

## PERCEPTUAL QUALITY ASSESSMENT OF HIGH FRAME RATE VIDEO

*Rasoul Mohammadi Nasiri, Jiheng Wang, Abdul Rehman, Shiqi Wang and Zhou Wang*

Dept. of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1, Canada  
Emails: {r26moham, j237wang, abdul.rehman, s269wang, zhou.wang}@uwaterloo.ca

### ABSTRACT

High frame rate video has been a hot topic in the past few years driven by a strong need in the entertainment and gaming industry. Nevertheless, progress on perceptual quality assessment of high frame rate video remains limited, making it difficult to evaluate the exact perceptual gain by switching from low to high frame rates. In this work, we first conduct a subjective quality assessment experiment on a database that contains videos compressed at different frame rates, quantization levels and spatial resolutions. We then carry out a series of analysis on the subjective data to investigate the impact of frame rate on perceived video quality and its interplay with quantization level, spatial resolution, spatial complexity, and motion complexity. We observe that perceived video quality generally increases with frame rate, but the gain saturates at high rates. Such gain also depends on the interactions between quantization level, spatial resolution, and spatial and motion complexities.

**Index Terms**— high frame rate video, video quality assessment, video compression, spatial complexity, motion complexity

### 1. INTRODUCTION

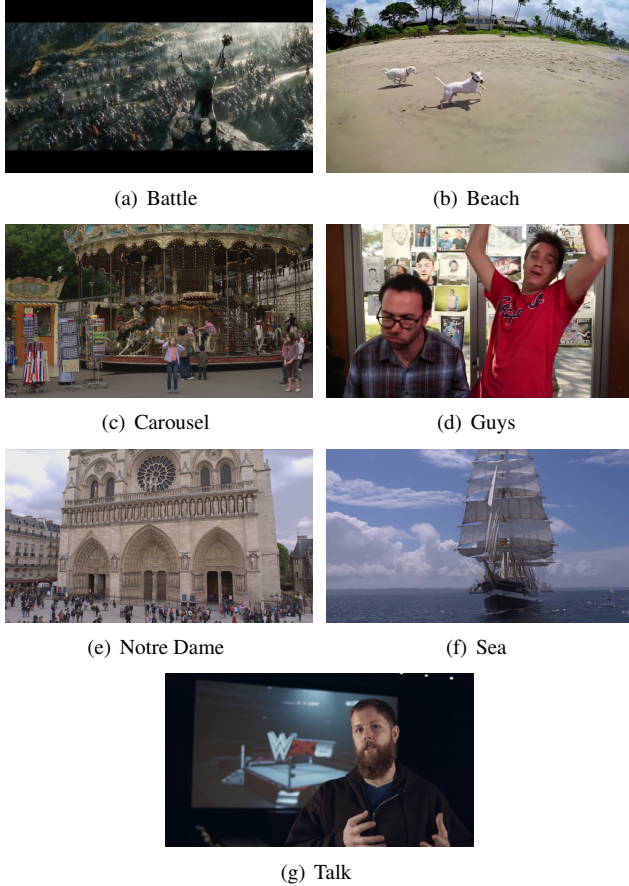
Frame rate refers to the number of frames displayed by a viewing device in one second. For example, the current standard in cinemas worldwide is 24 frames per second (fps). It is commonly believed that higher frame rates such as 48fps, 60fps, or 120fps can provide enhanced visual clarity and reduce flickering, stuttering, and motion blur during action scenes. Nevertheless, in the literature, the progress on high frame rate (HFR) video quality assessment (VQA) remains limited, making it difficult to judge the actual perceptual gain of switching to HFR video. Previous works on video frame rate mostly focused on the lower range of 5fps to 30fps. In [1], a comprehensive review of the effects of different frame rates on human performance was conducted. In [2], five videos were compressed using H.263 and H.264 encoders at bit rates ranging from 24k to 382k bits/second(bps) with frame sizes of QCIF and CIF and frame rates of 7.5 to 30 fps, followed by an extensive subjective test. An interesting observation was that for an optimal combination of frame rate

and frame size under a low bit-rate constraint, small frame size is often preferred while frame rate should typically be kept low (high) for video sequences with high (low) temporal activity. In [3], the impact of frame rate (30, 15, 7.5, and 3.75 Hz) and quantization level (QP = 28, 36, 40, and 44) on perceived video quality has been investigated. The subjective testing results indicate that at all QP levels, the mean opinion score (MOS) reduces consistently as the frame rate decreases and such a reduction in MOS is quite independent of the quantization level. In [4, 5, 6], computational VQA models considering both frame rate and quantization artifacts or frame resolutions have been proposed where all algorithms focused on a maximum frame rate of 30fps.

Some existing research focused on videos of different frame rate with special conditions. Low bit-rate videos were studied in [7, 8, 9], where in most cases the low bit rate videos have frame sizes of CIF or QCIF. A list of video quality models and their supported spatial and temporal resolutions can be found in [10]. In [11], various scalability issues, including temporal scalability, were discussed in the context of video coding. In [7, 12], temporal resolution was considered as a quality factor in specific applications of video broadcasting and distribution over the network.

Only a few works were devoted to the effect of frame rates beyond 30fps. In [13], the effects of frame rate (60, 15, 7.5, and 5 Hz) and spatial resolution on users playing First Person Shooter games were investigated. It was concluded that frame rate has a marked impact on both player performance and game enjoyment while spatial resolution has little impact on performance but moderate impact on enjoyment. Recently, a subjective study on frame rates up to 60fps for 3D videos were reported in [14, 15], where only specific indoor scenarios were chosen as the source content. In [16], a quality evaluation method has been proposed for quality of video in scalable coding application, where frame rates up to 50fps have been considered while the test videos content are limited. In general, systematic studies on subjective and objective quality assessment of HFR videos are still lacking, making it difficult to derive VQA models that have the potential to be generalized to all frame rate levels to guide the compression and delivery of HFR video content.

In this work, we first conduct a subjective quality assessment experiment on a database that contains videos com-



**Fig. 1.** Sample frames from the pristine videos used in the subjective study.

pressed at different frame rates, transform-domain quantization levels and spatial resolutions. We then carry out a series of analysis on the subjective data to investigate the impact of frame rate on perceived video quality and its interaction with quantization level, spatial resolution, spatial complexity, and motion complexity.

## 2. SUBJECTIVE STUDY

### 2.1. Video Database

The new Waterloo-IVC High Frame Rate Video Quality Database is created from 7 pristine 60fps source videos as shown in Fig. 1. The specifications of the test videos are given in Table 1. Previous work [15] suggested that the impact of high frame rate on video quality depends on spatial and temporal complexity. In this work, the raw video sequences are selected to represent different combinations of spatial content, object motion and camera motion. The details are given in Table 1.

The pristine video is compressed using an H.264 encoder at different quantization levels with various frame rates and

**Table 1.** Details of the source videos.

Resolution	Frames/Second	Length	Color Format
1920 × 1080	60FPS	10s	YUV4:2:0
Sequence	Object Motion	Camera Motion	Spatial Complexity
Battle	High	Yes	High
Beach	High	Yes	Low
Carousel	Medium	No	High
Notre Dame	Medium	Yes	High
Guys	High	No	Low
Sea	Low	Yes	Low
Talk	Low	No	Low

frame sizes. The details of all control parameters are given in Table 2. Two different resolutions are used: 480p and 1080p. 480p represents standard definition (SD) formats and 1080p is the most common High Definition (HD) format supported by all HDTV display devices. The 480p format has been generated from 1080p by down-sampling followed by bicubic interpolation.

In the database, different frame rates from low frame rate of 5fps to high frame rate of 60fps were generated for all combinations of quality, resolution, and test video content. The values of frame rate have been selected based on different needs. 30fps is a common frame rate in many current applications. 15fps and two other lower frame rates are often used to support lower bit-rate encoding as a compromise for limited storage space or transmission bandwidth. 60fps is the most common high frame rate being used. 45fps is the middle frame rate that is included to make a better spaced temporal resolution in the subjective test. Different frame rate has been generated using FFmpeg tool using dripping and duplicating method for generating different frame rates. Four different quality parameters have been used in order to cover different levels of compression artifact from low compression of QP=22 to high compression of QP=37. As such, for each source content, there are 6 (Frame rate) × 4 (QP) × 2 (Resolution) = 48 test video sequences. Altogether, there are totally  $48 \times 7 = 336$  video sequences in the database.

**Table 2.** Configurations used to generate test videos from the source video.

Parameter	Values
Frame Rate	5,10,15,30,45,60
Quantization Level	22, 27, 32, 37
Frame Size	640 × 480, 1920 × 1080

### 2.2. Remarks on the Database

There are three important features of the database. First, the database contains sequences with a wide range of frame rates from 5fps to 60fps, which allows us to directly examine the general trend of the impact of frame rate on perceptual video quality. The better coverage of the frame rates makes the database better suited to study a wider range of practical applications, and to better observe the general trend of quality variations as a function of frame rate that could be extrapolated.

lated beyond the frame rates currently being tested. Second, the database contains sequences with different combinations of spatial complexity, object motion, and camera motion, allowing us to study the interactions between frame rate and video content. Third, the database contains sequences with different compression levels and frame sizes, allowing us to investigate the trade-offs between frame rate, compression level, and spatial resolution.

Compared with the new database, existing databases in the literature are limited in one aspect or another. In [17], the authors attempted to consider time complexity with motion, but only videos with low spatial resolution ( $352 \times 240$ ) and frame rates (up to 30fps) were used. Similarly, in [2, 3, 5, 4, 18], only small resolution videos (CIF or QCIF size) were employed. In [9] only low bit rate videos are considered, which are not able to cover the HD cases where the bit rates are often much higher. In [14, 15], 60fps videos were studied, but the impact of spatial and temporal complexities on video quality was investigated separately, making it impossible to study the combined effect of complexities as well as variations in video content and quantization levels. In [19], the effects of quantization and frame rate were studied while the dimension of spatial resolution and content complexity were missing, making it difficult to build or test a complete model.

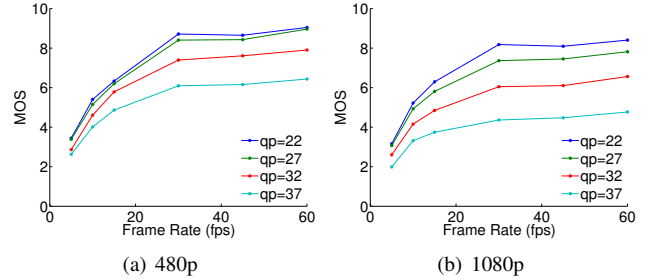
### 2.3. Subjective Test

The subjective test was conducted in the Lab for Image and Vision Computing at University of Waterloo. The test environment has no reflecting ceiling wall or floor, and was not insulated by any external audible and visual pollution. An ASUS LED monitor was used for the test. The details of the viewing conditions are given in Table 3. Twenty-five naïve subjects, 13 males and 12 females aged between 22 and 33, participated in the study.

**Table 3.** Viewing conditions of the subjective test.

Parameter	Value
Subjects Per Monitor	1
Screen Resolution	1920 × 1080
Screen Diameter	31.5"
Viewing Distance	30.00"
Screen Width	27.45"
Viewing Angle	49.2° H/28.9° V
Screen Height	15.44"
Pixels Per Degree	78.1/74.8 pixels(H/V)

The subjects were asked to evaluate their overall viewing experience – Video Quality (VQ) in this study. A single stimulus, 11-grade numerical categorical scale (SSNCS) protocol was employed. A general introduction was given at the beginning of the whole test, and more specific instructions and a training session were given afterwards. The video content of the training videos is similar but different from those in the formal test session. The parameters used to generate the training videos are also similar to the test video parameters.



**Fig. 2.** MOS versus frame rate for all test videos.

The subjects were asked to rate training videos until they fully understood the requirements and stabilized their rating strategies.

All stimuli were displayed in actual pixels, and in the case of 480p sequences, display regions outside the frames were filled with black pixels. A still gray image was displayed for 7 seconds after each test video for subject scoring. Each stimulus was shown once and the order of stimuli was randomized. Eighty-four videos were evaluated in one session. To reduce visual fatigue, each session was controlled to be within 20 minutes and sufficient relaxation periods (5 minutes or more) were given between sessions. The MOS for each test video was computed using scores of all users.

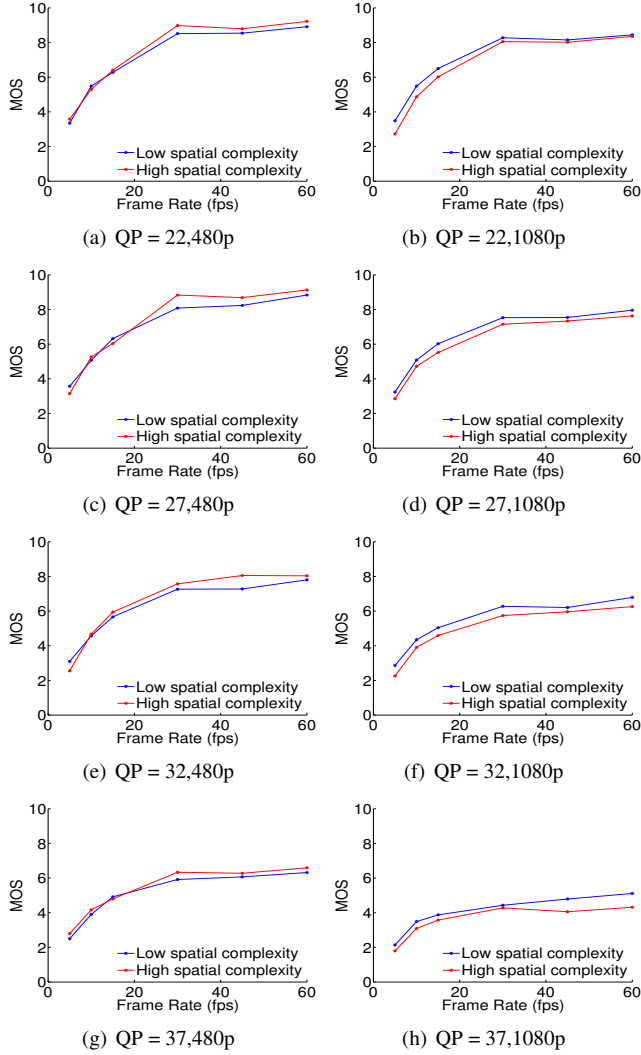
The rest of the paper focuses on the impacts of frame rate on perceived video quality with different quantization levels, different frame sizes, and different complexities of spatial content and motion. More detailed descriptions of the database and analysis of the other aspects of the subjective experiments will be reported in future publications.

## 3. KEY OBSERVATIONS

Based on the subjective test results, we have carried out a series of statistical analysis. Below we focus on a few main observations.

### 3.1. General trend of quality vs. frame rate

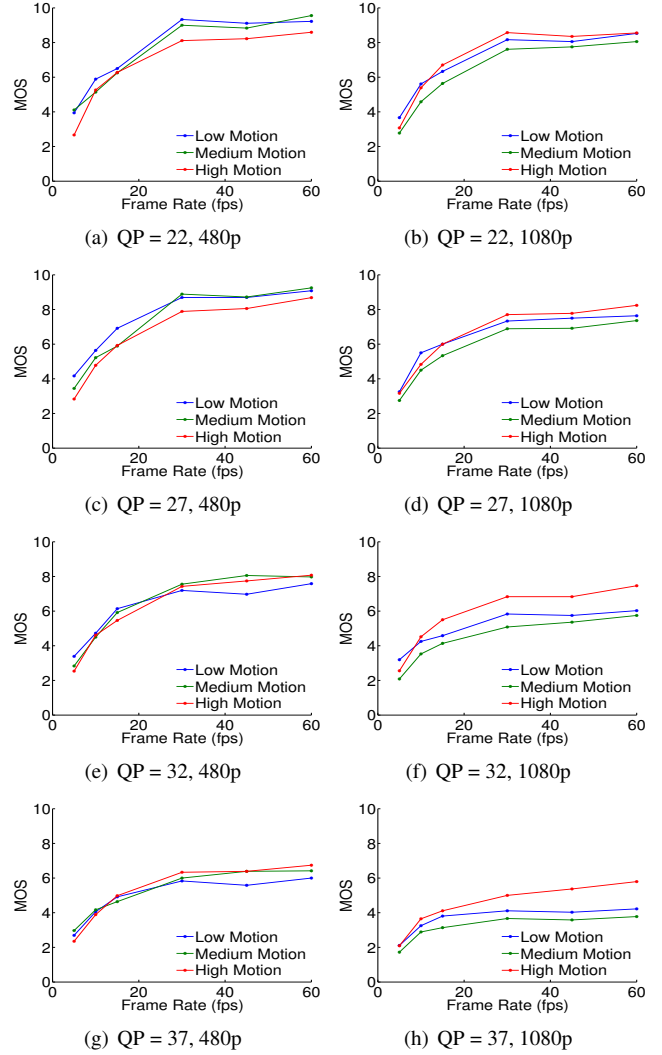
Fig. 2 shows the MOS values for all source sequences with respect to different quantization levels (QP values) and different frame sizes (480p or 1080p). It can be observed that there is a significant improvement in terms of MOS values from 5fps to 30fps, which is consistent with previous results [3, 15]. Such improvement decreases with increasing frame rate, especially after 30fps. Even though small, the improvement from 30fps to 60fps can still be clearly discerned, which justifies the value of going beyond 30fps. However, the general trend being observed here suggests that the quality improvement saturates at high frame rates, thus increasing frame rate beyond 60fps may not lead to distinguishable quality gain, depending on video content. Scrupulous observers may find that the improvement from 30fps to 45fps seems to be below expectation from the



**Fig. 3.** MOS versus frame rate for videos with low and high spatial complexities.

general trend. This may be because unlike 5fps, 10fps, 15fps, and 30 fps videos, the 45fps videos could not be generated directly by uniformly picking one of every integer number of frames from the source video sequences of 60fps. Instead, three of every four frames were picked, which affects the uniformity of frame time-spacing. An alternative way of creating 45fps video from 60fps ones is to temporally interpolate and insert new frames to satisfy the uniform time-spacing condition. However, the interpolation process will create additional quality degradations of the video.

Across distortion levels, it can be seen that the quality improvement decreases with the level of quantization, where QP = 22 (less compression, higher quality) shows the largest improvement and QP = 37 shows the least improvement. This implies that there is a competing relationship in terms of perceived video quality between reducing compression artifact



**Fig. 4.** MOS versus frame rate for videos with low, medium and high object motion.

and increasing frame rate. Previous work [5, 6] addressed this aspects for 5fps to 30fps videos and proposed certain computational VQA models compromising both factors. However, this trend saturates again in the range of 30fps to 60fps, which indicates that previously developed models need to be reexamined for their generalization ability to high frame rate levels.

### 3.2. The effect of spatial content

Fig. 3 reports the MOS values for different complexity levels of spatial content with respect to quantization level and spatial resolution. A similar general trend of quality versus frame rate is observed. An interesting point to notice here is that for the case of 480p videos, although the MOS curves corresponding to low and high spatial complexity videos almost overlap with each other from 5fps to 30fps, there is a

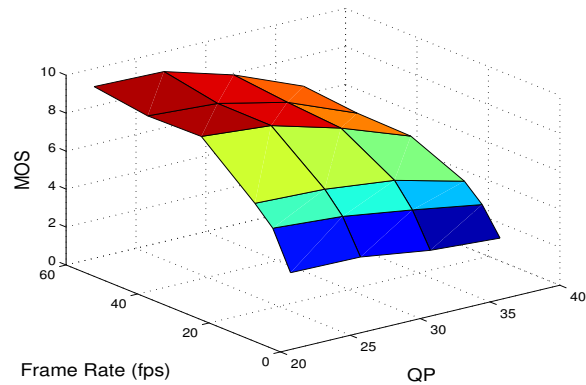
significant gap between them from 30fps to 60fps, where low complexity videos always obtain lower MOS values. One potential explanation is that high frequency, high texture complexity videos desire not only higher spatial sampling rate but also higher temporal sampling rate in order to accurately represent the complex content without strong (aliasing) artifacts, especially when there is motion associated with the complex textures. As a result, when the frame rate goes from low to high, humans recognize more quality improvement than that from relatively simple texture content. In the case of 1080p videos, the spatial resolution is already sufficient to precisely represent more complex content, and thus the benefit of moving towards high frame rate is less pronounced.

### 3.3. The effect of motion

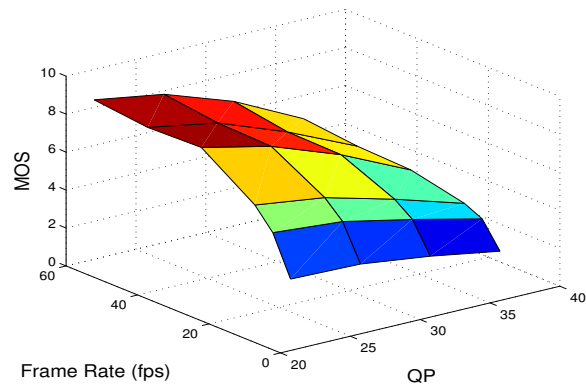
Fig. 4 reports the MOS values for different levels of object motion (low, medium and high) with respect to different quantization levels and different frame sizes. Based on previous studies (e.g. [14, 15]), it was expected that there exists some strong object motion dependency, i.e., with increasing frame rate, higher object motion videos would pronounce more improvements than lower ones. Surprisingly, this is not the case in our experiment, as no clear object motion dependency can be found in Fig. 4. Through more careful observations of the data and discussions with the subjects who did the experiment, we found two possible explanations. First, the uncertainty of human visual perception increases with the speed of motion [20, 21]. When the object motion is extremely high, the perceptual uncertainty becomes so high that further increasing frame rate would not help the visual system to capture more information from the scene. Second, in the case of low to moderate object motion, if they are accompanied by slow camera motion, humans tend to be more sensitive to temporal artifacts [22] and thus the effect of increasing the frame rate could be strong. It is also worth noting that the trend is independent of the quantization level.

### 3.4. Interactions between frame rate, quantization level and spatial resolution

The way the new database was built allows us to examine not only the impact of individual parameters including frame rate, quantization level, and spatial resolution on the overall video quality, but also their combined effect in a joint parameter space. Fig 5 (a) and Fig 5 (b) show the overall MOS score as a joint function of frame rate and quantization level, for 480p and 1080p resolution videos respectively. It can be seen that although increasing frame rate is generally helpful in improving the overall video quality, the speed of improvement depends on the quantization level. In other words, the overall quality improvement is not a simple additive effect of improving frame rate and reducing quantization. Their interactions need to be taken into account. A similar conclusion may be drawn when we include the spatial resolution parameter into



(a) 480p



(b) 1080p

**Fig. 5.** MOS as a function of frame rate and quantization parameter for 480p (a) and 1080p (b) videos.

the equation. Moreover, the results we presented earlier also show that spatial and motion complexities are adding more complications into the picture. Therefore, building a comprehensive objective quality prediction model that considers the impact of all parameters is a challenging but important task that desires deeper understanding and further investigation.

## 4. CONCLUSIONS

In this study we built a database that contains videos compressed at different frame rates, transform-domain quantization levels and spatial resolutions. We carry out a subjective study on the database and conduct a series of analysis to investigate the impact of frame rate on perceived video quality and its relationship with quantization level, spatial resolution, and spatial and motion complexities. Our investigation results in a number of interesting observations, which provide new insights for the future development of objective HFR VQA models, HFR video frame rate conversion algorithms, and HFR video compression and content delivery systems.

## 5. REFERENCES

- [1] J. Y. C. Chen and J. E. Thropp, "Review of low frame rate effects on human performance," *IEEE Trans. System, Man and Cybernetics, Part A: Systems and Humans*, vol. 37, no. 6, pp. 1063–1076, Nov. 2007.
- [2] G. Zhai, J. Cai, W. Lin, X. Yang, W. Zhang, and M. Etoh, "Cross-dimensional perceptual quality assessment for low bit-rate videos," *IEEE Trans. Multimedia*, vol. 10, no. 7, pp. 1316–1324, Nov. 2008.
- [3] Y. Ou, Z. Ma, T. Liu, and Y. Wang, "Perceptual quality assessment of video considering both frame rate and quantization artifacts," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 21, no. 3, pp. 286–298, 2011.
- [4] L. Janowski and P. Romaniak, "QoE as a function of frame rate and resolution changes," in *Future Multimedia Networking*, pp. 34–45. 2010.
- [5] Y. Ou, T. Liu, Z. Zhao, Z. Ma, and Y. Wang, "Modeling the impact of frame rate on perceptual quality of video," in *Proc. IEEE Int. Conf. Image Proc.*, San Diego, CA, Oct 2008, pp. 689–692.
- [6] Y. Ou, Z. Ma, and Y. Wang, "Modeling the impact of frame rate and quantization stepsizes and their temporal variations on perceptual video quality: A review of recent works," in *Proc. IEEE Int. Conf. on Information Sciences and Systems*, Princeton, NJ, March 2010, pp. 1–6.
- [7] Q. Huynh-Thu and M. Ghanbari, "Temporal aspect of perceived quality in mobile video broadcasting," *IEEE Trans. on Broadcasting*, vol. 54, no. 3, pp. 641–651, Sep. 2008.
- [8] Ming-Chen Chien, Ren-Jie Wang, Chien-Hsun Chiu, and Pao-Chi Chang, "Quality driven frame rate optimization for rate constrained video encoding," *Broadcasting, IEEE Transactions on*, vol. 58, no. 2, pp. 200–208, 2012.
- [9] Gayatri Yadavalli, Mark Masry, and Sheila S Hemami, "Frame rate preferences in low bit rate video," in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*. IEEE, 2003, vol. 1, pp. 1–441.
- [10] J. Joskowicz, R. Sotelo, and Ardao J. C. L., "Towards a general parametric model for perceptual video quality estimation," *IEEE Trans. on Broadcasting*, vol. 59, no. 4, pp. 569–579, Dec. 2013.
- [11] Cheon Seog Kim, Sung Ho Jin, Dong Jun Seo, and Yong Man Ro, "Measuring video quality on full scalability of h.264/AVC scalable video coding," *IEICE transactions on communications*, vol. 91, no. 5, pp. 1269–1278, 2008.
- [12] A. Khan, L. Sun, and E. Ifeachor, "QoE prediction model and its application in video quality adaptation over umts networks," *IEEE Trans. on Multimedia*, vol. 14, no. 2, pp. 431–442, Apr. 2012.
- [13] M. Claypool, K. Claypool, and F. Damaa, "The effects of frame rate and resolution on users playing first person shooter games," in *Proc. SPIE 6071, Multimedia Computing and Networking*, San Jose, CA, Jan. 2006.
- [14] A. Banitalebi-Dehkordi, M. T. Pourazad, and P. Nasiopoulos, "Effect of high frame rates on 3D video quality of experience," in *Proc. IEEE Int. Conf. on Cons. Electron.*, Las Vegas, NV, Jan. 2014, pp. 416–417.
- [15] A. Banitalebi-Dehkordi, M. T. Pourazad, and P. Nasiopoulos, "The effect of frame rate on 3D video quality and bitrate," *3D Research*, vol. 6, no. 1, pp. 1–13, Dec. 2014.
- [16] Jong-Seok Lee, Francesca De Simone, and Touradj Ebrahimi, "Subjective quality evaluation via paired comparison: application to scalable video coding," *Multimedia, IEEE Transactions on*, vol. 13, no. 5, pp. 882–893, 2011.
- [17] M Masry and Sheila S Hemami, "An analysis of subjective quality in low bit rate video," in *Proc. IEEE Int. Conf. on Image Processing*. IEEE, 2001, vol. 1, pp. 465–468.
- [18] Y. Wang, Z. Ma, and Y. Ou, "Modeling rate and perceptual quality of scalable video as functions of quantization and frame rate and its application in scalable video adaptation," in *Proc. Int. Packet Video Workshop*, Seattle, WA, May 2009, pp. 1–9.
- [19] John D McCarthy, M Angela Sasse, and Dimitrios Miras, "Sharp or smooth?: comparing the effects of quantization vs. frame rate for streamed video," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2004, pp. 535–542.
- [20] Z. Wang and Q. Li, "Video quality assessment using a statistical model of human visual speed perception," *Journal of the Optical Society of America*, vol. 24, no. 12, pp. B61–B69, Dec. 2007.
- [21] Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal Processing: Image Communication, special issue on Objective video quality metrics*, vol. 19, no. 2, pp. 121–132, Feb. 2004.
- [22] K. Zeng, T. Zhao, A. Rehman, and Z. Wang, "Characterizing perceptual artifacts in compressed video streams," in *Proc. SPIE 9014, Human Vision and Electronic Imaging XIX*, San Francisco, CA, Feb. 2014.