Frame Rate and Perceptual Quality for HD Video

Yutao Liu[1], Guangtao Zhai[2], Debin Zhao[1] and Xianming Liu[1]

[1] Department of Computer Science, Harbin Institute of Technology, Harbin, China

[2] Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai, China

{liuyutao2008, zhaiguangtao}@gmail.com, dbzhao@hit.edu.cn, xmliu.hit@gmail.com

Abstract. The frame rate (FR) of a video plays an important role in affecting the perceptual video quality. Most studies about the effect of FR on the video quality mainly focused on low frame rate, e.g. less than 30 frames per second (fps), at low resolutions like CIF or QCIF. As the video frame rate and resolution advance, we reconsider this issue and investigate the relationship between frame rate and the perceptual video quality under high frame rate and high resolution. In this paper, we discuss the impact of frame rate on the perceptual quality of High Definition (HD) video with high frame rates (up to 120fps) considered. Firstly, we design and conduct subjective experiment to construct the video dataset, which includes video sequences at different frame rates and the corresponding mean opinion scores (MOS) which represent the perceptual video quality. Based on the MOS results, we analyze how perceptual video quality changes as frame rate varies among different video sequences and propose some meaningful findings. The video dataset will be made publicly available. We deem that this study will enrich video quality assessment and benefit the development of high frame rate and high definition video business.

Keywords: video quality assessment, subjective experience, high frame rate, high definition

1 Introduction

It is well known that frame rate plays an important role in affecting the perceptual quality of a video. As a matter of fact, videos with different frame rates lead to different subjective experience for viewers. Another occasion is that special videos, like computer games or 3D videos, with improper frame rate may cause visual discomfort or visual fatigue that is harmful to people's health. Therefore, frame rate acts as an important video attribute which indeed affects the perceptual quality and study on the frame rate impact on the video quality occupies a necessary part of video quality assessment (VQA). Previous work on this point mainly assessed the video quality under several given frame rates, which rarely exceeded 30 frames per second (fps). However, as video frame rate advances, videos that own higher

adfa, p. 1, 2011.

frame rates than 30fps, like 60fps, 120fps (called high frame rate), have come into being nowadays. It can also be imagined that high frame rate will get another developing trend and become a new selling point in the movie industry. Therefore, it is needed to know how perceptual video quality changes as the frame rate arrives at higher values than 30fps that we have already been used to.

Previous works covering the impact of frame rate on the perceptual video quality have been conducted as follows: the work in [1] concerned three factors that affect the video quality which are spatial, temporal and amplitude resolutions and proposed functions to model the relationships between influential factors and the video quality. Zhai et al. in [2] investigated cross-dimensional perceptual quality assessment where frame rate was taken as an important parameter and presented some interesting observations. From another aspect, negative impact of frame dropping on video quality was investigated in [3]. In [4], the authors performed subjective experiments by varying the video's frame rate and quantization parameters and found the gap between PSNR and subjective experimental results. Based on this observation, they formed a more accurate subjective quality metric which combines frame rate, motion speed and PSNR to bridge the gap and got desired performance. Variable frame rate was also studied in [5] for the low bit rate video rate control scenarios.

Generally, the above works all take frame rate as an influential factor that will affect the perceptual video quality severely. However, the frame rates or resolutions of the videos employed in their study like 7.5fps or CIF (352×288) are relatively low, which will limit the applicability when it comes to video applications in high frame rate and high definition.

In this paper, we investigate how visual quality changes as frame rate varies, particularly choosing the high frame rate and high definition video sequences as our test materials. Firstly, subjective experiments are performed to construct the test video dataset containing video sequences at different frame rates and the corresponding mean opinion scores (MOS), which represent the videos' perceptual quality. Then we make sufficient statistical analysis on the obtained MOS values and present some meaningful findings. The video dataset of this work will be made publicly available for the research community. We expect that this study will enrich video quality assessment and benefit the development of high frame rate and high definition video business.

The rest of the paper is organized as follows. Section 2 details the design of the subjective experiment. The experimental results and analysis are presented in Section 3. At last, we conclude this paper and point out the future work in Section 4.

2 Subjective Experiment

2.1 Experiment materials

In this work, high frame rate and high definition videos were employed as our test materials for the subjective experiment. Specifically, the video sequences we adopted are from Ultra Video Group [6], which are "Beauty", "Bosphorus", "HoneyBee", "Jockey", "ReadySetGo", "ShakeNDry" and "YachtRide" respectively, with their snapshots shown in Fig. 1. The original frame rate of these video sequences is 120fps and the resolution is HD of 1080p (1920×1080). The spatial and temporal activities of these sequences cover a wide range, as we show in Fig. 2, where we utilized intra-frame variance and the mean of absolute inter-frame difference to represent the sequences' spatial and temporal activities respectively.

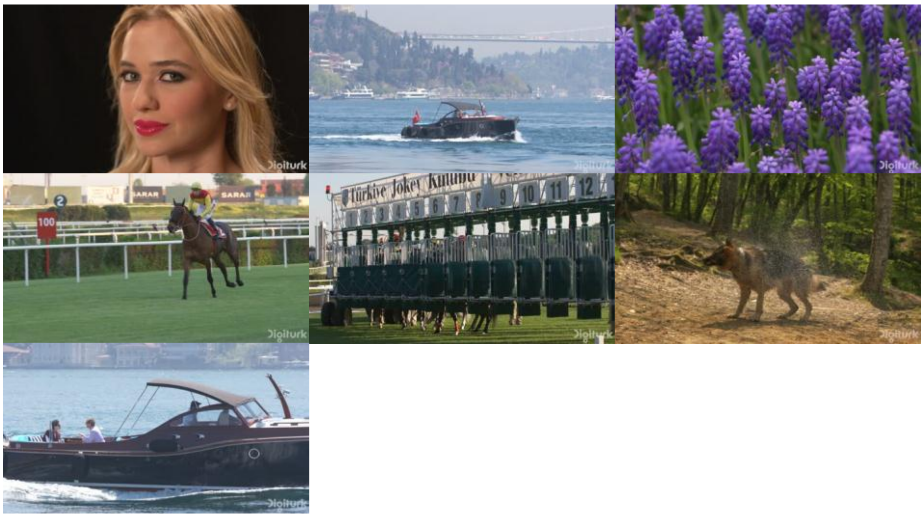
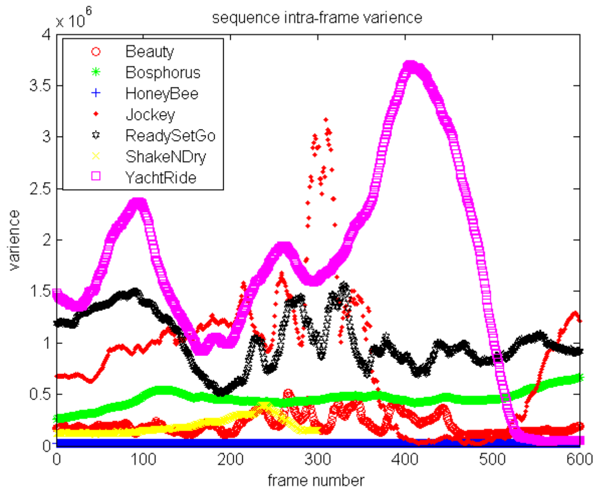
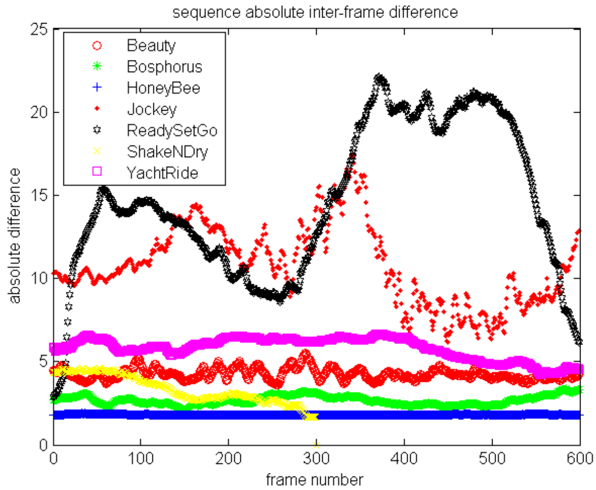


Fig. 1. Snapshots of the source video sequences (from left to right, top to bottom: "Beauty", "Bosphorus", "HoneyBee", "Jockey", "ReadySetGo", "ShakeNDry", "YachtRide").



(a) Intra-frame variance



(b) Mean of absolute inter-frame difference

Fig. 2. Intra-frame variance and absolute inter-frame difference for the sequences.

In our subjective experiment, we concerned the video frame rate at four different degrees, which are 15fps, 30fps, 60fps and 120fps respectively. The frame rate around 15fps seems to be the threshold of humans' satisfaction level [7]. 30fps is the frame rate that we have get used to. 60fps and 120fps are the emerging high frame rates. To emulate video sequences of lower frame rate than the original frame rate of 120fps, we down-sampled the video sequences on the time dimension and refilled the missing frames with their previous frames, the same as the method adopted in [8]. Fig. 3 gives an illustration to generate low frame rate videos, with the horizontal axis representing time dimension, N fps video sequence above and N/2 fps video sequence below.

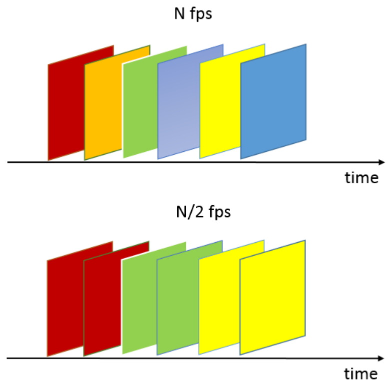


Fig. 3. The illustration of generating low frame rate video sequence.

2.2 Environment setup

We arranged the subjective experiment environment under the test conditions suggested in ITU-R BT.500-12 [9]. The illumination in the test room keeps low in order to avoid disturbance from other irrelevant light sources. The distance between the subject and the display is set about three times the height of the display. The display's refresh rate is 120 Hz and the resolution is 1920×1080. It should be noted that the size of the display is 22.9''. Both of the display refresh rate and resolution are set the same as that of the test video sequences, which guarantees normal playing of the test materials.

2.3 Experiment design

In our subjective experiment, we invited 20 inexperienced college students with normal vision or corrected to normal vision as our subjects for the test. In the process of the experiment, subjects watch the video sequences as usual and rate the overall video quality after each video sequence played. The scores rated by the subjects represents the perceptual quality for the sequences. The rating standard we adopted are 0-20, 20-40, 40-60, 60-80, 80-100, which stand for the video quality level of "bad", "poor", "fair", "good" and "excellent" respectively. The test video sequences are played randomly with frame rate unknown to the subjects. To guarantee the reliability of the rating procedure, each video sequence was played 3 times continuously. The subject is allowed to give a continuous score in [0,100]. Single stimulus evaluation method [9] without comparison to a reference version is employed during the whole test.

3 Experimental results and analysis

3.1 MOS results

After the subjective experiment, we got 20 raw scores rated in [0,100] from 20 subjects for each video sequence. The 20 raw scores represent the perceptual quality of the video sequence. However, different viewers have different psychological feelings to the same video sequence so that they may give different scores. For example, when watching the same video sequence of bad quality, one subject may rate 0, while the other subject may rate 15. Therefore, the obtained scores given by different subjects may drop into different intervals in [0,100]. To compare uniformly, we firstly normalized each viewer's scores into a unified interval [0, 1] by linear transform. Specifically, we found the maximum and minimum scores assigned by each subject. Then all the scores given by this subject can be normalized via:

$$n\_score = (score - min) / (max - min) \quad (1)$$

where n_score represents the normalized score, $score$ represents the raw score, min is the minimum score and max is the maximum score. By performing score normalization one subject by one subject, the raw scores were all normalized into [0, 1] and different subjects' experience can be compared uniformly. Then we averaged all the normalized scores for each sequence to get its final MOS. The results are shown in Table 1.

Table 1. MOS results

	Beauty	Bosphorus	HoneyBee	Jockey	ReadySetGo	ShakeNDry	YachtRide	AVG
@15fps	0.2142	0.0940	0.1259	0.0541	0.1789	0.4573	0.3283	0.2075
@30fps	0.7886	0.7398	0.5368	0.6017	0.6739	0.7387	0.6837	0.6805
@60fps	0.8177	0.8402	0.8057	0.8427	0.8878	0.8594	0.9731	0.8609
@120fps	0.8621	0.8901	0.8908	0.9093	0.9399	0.8822	0.9553	0.9042

3.2 ANOVA on MOS

To statistically verify that frame rate affects the perceptual video quality, we performed one-way analysis of variance (ANOVA) on the obtained MOS results. Specifically, we divided MOS values into four groups according to the four frame rates. For example, group 1 contains all the MOS values of 15fps. Then ANOVA was performed on the four MOS groups. The ANOVA results are listed in Table 2, as we can see the first column is the Sum of squares of the four groups. The second column is the degrees of freedom which is defined as the number of groups minus 1. The third column is the mean squares calculated by dividing the Sum of squares by the Degrees of freedom. The fourth column is F statistic and the last column gives the p-value. It can be observed that F statistic is much greater than 1 and the p-value is almost 0, which shows statistical significance, namely the MOS variance coming from different groups takes the dominating place and then verifies that frame rate affects the perceptual video quality severely.

Table 2. Results of ANOVA on MOS

Sum of squares	Degrees of freedom	Mean squares	F statistic	p-value
2.13597	3	0.71199	89.02	3.8288×10^{-13}

3.3 Model validation

The authors in [1] investigated the impact of spatial, temporal and amplitude resolution on the perceptual quality for compressed videos and proposed a video quality assessment model with considering these three factors. In this work, the spatial and amplitude resolution are fixed and we only focused on the temporal resolution (frame rate) effect, which in [1] was modelled by:

$$MNQT(t) = \frac{1 - e^{-\alpha_t (\frac{t}{t_{\max}})^{\beta_t}}}{1 - e^{-\alpha_t}} \quad (2)$$

where $MNQT(t)$ refers to the perceptual quality when the temporal resolution is t, t_{max} is 120fps here, β_t is set 0.63 according to [1], α_t controls the dropping rate as temporal resolution decreases, more details can be referred to [1].

We examined this model on our video dataset and calculated the correlation coefficient between the MOS values and the predicted MOS values obtained by equation (2). The correlation coefficient result is 0.81, while the average correlation coefficient is 0.95 in [1]. It should be noted that the max frame rate of the sequences tested in [1] is 30fps and the max frame rate of our test sequences is 120fps. Therefore, the results proves subjective experience changes differently as the video frame rate arrives at high values.

3.4 MOS results analysis

We show the MOS results visually in Fig. 4. It can be observed that the horizontal axis in each subfigure refers to frame rate and the vertical axis represents the value of MOS which reflects the overall perceptual quality of the video sequence. The red circles show the final results of our subjective experiment. The vertical bar means 95% confidence interval (CI). Additionally we list the average lengths of 95%

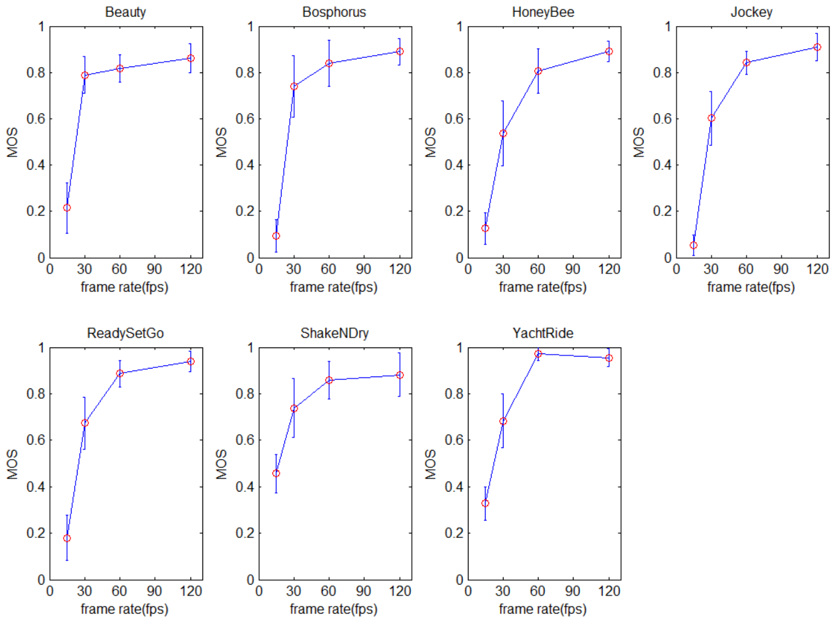


Fig. 4. MOS of the test sequences at different frame rates with 95% confidence interval.

Table 3. Average lengths of 95% CI at different frame rates

	@15fps	@30fps	@60fps	@120fps
95% CI length	0.1551	0.2353	0.1357	0.1132

CI at different frame rates in Table 3. As we can see in Fig. 4, the overall changing trend of the lines keeps much similar, namely, as the frame rate grows, the MOS values increase for all the test sequences, which means the perceptual quality of all the videos improves as frame rate increases. In addition, The MOS values are low at 15fps, most are less than 0.2, and the corresponding 95% confidence intervals are also small, that reflects the quality of videos at 15fps is poor and the viewers can easily sense the inferiority so that give relative low scores. However, the MOS values at 30fps increase quickly, most exceed 0.6, which means the frame rate of 30fps satisfies the subjects basically and the subjects tend to give much higher scores than that at 15fps. While the 95% confidence intervals at 30fps are generally larger than that at 15fps, which says there exists much more uncertainty to assess the video quality at 30fps compared to 15fps. It is interesting to find that the MOS values at 60fps are obviously higher than MOS values at 30fps. It should be noted that our subjective experiment is single stimulus without playing different frame rate videos for comparison, which implies that the perceptual video quality can improve at high frame rate that exceeds 30fps we've get used to. Similarly, the perceptual quality continuously improves when the frame rate reaches 120fps, while the growth rate slows down and the MOS improvements are less

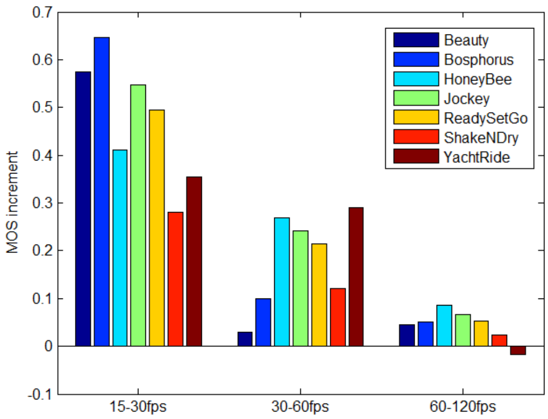


Fig. 5. The MOS increment results.

than that from 30fps to 60fps. All the confidence intervals at 120fps are also small and the subjects can consistently sense the superiority of 120fps.

Although the MOS changing trend as frame rate varies for all the videos exhibits much similarity, there still exists some differences for different videos. In particular, among different videos, the MOS increments from adjacent frame rates are different like MOS increments from 30fps to 60fps. We show the MOS increment results in Fig. 5, the first group at 15-30fps refers to the video frame rate at 30fps compared to 15fps. The second group is 60fps compared to 30fps and the third group is 120fps compared to 60fps. It can be clearly seen that most of the MOS increments in the first group exceeds 0.3, which means perceptual video quality at 30fps is much better than that at 15fps, as mentioned before. In the second group, the increments of “HoneyBee”, “Jockey”, “ReadySetGo” and “YachtRide” are above 0.2, while the increments of “Beauty”, “Bosphorus” and “ShakeNDry” are not significant. Combined with Fig. 2, the spatial and temporal energies of “Jockey”, “ReadySetGo” and “YachtRide” are higher than other sequences, which indicates high frame rate brings about obvious perceptual quality improvement of the videos with high spatial and temporal energies. Yet the quality of “HoneyBee” also improves a lot, in Fig. 2, “HoneyBee” yields the lowest spatial and temporal energies, which implies the spatial and temporal complexities of this video are the lowest. Therefore, we infer that subjects can also tell the perceptual difference of this kind of videos with simple scene and slight temporal changes as the video frame rate varies. While in the third group, the MOS of all the videos increases a little when the frame rate arrives at 120fps.

4 Conclusion

In this paper, we focused on the question of how frame rate affects the perceived video quality. Specially, we employed the emerging high frame rate and high definition videos as our video dataset. Subjective experiments were designed and conducted to assess the video quality under different frame rates. We performed thorough statistical analysis on MOS and presented some meaningful findings. In the future, we intend to extend our work from two aspects. Firstly, we will enlarge our high frame rate and high definition video dataset by including more videos like sports videos or computer game videos. Secondly, we will consider more factors that may have influence on the perceptual video quality as frame rate varies, like changing the video resolution or introducing some artifacts to the video.

5 Acknowledgement

This work is supported by the Major State Basic Research Development Program of China (973 Program 2015CB351804), the National Science Foundation of China under Grants 61300110.

6 References

1. Yen-Fu Ou, Yuanyi Xue and Yao Wang, "Q-STAR: A Per-ceptual Video Quality Model Considering Impact of Spatial, Temporal, and Amplitude Resolutions," Image Processing, IEEE Transactions on, vol.23, no.6, pp.2473-2486, June 2014.
2. Guangtao Zhai, Jianfei Cai, Weisi Lin, Xiaokang Yang, Wenjun Zhang and Etoh M., "Cross-Dimensional Perceptual Quality Assessment for Low Bit-Rate Videos," Multimedia, IEEE Transactions on , vol.10, no.7, pp.1316,1324, Nov. 2008.
3. Zhongkang Lu, Weis Lin, Boon Choong Seng, Kato Sadaatsu, EePing Ong and Susu Yao, "Perceptual Quality Evaluation on Periodic Frame-Dropping Video," Image Processing, 2007. ICIP 2007. IEEE International Conference on, vol.3, no., pp. III - 433, III - 436, Sept. 16 2007-Oct. 19 2007.
4. Feghali R., Speranza F., Demin Wang and Vincent A., "Video Quality Metric for Bit Rate Control via Joint Adjustment of Quantization and Frame Rate," Broadcasting, IEEE Transactions on , vol.53, no.1, pp.441,446, March 2007.
5. Hwangjun Song and Kuo C.-C.J., "Rate control for low-bitrate video via variable-encoding frame rates," Circuits and Systems for Video Technology, IEEE Transactions on , vol.11, no.4, pp.512,521, Apr 2001.
6. Ultra Video Group. http://ultravideo.cs.tut.fi/#main
7. Chen J.Y.C. and Thropp J.E., "Review of Low Frame Rate Effects on Human Performance," Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, vol.37, no.6, pp.1063-1076, Nov. 2007.
8. Zinner T., Hohlfeld O., Abboud O. and Hossfeld T.,"Impact of frame rate and resolution on objective QoE metrics," Quality of Multimedia Experience (QoMEX), 2010 Second International Workshop on , vol., no., pp.29,34, 21 -23 June 2010.
9. Rec. ITU-R BT.500-12, “Methodology for the subjective assessment of the quality of television pictures,” 2009.