEEE Society since 2000. He has served as an expert for the European Framework Programme and the Polish State Foresight Programme as well as a reviewer for several conferences publications and journals.

E-mail: leszczuk@agh.edu.pl
Department of Telecommunication
AGH University of Science and Technology
Mickiewicza Av. 30
30-059 Kraków, Poland

Piotr Romaniak received his MS.c. Eng. degree in Telecommunications (2006), currently is a Ph.D. student in the Department of Telecommunications at the AGH University of Science and Technology (Kraków, Poland). His area of interest includes knowledge about multimedia systems, content-based indexing, perceptual no-reference video and image quality assessment (objective and subjective techniques), video streaming and video delivery technologies (IPTV, VoD), security and watermarking issues. He was/is involved in international projects funded by Culture 2000, eContentPlus, FP6 and FP7, as well as nation projects funded by Polish Ministry of Science and Higher Education. He was also involved in few research activities: for polish top telecommunication services provider (assessment of the 3rd generation fax image quality, subjective and objective evaluation methods) and in cooperation with other educational entities (panoramic images concatenating, subjective tests – MOS). Currently he is working on 3DTV quality of experience assessment and 2D content quality optimization and assessment in surveillance systems (FP7 IP INDECT Project).

E-mail: romaniak@kt.agh.edu.pl
Department of Telecommunication
AGH University of Science and Technology
Mickiewicza Av. 30
30-059 Kraków, Poland

Zdzisław Papir is professor at Department of Telecommunications (AGH University of Science and Technology in Kraków, Poland). He received the M.Sc. degree in Electrical Engineering in 1976 and Ph.D. degree in Computer Networks both from the AGH University of Science and Technology. In 1992 he received the Dr Hab. degree from the Technical University of Gdańsk. He is currently lecturing on Signal Theory, Modulation and Detection Theory, and Modeling of Telecommunication Networks. During 1991–98 he made several visits at universities in Belgium, Germany, Italy, and US working on statistical modeling of telecommunication traffic. During 1994–95 he was serving for the Polish Cable Television as a Network Design Department Manager. Since 1995 he serves also as a consultant in the area of broadband access networks for the Polish telecom operators. He authored five books and about 60 research papers. The book *Telecommunication Traffic and Packet Network Congestion* was awarded by the Ministry of Science and Higher Education. He was involved in organization of several international conferences at home and abroad. Between 1999–2006 he was a guest editor for IEEE Communications Magazine responsible for the Broadband Access Series. He is a member of an editorial board of a bookseries *Global Information Society* (in Polish). He has been participating in several R&D IST European projects being personally responsible for statistical modeling of telecommunication traffic and subjects/users behavior for quality assessment of multimedia communication services.

E-mail: papir@kt.agh.edu.pl
Department of Telecommunication
AGH University of Science and Technology
Mickiewicza Av. 30
30-059 Kraków, Poland