When Qualinet meets MPEG
Klaus Diepold

Recent PhD theses
A report on the recent tests for MPEG 3D video compression technology, which were organized and performed by Qualinet.

Quality of Experience for Digital Cinema Presentation
Fitri Ranayu

Integral and Diagnostic Intrusive Prediction of Speech Quality
Nicolas Côté

Modelling Perceptual Quality and Visual Saliency for Image and Video Communications
Ulrich Engelke

No-Reference Image and Video Quality Assessment
Tomás Brandão

Video Quality Prediction for Video over Wireless Access Networks (UMTS and WLAN)
Asiya Khan
What's the role of referees in professional football games? Well, they must be essentially important since there is no game without referees. Football clubs invest a big amount of money into professional football, betting on a return of investment in case of success. Referees influence the outcome of the games substantially and hence they have a significant impact on the distribution of the money that is at stake. In spite of their obvious influence and their importance referees never reach the level of popularity (and don't forget the pay) as the professional football players and managers, even though the referees understand the art of football playing at a comparable level since they usually have the best position to watch the games. The referees bring on a distinctive set of capabilities and skills, which are largely overlooked or underestimated by the general public, and they are not allowed to make any mistakes or wrong calls.

What's that got to do with Qualinet? Well, the video quality evaluation experts somehow share the fate with the football referees. They are necessary to settle the question on who got the best video coding technology. The corresponding championship is called “MPEG standardization”. In this championship the proponents of new coding technology take on the role of the star players, their companies or research labs act as the clubs who invest a lot of money for winning the championship. All the fame and glory goes to the successful video coding experts in MPEG and their companies. But it is up to the video quality experts to make the right calls and taking on an active role in deciding the championship and the success in MPEG. It goes without saying that they better make no mistakes or they will have to take the blame for failures. Video quality experts also have their little championship – VQEG, but that league does not generate headlines. Nevertheless, video quality evaluation is a scientific discipline on its own, where the experts produce specific skills and capabilities, which are commonly underestimated and overlooked by the general video coding expert.

MPEG 3D Video Coding

More recently, MPEG has initiated another championship. MPEG defines what type of technology it intends to standardize and then issues a “Call for Proposals”, wherein all companies and research institutions around the world are invited to submit their technologies, which are considered to be candidates for the next round of standardization. Once the proponents have submitted their proposals MPEG needs to evaluate them and rank them according to a set of pre-defined evaluation criteria. Once the ranking is done MPEG selects the winner or the winners of this beauty contest as the core technology built into a new standard.

Now this time the call is about coding technology (vulgo “compression”) for 3D video. One of the major evaluation criteria is the subjective video quality a coding technology achieves under pre-determined coding conditions. So somebody has to do the job to evaluate all proposals – that's the call to the referees. Qualinet can offer a collection of video quality evaluation experts and laboratories. It was therefore natural for Qualinet to organize and perform the evaluation of the proposals on behalf of MPEG. After both organizations exchanged a number of liaison letters an agreement based on this idea was reached.
Assessing the visual quality achieved by any video coding technology is a scientific discipline in itself. Measuring the quality of the compressed video is an essential part of determining the compression performance of a codec. Test material is selected, coding conditions are set, and target bit rates are specified for each test sequence. The encoders compress the test material to produce coded representations of the video samples. Subsequently, the coded bit streams are decoded to create the decompressed video, the quality of which needs to be determined. Quality in this context can be measured fundamentally by two different methods.

**OBJECTIVE MEASUREMENTS:** As a unit of quality measurement, the root mean square error is calculated (RMSE). The original video signals are taken and the decoded versions are subtracted from them. The result is an error image. All error image pixel values are squared and summed up. The lower the RMSE, the better the quality. Engineers often prefer to deal with logarithmic quantities for pure convenience, and therefore the RMSE value is transformed into a logarithmic value, which is called PSNR (Peak Signal-to-Noise Ratio). A higher value for the PSNR implies a better quality image. If a quality measure is required for the entire sequence, then the PSNR values for all the frames in the sequence are averaged to produce one single number. This a simple calculation and leads to consistent results. However, the PSNR values do not correspond well with the visual quality that is perceived when people are actually looking at the images. It often happens that in spite of a higher PSNR value, a sequence is visually judged to be inferior. Similarly, it just as often happens that a video sequence shows a lower PSNR value while still being judged to be superior. PSNR is a mathematical tool that does not reflect human visual perception very well, but which is often used due to its simplicity.

In situations where comparisons are to be made between fundamentally different coding techniques, PSNR values become almost useless. This is because if coding techniques are fundamentally different, then they are likely to produce different types of artifacts and the distortions will look very different. In such situations, PSNR does not tell you very much and what it tells you is usually inaccurate.

**SUBJECTIVE TESTING:** A thorough performance assessment requires that a subjective evaluation procedure be performed. To do this, people are put in front of a monitor where they watch images and are then asked to vote on what they saw in a kind of blindfolded test. This way, the test subjects rank the codecs based on their subjectively perceived visual quality.

In the context of subjective testing for conventional 2D video there exist established standards and recommendations for how the experimental setup is supposed to look what type of monitor, the amount and type of ambient light, the color on the walls of the laboratory environment, the viewing distance, the voting procedure, and so on. The idea is to be able to create test conditions, which can be reproduced in any laboratory around the world and to minimize the influence of the laboratory setup. Finally, the choice of tools and techniques for doing the statistical analysis on the voting data is another important task. Detection and removal of outliers and other undesirable trends, as well as the validation of the results, are time consuming jobs that require experienced and skillful engineers.

All this subjective testing is made necessary by the lossy nature of media compression technology, and it is the reason why the quality assessment of coding performance
for lossy compression systems is a difficult and costly procedure. Clearly, doing an objective measurement instead would be much easier. After all, for doing a meaningful evaluation of coding technology, performing subjective tests is the only reliable way to go.

The Qualinet test laboratories

The Qualinet team performing the evaluation project includes 12 laboratories, which are distributed in Europe, North America and Asia. Most of the laboratories are members of Qualinet. See the table for a complete list of participating laboratories and their respective test managers.

The test laboratories participating in the evaluation of the responses to the 3DV Call for Proposals had to satisfy a minimum set of technical requirements. For example, once they needed to provide specific monitors such as Hyundai S465D 46” for stereoscopic viewing with passive glasses, and a Dimenco BDL5231V3D 52” for auto-stereoscopic display and viewing without glasses. Among other technical specifications, the laboratories were asked to provide a test area of about 5 x 4 meters with no reflecting ceiling, walls and floor, and which is insulated from external audible and visual pollution.

These 12 participating laboratories had varying levels of experience in performing such quality tests for 3D video or stereoscopic video, ranging from a few small-scale subjective quality tests all the way to about 40 formal and informal test campaigns accumulated throughout the past 12 years. Quite a number of test labs have extensive experience in performing subjective quality tests also for earlier MPEG activities such as the tests for the AVC, MVC, SVC and HEVC standards. However, for all participants this recent test project, which included 12 international laboratories, was by far the largest test campaign they were ever involved in.

Since coding for 3D video is a relatively recent technical task it is a question what’s the state-of-the-art in testing 3D video material subjectively. Experts in Qualinet feel that despite some interesting recent progress, the processes for assessing the subjective quality for 3D media is not mature yet. State-of-the-art subjective 3D video quality testing is often following traditional 2D test methodologies. However, 3D scenarios can produce a very different viewing experience. Besides a further development of the test methodologies experts express a need for more technology. One major priority is a “reference” stereoscopic display. The currently available active display technologies can present images at full resolution, but they produce too much ghosting. The currently available passive display technologies suffer less from ghosting, but they are able to show only half the image resolution. So the tests have to deal with an influence of the display technology on the perceived quality, while the coding experts want to isolate the impact of the coding from the display technology.

Numerous experts in Qualinet expressed the opinion that there are still a large number of open challenges to be further investigated. On the other hand, there is still a shortage of systematic research work pushing

<table>
<thead>
<tr>
<th>Test Laboratory</th>
<th>Country</th>
<th>Test Manager</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technische Universität München</td>
<td>Germany</td>
<td>Christian Keimel</td>
</tr>
<tr>
<td>Universidad Politécnica de Madrid</td>
<td>Spain</td>
<td>Narciso Garcia</td>
</tr>
<tr>
<td>Canadian Research Center, Ottawa</td>
<td>Canada</td>
<td>Filippo Speranza</td>
</tr>
<tr>
<td>Yonsei University</td>
<td>Korea</td>
<td>Chulhee Lee</td>
</tr>
<tr>
<td>France Telecom - Orange</td>
<td>France</td>
<td>Gwenael Le Lay</td>
</tr>
<tr>
<td>Queen Mary University of London</td>
<td>United Kingdom</td>
<td>Panos Kudumakis, Naeem Ramzan</td>
</tr>
<tr>
<td>ACREO AB</td>
<td>Sweden</td>
<td>Kjell Brunnström</td>
</tr>
<tr>
<td>Norges Teknisk-Naturvitenskapelige Universitett, Trondheim</td>
<td>Norway</td>
<td>Andrew Perkis</td>
</tr>
<tr>
<td>University of Poitiers</td>
<td>France</td>
<td>Chaker Larabi</td>
</tr>
<tr>
<td>University of British Columbia, Vancouver</td>
<td>Canada</td>
<td>Panos Nasiopoulos</td>
</tr>
<tr>
<td>École Polytechnique Fédéral de Lausanne</td>
<td>Switzerland</td>
<td>Touradj Ebrahimi, Francesca De Simone, Martin Rerabek</td>
</tr>
<tr>
<td>SISVEL, None Torinese</td>
<td>Italy</td>
<td>Vittorio Giovara, Alessandra Mosca</td>
</tr>
<tr>
<td>Fondazione Ugo Bordoni, Rome</td>
<td>Italy</td>
<td>Giancarlo Gaudino, Cristina Delogu, Vittorio Baroncini</td>
</tr>
</tbody>
</table>
the state of the art related to 3D subjective testing. This is good news for young researchers who plan to start working on the topic with the goal to develop a career in this domain. The researchers in Qualinet are certainly driving this initiative in Europe and possibly in the world. So Qualinet offers a great opportunity to launch a research career since it provides an international network of experts in the field and many topics and questions to work on.

MPEG received 23 submissions in response to the call for proposals, including 12 AVC-compatible and 11 HEVC-compatible proposals. The list of proponents included Nokia, Sony, Qualcomm, Samsung, LG Electronics, Philips, Fraunhofer HHI, Disney Research, Poznan University, Ghent University, RWTH Aachen, and Nagoya University.

Numbers indicating the effort

In order to estimate the logistic effort for the test we can begin to determine the number of stereo AVI files which had to be produced for the formal subjective assessment test. MPEG specified the evaluation tests to consist of two test categories (forward compatibility with AVC and forward compatibility with HEVC) and to consider 4 test scenarios (stereoscopic 3 views with centre view pairs, 3 views with randomly selected view pairs and auto-stereoscopic viewing with 28 views). On top of this MPEG agreed to perform the tests for two classes, each class consisting of 4 test sequences where each sequence had to be coded at four different bitrates. Based these numbers one can easily calculate that the formal subjective quality evaluation of one proposal included the assessment of 32 stereo video files for each test scenario. This lead to a total of 32 x 4 = 128 stereo video files. For each Proposal, 32 out of the 128 video files were to be produced by the test coordinator, after he randomly selected one view pair for the test.

The large number of proposals that were received combined by the various viewing conditions lead to an explosion of the test conditions and hence to a huge effort to perform the subjective tests. Let’s have a look at some numbers to provide an impression of the actual size and complexity of the 3D video coding test project. On average about 100 test subjects (i.e. actual people sitting in a video laboratory watching very exciting (!) video clips and voting on their quality) participated in the tests per laboratory. This adds up in an overall participation of well over 1000 test persons worldwide. Some laboratories reported that it was quite a challenge to find and activate a sufficient number of subjects to make their test project meaningful. A single viewing session took about 15 minutes of time for a subject. Typical test sessions consist of two viewing sessions of 15 minutes duration each plus a break of 15 minutes in between viewing sessions plus an initial training session, which also may take about 5 minutes, that is a total of about 50 minutes to an hour for a test session.

The laboratories also employed different remuneration and motivation schemes; some paid hourly rates to the test persons, others were giving out tickets to movie theatres or feeding students with pizza and soft drinks. In addition to the efforts carried by the test subjects about 3 people on average per laboratory worked for the tests during the project period ranging between 25 working days and 3 months. Based on the available data we can estimate the overall effort to be well over 72 person months for the technical staff and practically thousands of viewing hours for the subjects.

For a typical workload of 20 test sessions a total of 560 video files had to be processed and presented. Taking into account the test scenario using Dimenco’s auto-stereoscopic display, which asks for 28 intermediate views results in an impressive amount of
processed video data to instrument the test sessions. Extra people were hired to help monitoring the tests sessions transferring the scores from paper into an electronic version. All in all these numbers show that executing the 3D tests was a major effort for all parties involved.

The amount of video data necessary to perform the tests exceeded the limits of what could be distributed via data transfer. Preparing hard disks and shipping them to the test laboratories was the way to distribute the video data. This process worked ok most of the time, but preparing and shipping all the hard disks caused a significant amount of work and on top of that there were also cases where hard disks got mysteriously lost on their way.

One of the major challenges was the organizational effort that had to be done even before the tests were actually performed. The experts had to agree on a number of practically critical points, which were addressed rather late with respect to the test campaign’s original schedule. Among the issues to be resolved was the question of the formats for displaying the stereoscopic video sequences. The stereo material provided by the proponents came as side-by-side video files. For the displays to play back the videos correctly on a stereo screen the stereo content was supposed to hold the left and right view as interlaced fields. However, this fundamental step was completely missing in the MPEG plan. This resulted in a time consuming activity to process all the stereoscopic data in order to create the correctly interlaced stereoscopic files to be used by all the labs, and to distribute the resulting files to allow for a simple direct rendering of the images. Many of these practical points were not specified in the MPEG document that was made available to the Qualinet experts, and which described the subjective test effort.

On top of that MPEG had decided to use brand new display equipment, such as an auto-stereoscopic display produced by Dimenco that has not been used much for subjective assessments and for which there existed little experience for the best settings of parameters (not even Dimenco knew). Fixing all these issues took a significant amount of time, which in turn resulted in a tight schedule for all participants to actually perform the tests.

Throughout this process the Qualinet experts gained substantial experience, which is very valuable for executing such a complex testing task. In the aftermath, the Qualinet experts will analyze these aspects and hope to provide input to MPEG to improve the proposal specifications for future tests.
Test methodologies

For the tests the experts did not invent or apply new testing methods and procedures. Since the only focus of the test was on coding distortions the MPEG members agreed on a standard 2D test methodology for codec performance evaluation. Qualinet experts felt that there are several more efficient and better evaluation methodologies that could have been used instead. However, the evaluation methodology had to be as close as possible to the conventional evaluation methodologies as described in the corresponding international standards, to avoid discussions about the validity of the test results if a more bold methodology was adopted. The credibility of the results is paramount to MPEG as a lot is at stake for the proponents of new coding technology.

There comes the question if testing 3D video is all that different from testing conventional 2D video. Let’s mention a few differences. For once the quality of 3D content is determined by two major influence factors: the image quality (this is same to 2D quality) and the depth perception. Furthermore, 3D video content creates particular viewing experiences such as immersion. Also, several recent published research show that human subjects have difficulties in precisely assessing average quality in 3D video. The issue of accommodation to 3D viewing is another important issue as it takes a few minutes for test persons to get used to 3D, whereas in 2D this process is much faster. This observation could put in question the 10 seconds duration that typically test videos have when scores are generated.

The Qualinet experts also discussed questions such as the correct viewing distance, how this is determined and how this differs from 2D viewing conditions. The stereoscopic effect depends upon the angular (retinal) disparity in the eye, which in turn depends upon the magnitude of the images on the screen, the disparity and the viewing distance. As a result there is not such a thing as a special 3D viewing distance. However, the community needs to agree on some guidelines to set the distance. For stereoscopic displays experts proposed to use the same distance as for 2D since people will be using the same display for 2D and 3D and the baseline comparison of 3D performance is actually the 2D performance (the 3D technology is supposed to be ‘better’ than 2D). For auto-stereoscopic displays it’s the characteristics of the display that determines the optimal viewing distance.

Another difference between 2D and 3D is that there are additional parameters, which need to be taken into account via additional test conditions. This leads to an explosion...
of the test complexity and hence creates new challenges for designing experiments to arrive at an efficient and effective test schedule.

**Challenges for future work**

Maybe in the future the test design could be supported in a systematic way by new computational tools. Also, new methods for actually doing the data analysis after the tests are completed could contribute to improve the situation for tests with increasing test complexities. Further items for future work on test methodologies includes questions if it is sufficient to adapt standard test procedures in a minimal way or if we need entirely new testing methods. A pragmatic approach is to extend and adapt the currently used 2D methodologies to the study of 3D. The core procedures for testing then could be the same for 2D and 3D. However, there is a need to invent recommendations, which address the preparation and the presentation of the videos under test. The most challenging issue is how to measure the perceived depth and connect it with the perceived quality. The 3DV tests for MPEG focused only on assessing the mere quality of the compressed video signal rather than considering the quality of an entire 3D experience. Studies focusing on understanding user’s quality perception of 3D content, including for example monitoring of focus of attention mechanisms and audio-visual interactions, are the challenges for the future.

From this point of view Qualinet experts commented that MPEG should consider also to issue a call for proposal with respect to the test methodology to be used in future MPEG tests. This call should be issued during the planning phase of the test campaign with an active involvement of the Qualinet experts.

Initial cross correlation analysis revealed a surprisingly good agreement of the test results between all participating laboratories. Qualinet experts currently plan to perform some in-depth cross correlation analysis in order to deliver an empirical basis for a certification process for test laboratories or a handbook of best practices on how to do 3D tests. This way the community can increase the level of trust concerning the reproducibility and validity of tests performed in different laboratories at different times.

As usually in the case of standardization studies, the quality level of the test material overall was quite low compared to what a user expects to watch in a real scenario. It is commonly understood how difficult it is to get uncompressed video sequences to be used for testing purposes. Thus, sharing the test material (i.e. raw views submitted by the proponents) would be very useful. The goal of this test was to evaluate the performance of several new 3D video coding algorithms. As the tests could clearly distinguish between good and better proposals the video material was good enough to reach the goal of the tests. But if the type of video material was realistic enough to represent the anticipated 3D applications and if the bit rates selected by MPEG were adequate are open questions. In any case, the community constantly requires more special 3D content for testing that is also freely available and comes with the appropriate copyright statements. It is a clear goal of the current work in Qualinet to create an ftp site to share such material to support further research work in this domain.

Besides the technical and scientific work and findings a test campaign of this scale also produces a couple of entertaining anecdotes. For example, one lab manager reported that they surprised a few of the test persons as the results of the stereo acuity test by finding out they have no depth sense. Another entertaining episode during the preparation of the tests was the first coordination meeting, which was organized as a Skype conference call. As it often happens with Skype calls participants drop out of the conference call and try to come in again to continue discussing. The very nature of Skype conference calls resulted in a situation where the people were discussing
in two different conferences at the same time unaware of the existence of another conference where the other half was discussing. As a conclusion of this experience Qualinet switched to a more professional system to convene large-scale conference calls.

**After the tests is before the tests**

After the tests were completed MPEG and Qualinet issued a liaison letter to continue their collaboration. One central item is to allow the use the raw data from the 3DV evaluation tests for ongoing research work. This opens the way for more research work using such a comprehensive body of experimental data. One of the immediate goals for the further cooperation between MPEG and Qualinet is to develop an objective quality metric that can be used in the ongoing MPEG core experiment work to computationally assess and compare the quality provided by new coding tools without going full circle with formal subjective tests. Initial work in this direction has started, but more work lies ahead before a conclusive answer can be given to this question.

Video quality assessment experts in Qualinet are still facing numerous technical and scientific challenges ahead, which they will master eventually. If they will become famous for these accomplishments is yet another open question. Looking at the similarity between international football and the world of video it is noteworthy to see that there is at least one referee who made it to become famous – Italy’s Pierluigi Colina. So there is hope that one day some quality assessment expert may become as famous as the known coding experts.

**Acknowledgement**

The author thanks the test managers who were kind enough to share their experiences and their numerous insights gathered throughout the MPEG evaluation tests. Without their input this article would not have been possible.
Recent PhD Theses

Integral and Diagnostic Intrusive Prediction of Speech Quality

Author: Nicolas Côté
URL: http://www.springer.com/engineering/signals/book/978-3-642-18462-8/

In the thesis a new class of diagnostic prediction models for the perceived quality of transmitted speech is described. Such methods simulate the speech perception process employed by human subjects during auditory experiments on the basis of perceptually relevant dimensions. Recent experimental studies have shown that listeners make use of such perceptual dimensions to judge on the quality of speech signals. In order to represent the signal at a higher stage of perception, the author developed a new model, called “Diagnostic Instrumental Assessment of Listening quality (DIAL)”. It includes a perceptual and a judgmental model, and thus simulates the whole quality judgment process. The model compares a speech signal either transmitted through or processed by a system to a corresponding clean signal.

DIAL estimates an overall quality value on a unique quality scale for all telephone networks from traditional analog up to S-WB telephony. In addition, DIAL is the first instrumental model providing a diagnostic profile on four perceptual dimensions: coloration, loudness, discontinuity and noisiness. Except for strong discontinuities, DIAL predicts very well the speech quality of different speech processing and transmission systems: Pearson coefficients show a correlation of 0.936 over 55 Narrow-band (NB, 300–3 400 Hz) databases and 0.910 over 39 WB and S-WB databases. With these results DIAL outperforms the long-standing standard WB-PESQ. In addition, DIAL provides relatively good quality dimension estimations, especially for the coloration dimensions.

The book provides a detailed description of the approach and the implementation of the DIAL model. In addition, it presents an outstanding comparison to existing standards, which makes it a valuable resource for the interested practitioner and expert alike.

Quality of Experience for Digital Cinema Presentation

Author: Fitri Ranayu
Supervisors: Andrew Perkis and Touradj Ebrahimi, NTNU Trondheim, Norway
Published: November 2011

Quality has not been used explicitly to drive the Digital Cinema roll out. The open question still remains what role the “Quality of Experience” (QoE) plays for the adoption of Digital Cinema. In this context higher dynamic range for each pixel and significantly higher pixel counts add additional distinctive factors influencing QoE.
assessment for Digital Cinema presentations. The focus of this thesis was to perform formal tests assessing the subjective quality of digital cinema presentations, which were conducted in a commercial cinema in Trondheim, Norway (see Fig. 3).

For images we conducted subjective image quality assessments for Digital Cinema. We obtained the parameters of a Multi-Scale Structural Similarity (MS-SSIM) objective metric for Digital Cinema presentation. We analysed the performance of several objective metrics including MS-SSIM with original parameters and parameters obtained from our experiment in the Digital Cinema. The results show that in the case of Digital Cinema, MS-SSIM does not exhibit the same type of performance that has been reported in the literature, when compared to PSNR metric.

For motion pictures we conducted subjective motion pictures quality assessment for Digital Cinema. The collected subjective data is used to analyse the performance of two compression algorithms (JPEG 2000 and AVC/H.264) for a Digital Cinema environment; the results showed that temporal compression schemes like H.264/AVC have high coding efficiency not only at SD resolutions, but also at high resolutions for Digital Cinema presentation.

For audio-visual presentations we performed subjective experiments of audiovisual contents for Digital Cinema. In addition, we investigated the multimodal effect on perceived quality in a Digital Cinema environment. As a major result we showed that the presence of audio (low or high quality) does not influence the visual quality judgement.

Modelling Perceptual Quality and Visual Saliency for Image and Video Communications

Author: Ulrich Engelke
Supervisor: Hans-Jürgen Zepernick, Blekinge Institute of Technology, Sweden
Published in 2010
URL: http://www.dissertations.se/dissertation/5b512c00ae/

This thesis focuses on subjective and objective assessment methods of perceived visual quality in image and video communication. The content of the thesis can be broadly divided into four parts.

Firstly, the focus is on the development of image quality metrics that predict perceived quality degradations due to transmission errors. The metrics follow the reduced-reference approach, thus, allowing to measure quality loss during image communication with only little overhead as side information. The metrics are designed and validated using subjective quality ratings from two experiments. The distortion assessment performance is further demonstrated through an application for filter design.

The second part of the thesis then investigates various methodologies to further improve the quality prediction performance of the metrics. In this respect, several properties of the human visual system are investigated and incorporated into the metric design. It is shown that the quality prediction performance can be considerably improved using these methodologies.
The third part is devoted to analyzing the impact of the complex distortion patterns on the overall perceived quality, following two goals. Firstly, the confidence of human observers is analyzed to identify the difficulties during assessment of the distorted images, showing that indeed the level of confidence is highly dependent on the level of visual quality. Secondly, the impact of content saliency on the perceived quality is identified using region-of-interest selections and eye tracking data from two independent subjective experiments. It is revealed, that the saliency of the distortion region indeed has an impact on the overall quality perception and also on the viewing behavior of human observers when rating image quality.

Finally, the quality perception of H.264/AVC coded video containing packet loss is analyzed based on the results of a combined subjective video quality and eye tracking experiment. It is shown that the distortion location in relation to the content saliency has a tremendous impact on the overall perceived quality. Based on these findings, a framework for saliency aware video quality assessment is proposed that strongly improves the quality prediction performance of existing video quality metrics.
Since human viewers are the target consumers for video communications products, they are naturally the most reliable source for quality assessment. However, gathering video quality assessment data from the human viewers is not a straightforward task, since it requires the completion of subjective quality assessment tests, which are hard to perform and cannot be used in real-time applications. An alternative to the subjective quality assessment procedures is to automatically score video quality using objective metrics.

The research work described in the Thesis is focused on the development of new algorithms for no-reference image and video quality assessment. Typical approaches for no-reference metrics usually try to estimate artifacts that result from lossy video encoding and/or from transmission losses. In this Thesis, a different philosophy is followed: the key idea is to estimate the quality of the encoded image and video data by firstly estimating local errors relatively to the original (unknown) media, and then weighting those errors using a perceptual model. This Thesis deals with the distortion caused by lossy encoding processes only, namely source coding and transcoding. The effect of transmission losses (i.e., packet losses in IP networks) has not been considered. Nevertheless, the ideas that are presented in the Thesis can be applied to a complete system, where transmission is also taken into account. The Thesis starts by proposing a technique based on watermarking, that evolves to a technique based on natural image statistics only. The results produced by the metrics are close to and well correlated with subjective quality assessment data.

**Video Quality Prediction for Video over Wireless Access Networks (UMTS and WLAN)**

Author: Asiya Khan  
Supervisor: Lingfen Sun, University of Plymouth, U.K.  
Published: December 2011  
URL: http://pearl.plymouth.ac.uk:8080/pearl_xmlui/bitstream/handle/10026.1/893/AKhan_thesis%5b1%5d.pdf?sequence=6

The success of video applications over wireless access networks very much depends on meeting the user's Quality of Service (QoS) requirements. Thus, it is highly desirable to be able to predict and to control video quality to meet user's QoS requirements.
Video quality is affected by distortions caused by the encoder and the wireless access network. The impact of these distortions is content dependent, but this feature has not been widely used in existing video quality prediction models.

The main aim of the project is the development of novel and efficient models for video quality prediction in a non-intrusive way for low bitrate and resolution videos and to demonstrate their application in QoS-driven adaptation schemes for mobile video streaming applications. This led to five main contributions of the thesis. The contributions are described in detail below the figure.

A thorough understanding of the relationships between video quality, wireless access network (UMTS and WLAN) parameters (e.g. packet/block loss, mean burst length and link bandwidth), encoder parameters (e.g. sender bitrate, frame rate) and content type is provided. An understanding of the relationships and interactions between them and their impact on video quality is important as it provides a basis for the development of non-intrusive video quality prediction models. A new content classification method was proposed based on statistical tools as content type was found to be the most important parameter. Efficient regression-based and artificial neural network-based learning models were developed for video quality prediction over WLAN and UMTS access networks. The models are light weight (can be implemented in real time monitoring), provide a measure for user perceived quality, without time consuming subjective tests. The models have potential applications in several other areas, including QoS control and optimization in network planning and content provisioning for network/service providers.

The applications of the proposed regression-based models were investigated in (i) optimization of content provisioning and network resource utilization and (ii) A new fuzzy sender bitrate adaptation scheme was presented at the sender side over WLAN and UMTS access networks.

Finally, Internet-based subjective tests that captured distortions caused by the encoder and the wireless access network for different types of contents was designed. The database of subjective results has been made available to research community as there is a lack of subjective video quality assessment databases.

The next issue of the Qualinet Newslet will be published on Oct.1, 2012.
The article „A Quest for the Definition of „Quality of Experience”“ will report on the challenge and the processes to define the term „QoE“. 