Impact of Frame Rate and Resolution on Objective QoE Metrics

Thomas Zinner*, Oliver Hohlfeld†, Osama Abboud‡, Tobias Hossfeld*

* University of Wuerzburg, Institute of Computer Science, Chair of Communication Networks, Wuerzburg, Germany
eMail: [zinner|hossfeld]@informtik.uni-wuerzburg.de

† Chair Intelligent Networks, Technische Universitaet Berlin / Deutsche Telekom Laboratories, Germany
eMail: oliver@net.t-labs.tu-berlin.de

‡Multimedia Communications Lab, Technische Universitaet Darmstadt, Germany
eMail: osama.abboud@kom.tu-darmstadt.de

***Abstract*—Video streaming applications are a major driver for the evolution of the future Internet. In this paper we introduce a framework for QoE management for video streaming systems based on the H.264/SVC codec, the scalable extension of H.264/AVC. A relevant feature is to control the user perceived quality of experience (QoE) by exploiting parameters offered by SVC. A proper design of a control mechanism requires the quantification of the main influence parameters on the QoE. For this purpose, we conducted a measurement study and quantified the influence of i) video resolution, ii) scaling method, iii) video frame rate and iv) video content types on the QoE by means of the SSIM and VQM full-reference metrics. Further, we discuss the trade-off between these different control knobs and their influence on the QoE.**

I. Introduction

In Next Generation Networks (NGNs), video streaming is expected to be the killer application dominating the traffic share worldwide. According to [10], Internet video will account for over 60 % of all consumer Internet traffic in 2013 and will generate over 18 exabytes per month. The user will demand high-quality image resolutions that may require bandwidths larger than what is supported in the current Internet architecture. Massive investments by network and service providers are one pathway to cope with the emerging challenges. In this paper, we motivate an alternative approach referred to as *QoE management* which will lead to much more economic and efficient use of the available resource while improving the quality of experience (QoE) for end users.

In particular, QoE management includes a) the monitoring of the current situation from the network's and the user's point of view as well as b) control mechanisms to dynamically adapt the video system to deliver the optimal QoE. The *monitoring of the current situation* aims at retrieving information about (i) the network environment of the user, like fixed FTTH or DSL access, or wireless WiMAX, WLAN or 3G access; (ii) the current network conditions like available end-to-end througphut or packet loss; (iii) user related parameters, like the capabilities of the end device (CPU power, resolution) or SLAs with the network or service operator; (iv) service and application specific information, like used video codec or type of content (sports clip, music clip, news, etc.). Similar investigations for the QoE have been conducted for voice and web traffic in [7], [9].

The work has been conducted in the project G-Lab, funded by the German Ministry of Educations and Research (Förderkennzeichen 01 BK 0800, G-Lab) and within the Internet Research Center (IRC) Wuerzburg. The authors alone are responsible for the content of the paper.

The *QoE control mechanism* takes into account the monitoring information and adjusts corresponding influence factors. For video streaming systems, the dynamic adapation of the video quality according to the current situation can be smartly realized with the de facto state of the art video codec H.264 and its scalable extension (H.264/SVC). This extension provides an integrated solution for different temporal, spatial and quality scalabilities and a seamless switching between, e.g., different resolutions or frame rates. The concept enables an adoption of the delivered video quality to the available bandwidth. In case of network problems like congestion, the resolution, and thereby the necessary bandwidth, could be reduced in order to avoid packet loss and the emerging video quality degradation. In this context, the question occurs how the end user perceives the actual quality of the delivered video. In particular, we investigate, based on objective metrics, if a user is more satisfied with i) a low resolution, but a smooth video playout or ii) a high resolution at the cost of quality degradations due to packet loss in the network or a lower frame rate.

Figure 1 illustrates QoE regions acceptable from a users perspective in a spider plot. The different axes denote the influence of the control knob settings provided by SVC on the QoE. A highly sophisticated QoE control mechanism determines the optimal settings in terms of best QoE with minimal costs. The focus of the paper is the quantification of the acceptable area, i.e. of different influence factors and control knob settings on the QoE and the required resources (bandwidth, CPU). This is mandatory in order to design appropriate QoE control mechanisms. We rely on objective QoE metrics like SSIM and VQM which allow to conduct measurement studies and

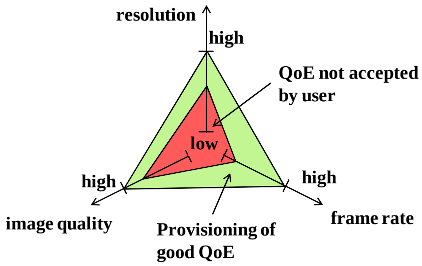


Fig. 1. Acceptable area of QoE control knob settings

to derive simple relationships applicable in QoE control. In particular, we take a closer look at the influence of content type of the video, its resolution, as well as the scaling method and the frame rate. The scaling method is an additional control knob on application layer. User prefer to watch a video clip in an adequate size [12], that means, they will scale up the video, if possible, to be displayed on full screen. For resizing, the various interpolation algorithms differ in their computational complexity at the user device as well as the achieved video quality.

The main contribution of the paper is twofold. First, we conduct extensive measurement studies for the quantification of the QoE for different control knobs (video resolution, scaling method and frame rate) and video content types. Second, we correlate these results with the used bandwidth in order to characterize the quality degradation caused by bandwidth reduction. The measurement results extend results investigated in [23] and are a further step to define thresholds for a highly sophisticated QoE control, e.g. for a p2p based VOD system as proposed in [14].

The remainder of this paper is structured as follows. Section II gives comprehensive background on scalable video coding and existing work which links main influence factors of different video dimensions on the user perceived quality. In addition, metrics for quantifying QoE are briefly introduced. By means of laboratory measurement studies, we quantify the influence factors on QoE which allow to design control mechanism in the QoE management framework. The measurement methodology is discussed in Section III, while the results are presented in Section IV. Finally, Section V concludes this work and gives an outlook on future work.

II. Background

A. Scalable Video Coding

The video codec H.264/SVC, cf. [11], [16], is based on H.264/AVC, a video codec used widely in the Internet, for instance by video platforms (e.g., YouTube, GoogleVideo) or video streaming applications (e.g., Zattoo). H.264/AVC is a so called single-layer codec, which means that different encoded video files are needed to to support heterogeneous end user devices. The Scalable Video Coding (SVC) extension of H.264/AVC enables the encoding of a video file at different qualities within the same layered bit stream. This includes besides different resolutions also different frequencies (frames displayed per second) and different qualities w.r.t. Signal-to-Noise Ratio (SNR). Different qualities can be considered as a special case of spatial scalability with identical picture size for base and enhancement layers. These three dimensions are denoted to as spatial, temporal and quality scalability. Figure 2 gives an example of different possible scalabilities for a video file. The scalable video file can be watched in three different temporal resolutions (15Hz, 30Hz, 60Hz), three different spatial resolutions (CIF, SD, HD) and three different quality resolutions (Q0, Q1, Q2). The left bottom "subcube", CIF resolution with 15 Hz and quality Q0, is the base layer which is necessary to play the video file. Based on this layer different enhancement layers permit a better video experience with a higher resolution, better SNR or higher frame rate, respectively. The more subcubes along any of the three axes are available the higher the quality in this respect is. If all subcubes are available the video can be played back in highest quality. If all subcubes within quality Q0 are available, the video can be played back in HD-resolution with 60 Hz, but only with a low SNR. This concepts allows an adaptation of the video quality to the service parameters, for instance to the connection throughput. Due to the integration of different layers within one video file a seamless switch between different layers is possible. Thus, the bandwidth of a video stream may be adjusted to the network conditions. If the offered end-to-end throughput is not sufficient for playing back the video file in maximum quality it is possible to reduce the delivered frame rate, image quality or resolution. Therefore the bandwidth of the video stream can be reduced, that means the bandwidth is adopted to the offered network quality of service parameters. The influence of a bandwidth reduction on the user perceived quality is discussed in the next subsection. It has to be noted that no bandwidth adaption to network conditions leads to artifacts or missing frames in case of UDP-based video streaming or to stalling of the video in case of TCP.

B. Main Influence Factors of Different Video Dimensions on the User Perceived Quality

When using a video codec like H.264, bandwidth reduction is usually achieved by one of the following ways (i) reduce the

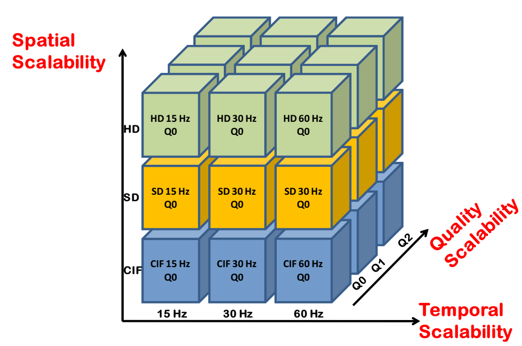


Fig. 2. SVC Cube, illustrating the possible scalability dimensions for a video file

image resolution of the video, (ii) decrease the image quality due to higher image compression rates (larger quantization), or (iii) reduce the frame rate (fps). Recently user surveys have been conducted investigating the impact of these influence factors on the subjective quality of digital video, especially in the context of mobile environments. In [5] Buchinger et al. described the interconnection between the compression rate and the frame rate in mobile environments. It turns out, that, for a given resolution, users prefer a video higher image quality, i.e. lower compression rate, and low frame rate instead of a video with medium picture quality and high frame rate. Similar investigations have been carried out by McCarthy et al., cf. [13]. For their experiments they showed test videos on desktop computers and palmtops in two different resolutions, 352x244 for the desktop experiments and 176x144 for the palmtop experiments. The conducted surveys confirm, that users tend to neglect a reduction of the frame rate, but that a decrease of picture quality leads to dissatisfied users. Our work differs from the approaches mentioned above, since we investigate H.264 encoded video sequences with higher resolutions. Further, we use objective metrics for determining the Quality of Experience instead of subjective ones like MOS or acceptability.

The issue how a video clip with given resolution should be displayed on the screen is discussed by Knoche [12]. He discovered that the video has to be displayed at an adequate size. This includes that users prefer to resize the video picture to the highest possible size with still a sufficient image quality. Similar findings are discussed by Bae et al [4]. They investigated that human observers are willing to accept some visible distortion in order to obtain higher resolutions. His work does not include an investigation of resizing mechanisms. Usually, this is done by the player either with simple mechanisms like nearest neighborhood interpolation or more sophistic mechanisms like cubic or bicubic polynomial interpolation. The more complex the resizing algorithm, the more expensive is it in terms of CPU and energy consumption. On the other hand a more complex algorithm might increase the user perceived quality. An investigation of this issue is performed with full reference models in Section IV-A. Yamagishi [22] discuss the influence of the coded bit rate on the video quality. This differs from our contribution since we detail the different scalability mechanisms provided by H.264/SVC.

C. Quantifying Quality of Experience

Quality of Experience is defined as the subjectively perceived acceptability of a service [6]. The perceived quality can be investigated in subjective tests, where presented stimuli—such as impaired video sequences—are rated by subjects under controlled conditions. The obtained rating expresses the *subjective Quality of Experience* (sQoE), typically described by the Mean Opinion Score (MOS).

However, subjective tests are time-consuming, expensive and have to be undertaken manually, which does not allow for automatic quality ratings by software. This aspect motivates objective metrics, which are designed to correlate with human perception, and, thus avoid cost and time intensive empirical evaluations. Estimates for the quality obtained by metrics are called *objective Quality of Experience* (oQoE). A more comprehensive discussion on this subject can be found in [18].

TABLE I
MAPPING OF OQOE TO SQOE

MOS	SSIM	VQM
5 (excellent)	> 0.99	< 0.2
4 (good)	≥ 0.95 & < 0.99	≥ 0.2 & < 0.4
3 (fair)	≥ 0.88 & < 0.95	≥ 0.4 & < 0.6
2 (poor)	≥ 0.5 & < 0.88	≥ 0.6 & < 0.8
1 (bad)	< 0.5	> 0.8

Quality metrics can be classified into three categories by the required amount of reference information [21]: *Full-reference* (FR) metrics are based on frame-by-frame comparison between a reference video and the video to be evaluated; *No-reference* (NR) metrics have to make assumptions about the video content and distortions, e.g. by evaluating the blockiness of a frame, as a common artifact in block-based compression algorithms such as MPEG; *Reduced-reference* (RR) metrics evaluate the test video based on a subset of features previously extracted from the reference video. Based on the complex nature of cognitive aspects and the human visual system, objective quality metrics do not capture its entire complexity and focus on aspects, which have been shown to correlate well with human perception in subjective tests. Thus, they are biased by model limitations and limited in their performance.

In this paper, we focus on two full reference metrics, SSIM and VQM, due to the availability of the unimpaired reference video in laboratory conditions.

In principle, the subjective judgement of video quality depends on factors such as content (e.g. interview, soccer match, movie) or context (e.g. viewed on a mobile device, HDTV capable screen).

The Structural Similarity Index Metric (SSIM) [19] introduced by Wang et al. is motivated by the assumption that human visual perception is highly adapted for extracting structural information. It has been shown to have a high correlation with image [19] and video quality [20].

The Video Quality Metric (VQM) [15] is a standaradized method of objectively measuring video quality. Validation tests by the International Video Quality Expert's Group (VQEG) showed a very high correlation between subjective user surveys and the objectively measured results. The VQM method is adopted as international ITU Recommendations ITU-T J.144 and ITU-BT.1683 since 2004.

According to Seshadrinathan et al, cf. [17], VQM has a higher correlation to the subjective Quality of Experience than SSIM. Nevertheless we use both metrics for our investigation, since results can be generated much faster with SSIM than with the VQM reference software, cf. [1].

Based on results obtained for still images in [19] and for videos in [15], we introduce a mapping of SSIM and VQM (oQoE) to a nominal 5-point MOS scale (sQoE) according to Table I for expressing an approximation of sQoE.

III. Measurement Methodology

This section describes the used video sequences, how this video sequences were encoded and which quality estimation softwares we used.

A. Video Clips and Encoding

As video clips we used blue sky, crowd run and park joy, cf. Table II, in y4m format with a resolution of 1080p, provided by xiph.org [3]. We encoded the videos in H.264/SVC with the JSVM software Version 9.15 in different spatial and temporal layers. The base layer comprises a resolution of 480x270 pixels with a frame rate of 1.875 frames per second (fps). This layer is extended by several temporal extension layers with 3.75, 7.5, 15 and 30 fps and by three spatial enhancement layers increasing the resolution to 640x360, 960x540 and 1216x684 pixels. The maximum video quality of 1216x684 pixels with 30 fps is achieved if all layers are available. We computed the average bandwidth requirements for each of the different layers. Based on these requirements we discuss the trade-off between bandwidth and estimated video quality in section IV-C.

B. Quality Estimation

As models for evaluating the user perceived quality we used SSIM and VQM metric. An efficient implementation of SSIM is provided by the MSU Video Quality Measurement Tool [8], which we used during the course of this work. The tool evaluates the SSIM values on a frame-by-frame basis; we use the average over all frames as indicator for the video quality. For computing the VQM value we used the ITS Video Quality Metric Software [1]. The different spatial resolutions were computed using the uncompressed y4m files and Virtual Dub [2]. In order to compare the different resolutions, we scaled up the videos to the reference resolution, using both, bicubic or nearest neighborhood interpolation. In order to compare different frame rates we emulated the lower frame rate by removing a temporal layer and refilling the missing frames with predecessor frames, cf., Fig. 3. The video depicted in Fig. 3 has a frame rate of n fps, whereas the black frames contribute a frame rate of $n/2$ fps. The same video clip with $n/2$ fps consists of half the number of frames than before, i.e. half of the frames are not displayed. The other frames are displayed longer on the screen, and the duration of the clip remains constant.

This approach allows to approximate the impact of a frame rate reduction on objective metrics like SSIM without investigating aspects like jerkiness that are not modeled in the respective metric. This provides only a rough approximation and does not completely reflect the subjective user experience, as aspects like jerkiness are not considered. Nevertheless, this investigation allows us to get first insights into the impact of this control knob on the video quality degradation as perceived by popular objective QoE metrics.

We compared each of the encoded SVC layers and several intermediate layers with the highest possible quality, i.e. 1216x684 pixels with 30 fps assuming that a user requests this quality and uses an equivalent display for playing back the video clip.

TABLE II
Properties of reference sequences

Name	blue sky	crowd run	park joy
# Frames	216	499	499
Frame rate	30	30	30
Average bandwidth (Mbyte/s)	0.82	1.54	1.85
Length (sec)	7.2	16.63	16.63
Motion type	low-medium	medium	medium

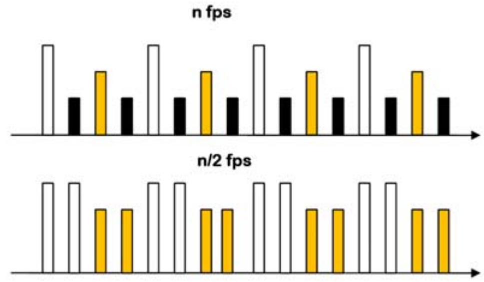


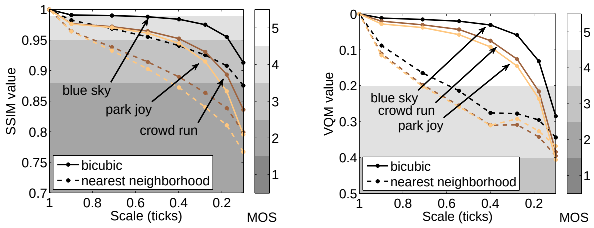
Fig. 3. Video with different frame rates

IV. Results

In this section we discuss the results obtained with our experiments.

A. Objective Quality of Experience for Different Resolutions

Now we want to discuss the influence of lower resolutions on the user perceived quality of experience. As additional influence we investigate this behavior for different interpolation mechanisms, nearest neighborhood and bicubic interpolation. The results are illustrated in Fig 4 for the objective quality estimators SSIM and VQM mapped to the subjective MOS scale. For both figures, the x-axis shows different resolutions, whereas the scale of the axis is proportional to the number of pixels of each resolution. For instance, a scale of 0.5 denotes half of the maximum resolution resulting in 860x484 pixels. The y-axis denotes the used objective metric, and the gray colored areas illustrate the corresponding MOS values. A darker gray scale indicates a lower MOS value, implying a lower quality. The dashed lines show the results for nearest neighborhood interpolation, the solid lines for bicubic interpolation. Fig 4(a) depicts the SSIM and the corresponding MOS value. It can be seen, that, regardless of the content, a decrease of the resolution yields in a decrease of the SSIM metric and accordingly to the MOS value. Further, it can be seen, that nearest neighborhood interpolation performs always worse than bicubic interpolation. Similar results are obtained for the VQM metric which is displayed in Fig 4(b). The VQM results differ from the results obtained with SSIM in two points. First, VQM does not differentiate the investigated interpolation types for small resolutions, i.e. a resolution of



(a) SSIM (b) VQM

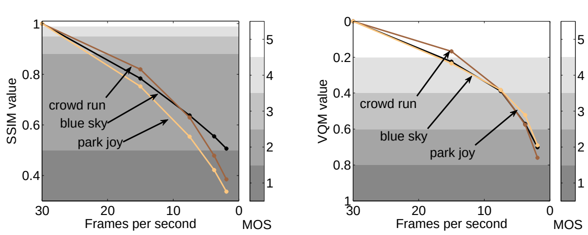
Fig. 4. Objective comparison of different resolution pairs

384x216 pixels. Further VQM estimates the user perceived quality better than SSIM. While SSIM classifies low resolutions with MOS values ranging from poor to fair (2-3), VQM classifies these resolutions with MOS value of good (4). For high resolutions, the differences between both metrics are not as significant. We can conclude, that both objective metrics can be used to classify the degradation of video content in case of lower resolutions. However, there are significant differences in the qulaity estimation of both metrics. Thus, a subjective assessment for the investigated scenarios is needed in order to link the objective metrics to the real user perceived quality. Both metrics estimate different upscaling mechanisms unequal and show similar behavior for the conducted study. Additionally, we have seen that there is a noticeable influence of the content on the oQoE metrics, i.e. if we want to control the user perceived quality we have to take the content into account.

B. Objective Quality of Experience for Different Frame Rates

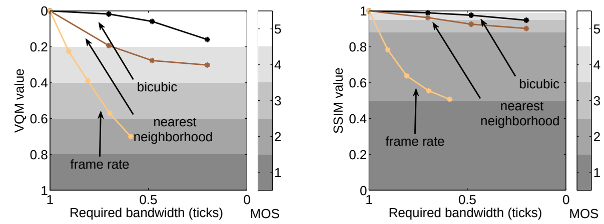
This subsection deals with the influence of lower frame rates on the user perceived quality, cf. Section III-B. The results for the objective metrics and the corresponding subjective MOS values are depicted in Fig 5. In both subfigures, the y-axis denotes the objective QoE value, while the different MOS values are illustrated by the areas with different gray colors.

The SSIM value, depicted in Fig 5(a) decreases with a lower frame rate. The same holds for the VQM metric, which is depicted in Fig 5(b). However, halving the frame rate to 15 frames per second yields to MOS value fair (3) for SSIM and to a MOS value of good (4) or better in case of VQM. The behavior, that SSIM estimates the influence of lower frame rates on the user perceived quality worse than VQM holds



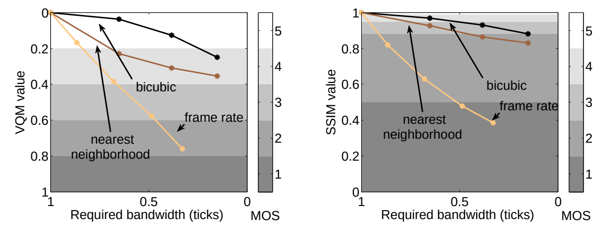
(a) SSIM (b) VQM

Fig. 5. Objective comparison of different frame rates



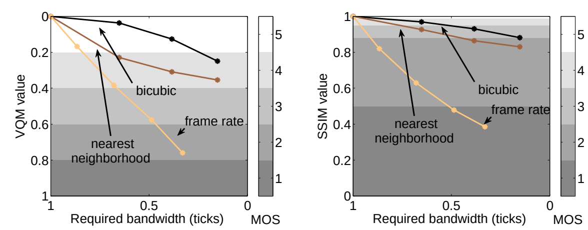
(a) VQM (b) SSIM

Fig. 6. Video clip blue sky



(a) VQM (b) SSIM

Fig. 7. Video clip crowd run



(a) VQM (b) SSIM

Fig. 8. Video clip park joy

for the other investigated frame rates too. Nevertheless the difference in terms of computed MOS value is smaller than compared to the evaluation for different resolutions in IV-A. Additionally it has to be noticed, that the content has an influence on the estimated oQoE metric although the difference in terms of the used metric is not as significant as compared to the influence in case of different resolutions.

C. Trade-Off between Bandwidth, Frame Rate and Resolution

In this section we evaluate the trade-off between required bandwidth, video frame rate and video resolution. For this case study we rely on VQM and SSIM. The Figures 6, 7 and 8 discuss this trade-off for three different video clips for both objective quality of experience metrics.

In all figures, the x-axis displays the bandwidth savings referring to the bandwidth of the maximum quality. The y-axis displays the VQM/SSIM value, and the different gray areas the mapping to MOS value.

The results for the video clip blue sky are depicted in Fig 6. For VQM, cf. Fig 6(a), it can be seen that the quality degradation in terms of VQM is smaller for lower resolutions than for lower frame rates, independent from the interpolation mechanism. This also holds for this video clip in case of an evaluation with SSIM, cf. Fig 6(b). Equal results can be seen for the video clip crowed run, cf., Fig 7, and for the video clip park joy, cf., Fig 8. It should be noted, that the SSIM metric again estimates the quality degradation worse than the VQM metric.

These results suggest, that bandwidth savings for the investigated scenarios have to be achieved by reducing the resolution. In case of a frame rate reduction less bandwidth could be saved and the video experience would be disturbed to a greater extend. These findings can be used to control the user perceived quality of a p2p video on demand system using H.264/SVC as proposed in [14].

However, the obtained measurements have to be conducted with other content and realistic longer video clips. Further, a subjective assessment of the investigated scenarios has to be carried out in order to validate the results and link the objective metrics to real user perceived quality for the discussed use-cases.

V. Conclusion

In this paper we proposed a framework for QoE mamagement for content distribution systems based on H.264 SVC. Bandwidth adaption with SVC can be achieved seamlessly in three ways: (i) reducing the video resolution, (ii) decreasing the image quality, or (iii) reducing the frame rate. Additionally to these three control knobs the scaling method on application layer has to be taken into account as control knob. We conducted a measurement study to quantify the *objective Quality of Experience* and linked the results to corresponding *MOS* values. For that, we focused on the scaling method for different resolutions and on the impact of different frame rates. We showed that objective metrics are able to distinguish between quality levels for different resolutions and frame rates.

Second, our results show, that SSIM and VQM can be used to quantify the behavior of different content on the QoE. Further, we showed that video sequences with lower resolution perform better than video sequences with lower frame rate with respect to VQM. Future work will extend this investigation with respect to higher qualities and higher frame rates of the video content. Additionally longer video sequences and other types of video sequences have to be evaluated. Also user surveys have to be conducted in order to validate the results and for linking the objective metrics to the real user perceived experience. The obtained results will be combined with the present results in an overall QoE control mechanism for SVC based on the integrated SVC scalabilities. This control mechanisms will permit an adaption of the video content to the network conditions and allow a maximization of the user perceived quality.

Acknowledgments

We deeply want to thank Benjamin Kleine, Dirk Rauscher, Michael Lang, Michael Denkler and Hans Henjes for their assistance during the course of this work.

References

[1] Video quality metric software. http://www.its.bldrdoc.gov/vqm/.

[2] Virtual dub. http://www.virtualdub.org/.

[3] Xiph.org test media. http://media.xiph.org/video/derf/.

[4] Soo Hyun Bae, Thrasyvoulos N. Pappas, and Biing-Hwang Juang. Subjective evaluation of spatial resolution and quantization noise tradeoffs. *Trans. Img. Proc.*, 18(3):495–508, 2009.

[5] S. Buchinger and H. Hlavacs. Subjective quality of mobile mpeg-4 videos with different frame rates. *Journal of Mobile Multimedia (JMM)*, 1(4):327–341, 2006.

[6] ITU-T Delayed Contribution D.197. Definition of Quality of Experience, 2004.

[7] M. Fiedler, T. Hoßfeld, and P. Tran-Gia. A Generic Quantitative Relationship between Quality of Experience and Quality of Service. Submitted, 2009.

[8] Graphics and Moscow State University Media Lab, CMC department. Net:netem. http://www.linuxfoundation.org/en/Net:Netem.

[9] T. Hoßfeld, P. Tran-Gia, and M. Fiedler. Quantification of Quality of Experience for Edge-Based Applications. In *20th International Teletraffic Congress (ITC20)*, Ottawa, Canada, jun 2007.

[10] Cisco Systems Inc. Cisco Visual Networking Index: Forecast and Methodology, 2008-2013. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360.pdf, June 2009.

[11] ITU-T Rec. & ISO/IEC 14496-10. Advanced Video Coding for Generic Audiovisual Services.

[12] H. Knoche and M. Sasse. The Big Picture on Small Screens Delivering Acceptable Video Quality in Mobile TV. *ACM Trans. Multimedia Comput. Commun. Appl.*, 5(3):1–27, 2009.

[13] J. McCarthy, M. Sasse, and D. Miras. Sharp or Smooth?: Comparing the Effects of Quantization vs. Frame Rate for Streamed Video. In *CHI '04: Proceedings of the SIGCHI conference*, pages 535–542, New York, NY, USA, 2004. ACM.

[14] Simon Oechsner, Thomas Zinner, Jochen Prokopetz, and Tobias Hoßfeld. Supporting Scalable Video Codecs in a P2P Video-on-Demand Streaming System. In *21th ITC Specialist Seminar on Multimedia Applications - Traffic, Performance and QoE*, Miyazaki, Jap, March 2010.

[15] M.H. Pinson and S. Wolf. A new standardized method for objectively measuring video quality. *IEEE Transactions on broadcasting*, 50(3):312–322, 2004.

[16] H. Schwarz, D. Marpe, and T. Wiegand. Overview of the scalable video coding extension of the h.264/avc standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(9):1103–1129, 2007.

[17] Kalpana Seshadrinathan, Rajiv Soundararajan, Alan C. Bovik, and Lawrence K. Cormack. A subjective study to evaluate video quality assessment algorithms. volume 7527, page 75270H. SPIE, 2010.

[18] B. Staehle, A. Binzenhöfer, D.Schlosser, and B. Boder. Quantifying the influence of network conditions on the service quality experienced by a thin client user. In *14. GI/ITG MMB*, Dortmund, Germany, apr 2008.

[19] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13:600–612, 2004.

[20] Z. Wang, L. Lu, and A. Bovik. Video quality assessment based on structural distortion measurement. *Signal Processing: Image Communication*, 19(1), January 2004.

[21] S. Winkler. Video Quality and Beyond. In *Proc. European Signal Processing Conference*, 2007.

[22] K. Yamagishi and T. Hayashi. Parametric packet-layer model for monitoring video quality of iptv services. In *ICC*, pages 110–114, 2008.

[23] Thomas Zinner, Osama Abboud, Oliver Hohlfeld, and Tobias Hossfeld. Towards QoE Management for Scalable Video Streaming. In *21th ITC Specialist Seminar on Multimedia Applications - Traffic, Performance and QoE*, Miyazaki, Jap, 2010.