PERCEPTUAL QUALITY OF VIDEO WITH FRAME RATE VARIATION: A SUBJECTIVE STUDY

Yen-Fu Ou, Yan Zhou, Yao Wang

Department of Electrical and Computer Engineering, Polytechnic Institute of New York University, Brooklyn, NY 11201
Emails: {you01, yzhou01}@students.poly.edu, yao@poly.edu

ABSTRACT

This work investigates the impact of periodic frame rate variation on perceptual video quality. Among many dimensions of frame rate variation, as a first step we focus on videos in which two frame rates alternate over fixed intervals. We present subjective test results, and analyze the influence of several factors (including the average frame rate, the frame rate deviation, and the video content) on the perceptual quality.

Index Terms— perceptual video quality, frame rate, jitter, quality metrics.

1. INTRODUCTION

In wireless video streaming, due to the limited sustainable bandwidth of a receiver, a video often has to be coded (or transcoded or extracted from a scalable stream) at a reduced frame rate and/or frame size, so that each coded frame has adequate quality. A critical issue is how to choose the appropriate spatial, temporal and amplitude resolutions (STAR), so as to achieve the best trade-off between picture quality and motion fluidity in the delivered video (Here we use amplitude resolution to refer to quantization parameter or QP). Another challenging problem is that the sustainable bandwidth of a wireless link often fluctuates in time, calling for adaptation of frame rate, frame size and QP. One naive approach would be to find the STAR that optimizes the perceptual quality over each short time duration based on the available instantaneous bandwidth. This will however create a video with rapidly fluctuating STAR, which may be annoying to the viewer. For example, variation in frame rate can cause visually annoying jitter artifacts. It is important to understand how does the variation of the STAR, individually and collectively, affect the perceived quality. Such understanding would enable us to impose proper constraints on the variation of the STAR, when adapting the STAR based on the time-varying bandwidth.

Take for example a hypothetical case where the available bandwidth alternates between $R_h$ and $R_l$, and the frame rates that can lead to the best perceived quality for constant rate video at $R_h$ and $R_l$ are $f_h$ and $f_l$, respectively. In this situation, it is better to code the video with alternating frame rates of $f_h$ and $f_l$, or would it be better to stay at $f_l$? More generally, one may want to vary not only the frame rate, but also the frame size and QP to meet the instantaneous rate constraints. In this paper, we focus on the impact of frame rate variation on the perceived quality, while keeping the frame size and QP fixed. Among the many dimensions of frame rate variation, we consider the the simple case where the frame rate alternates between $f_l$ and $f_h$, with each frame rate staying over a constant time duration $T$. We conduct subjective tests where viewers are asked to rate the quality of video with varying $f_l$ and $f_h$ for a chosen fixed $T$. We study the effect of $f_h$, $f_l$, their difference $Δf = f_h - f_l$ and ratio $f_l/f_h$ on the perceived quality. We include a variety of videos, to further assess the influence of the video content. This study directly addresses the questions we raised for the hypothetical example given earlier. But it also shed lights for more complicated cases where the frame rate may vary among more than two levels and the variation may not follow a periodic pattern.

There have been several studies regarding the influence of temporal and amplitude resolutions, individually or jointly, on the perceptual quality [1, 2, 3, 4]. Some of these works (e.g. [1, 2]) consider the case where the frame rate and QP are fixed in the entire video, whereas others (e.g. [3, 4]) consider the impact of frame rate variation, due to non-uniform and bursty packet losses. This type of frame rate variation however is quite different than the type we are concerned with.

This paper is organized as follows. Section 2 describes our subjective test configurations. Section 3 presents the subjective test results and our observations, and also proposes a model that relates to the perceived quality with the frame rate variation. Section 4 investigates the statistical significance of impact of frame rate variation and video content on perceptual quality using the ANOVA technique. Finally Section 5 concludes the paper.

2. SUBJECTIVE TEST SETUP

2.1. Testing Material

Our experiment is conducted using five video source sequences, Akiyo, Foreman, Football, Ice, Waterfall, all in CIF (352 × 288) resolution at original frame rate 30 fps, which are chosen from JVT (Joint Video Team) test sequence pool [5]. All these sequences are coded using JVT scalable video model (JSVM912) [6]. For each sequence, one bitstream is generated with four temporal layers corresponding to frame rates of 30, 15, 7.5, 3.75Hz, and each temporal layer in turn has three quality layers created with QP equal to 28, 36 and 40, respectively, using the coarse grain scalability (CGS).

We introduce temporal variation such that two frame rates will switch back and forth periodically every two seconds through the entire 8 seconds video. Let $f_h$ and $f_l$ denote the higher and lower frame rate of the video. Table 1 details all the test configurations,
which leads to a total of 150 processed (encoded and decoded) video sequences (PVS).

Table 1. Testing configuration

<table>
<thead>
<tr>
<th>QP</th>
<th>(f_h(\text{Hz}))</th>
<th>(f_l(\text{Hz}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>28/36/40</td>
<td>30/30/15/7.5/3.75</td>
<td>15/15/7.5/3.75</td>
</tr>
<tr>
<td>15</td>
<td>7.5/3.75</td>
<td></td>
</tr>
<tr>
<td>7.5</td>
<td>3.75</td>
<td></td>
</tr>
</tbody>
</table>

2.2. Experiment Configuration

The subjective quality assessment is carried out by using a protocol similar to ACR (Absolute Category Rating) described in [7]. Basically, each viewer is presented a series of video in a random order, and the viewer is asked to rate each video in the range of 0 to 100. Thirty three (27 males and 6 females) non-expert viewers who had normal or corrected-to-normal vision acuity participated in one or two subgroup tests. There are 16 ratings for each processed video sequence. Most of the viewers are engineering students from Polytechnic Institute of New York University, with age 21 to 33.

In order to remove “noisy” ratings or outliers, we adopted, with some modification, the screening method recommended by BT.500-11 [8]. After screening there are on average 15 user ratings for each processed video sequence, and the average standard deviation using a scale of 0 to 100 is 10 (ranging from 1-15).

3. SUBJECTIVE TEST RESULTS

Although the subjective test included processed video clips with different QPs, we will only present the results for clips obtained with QP=28 for the lack of space. The test video and subjective test results are available at http://vision.poly.edu/.

3.1. Impact of Constant Frame Rate

First we investigate the influence of the frame rate on the perceptual quality of a video with a constant frame rate, i.e., \(f_h=f_l\). Figure 1 (a) shows measured MOS vs. frame rate of the testing sequences. As expected, the MOS reduces as the frame rate decreases. Based on our previous work in [2], we model the impact of constant frame rate by,

\[
Q_e(f) = Q_{\text{max}} \frac{1 - e^{-b f/f_{\text{max}}}}{1 - e^{-b}}, \tag{1}
\]

where \(f\) represents the frame rate, \(f_{\text{max}}\) is the maximum frame rate \((f_{\text{max}} = 30 \text{ fps})\) and \(Q_{\text{max}}\) represents the quality rating for a video at frame rate \(f_{\text{max}}\). As can be seen from Fig. 1(b), this model fits with measured data quite well. Note that the parameter \(b\) characterizes how fast the quality drops as the frame rate reduces.

3.2. Impact of Frame Rate Variation

We now consider sequences in which the frame rate alternates between \(f_h\) and \(f_l\). We first discuss, under the same average frame rate, \(f_{\text{avg}}=(f_h+f_l)/2\), how does frame rate variation affects the perceived quality. Figure 2 shows that, when the \(f_{\text{avg}}\) is the same, the MOS for a video with a constant frame rate is higher than that with frame rate variation. Note that the brown solid lines in Fig. 2 are predicted quality of constant-frame-rate videos based on Eq. (1). The degradation due to frame rate change is more severe when \(\Delta f = f_h - f_l\) is higher (e.g., MOS difference between (30,7.5) and constant frame rate of \((30+7.5)/2=18.75\) is greater than MOS difference between (30,15) and constant frame rate of \((30+15)/2=22.5\)). This result is as expected, as large frame rate variation induces noticeable jitter. We note that for "Foreman", the fact that the MOS at \((f_h,f_l)=(15,7.5)\) is higher than at (15,15) may be due to inconsistency in viewer readings. In the following discussions, we ignore this particular data point. It is interesting to note that points corresponding to \(f_l = f_h/2\) are quite close to the operational quality-frame rate curves achievable by using constant frame rates, for most of the sequences. But those with \(f_l\) lower than \(f_h/2\) are much below the curve.

We also compare quality of the videos with different frame rate variations under the same average bit rate in Figure 3. It is clear that, a constant frame rate video has a better quality than a video with frame rate variation (especially when the frame rate variation is large), under the same average bit rate.

We next look at when \(f_l\) is fixed due to the lowest available bandwidth, whether alternating between \(f_h\) and \(f_l\) leads to better quality than staying at \(f_l\), when the available bandwidth fluctuates between the lowest bandwidth and a high bandwidth. This is the question we raised in the introduction as a motivation for this study. We plot normalized MOS (MOS\((f_h,f_l)/\text{MOS}(f_h,f_h)\)) against frame rate ratio \((f_h/f_l)\) in Figure 4, where it shows that alternating between \(f_h\) and \(f_l\) is generally better than staying at \(f_l\) when \(f_h/f_l \leq 2\). The slope of improvement reduces as \(f_l\) increases, and the degree of improvement depends on the motion characteristics of the video (e.g., Football and Ice have higher slope). When \(f_h\) is more than double of \(f_l\), the quality improvement is inconsistent. This suggests that when the lowest bandwidth limits the lowest frame rate to \(f_l\), even when available bandwidth at a later time allows a frame rate beyond \(2f_l\), it may be better to limit the highest frame rate to \(f_h = 2f_l\).

We next look at when \(f_h\) is fixed, how the quality changes with different \(f_l\). Figure 5 demonstrates the normalized MOS, MOS\((f_h,f_l)/\text{MOS}(f_h,f_h)\), against delta frame rate, defined as \(\Delta f = f_h - f_l\), to study the impact of the strength of frame rate variation. We observe that higher \(f_h\) has slower dropping trend than lower \(f_h\) along the \(\Delta f\) trajectory. This again implies that when \(f_h\) is already low, further inducing frame rate variation leads to more visual distortion, than when \(f_h\) is higher.

Instead of measuring the frame rate variation strength by the difference in frame rate, \(\Delta f\), Figure 6 uses the frame rate ratio, defined as \(f_l/f_h\). Fig. 6 shows how does the normalized MOS decreases with ...
Fig. 2. Measured MOS vs. average frame rate. The error bars are 95% confidence interval.

Fig. 3. Measured MOS vs. bit rate. The error bars are 95% confidence interval.

Fig. 4. MOS($f_l, f_h$)/MOS($f_l, f_l$) v.s. $f_h/f_l$ for different $f_l$.

Fig. 5. MOS($f_l, f_h$)/MOS($f_h, f_h$) v.s. $f_h - f_l$ for different $f_h$.

Fig. 6. MOS($f_l, f_h$)/MOS($f_h, f_h$) v.s. $f_l/f_h$ for different $f_h$. 
the frame rate ratio. It is interesting that the dropping trends under different $f_h$ are similar. This implies that the impact of frame rate variation can be quite well captured by the frame rate ratio, independent of $f_h$. We found that the variation of quality with the frame rate ratio can be modeled quite accurately using a power law function, i.e.,

$$\overline{Q_p}(f_h, f_l) = MOS(f_h, f_h) \left( \frac{f_l}{f_h} \right)^{\alpha} \quad (2)$$

where $\alpha$ is the model parameter. Fig. 6 shows both the measured quality (with points) as well as the predicted one (with curves) using this simple model.

### 4. STATISTICAL ANALYSIS

The results presented in Sec. 3 show that frame rate variation and video content both affect the perceptual quality. To evaluate whether the changes in quality ratings due to these factors are statistically significant, we perform the two-way Analysis of Variance (ANOVA) [9]. With ANOVA, we compute the probability (p-value, which is derived from the cumulative distribution function of F based on the $F$-value) of the event that the difference in MOS when a particular variable is changed is due to chance. If this probability is based on the $F$-value, which is derived from the cumulative distribution function of $F$ based on the $F$-value) of the event that the difference in MOS when a particular variable is changed is due to chance. If this probability is low ($p$-value < 0.05), we consider this variable as having statistically significant influence on MOS.

Instead of performing ANOVA on all possible frame rate variation cases, we focus on a few interesting pairs of frame rate variation, listed in Table 2. For each considered case, we evaluate the $p$-value due to frame rate variation ($\Delta f$), video content (i.e. across five video sources), and their interactions, respectively. Table 2 shows that, for cases $P_1$ - $P_5$, frame rate variation has statistically significant impact on the subjective ratings; whereas for case $P_6$, its impact is statistically insignificant.

The effect of video content is not significant for $P_1$, but it is for $P_2$-$P_5$. This suggests that the video content does not affect viewer ratings for videos at $P_1$, $P_2$-$P_5$, but it significantly influence viewer ratings at lower frame rate. Finally, the two-way ANOVA interaction between frame rate variation and content is not significant (all $p$-values are larger than 0.05) for $P_1$ - $P_5$. This indicates that the perceived degradation due to frame rate variation is largely independent of video content. This result is somewhat surprising and should be verified with subjective testing involving more test sequences and viewers.

### 5. CONCLUSION

In this paper, we report the results of our subjective experiments to investigate the impact of periodic frame rate variation on the perceived video quality. We observed several interesting trends. For example, under the same average frame rate or the same bit rate, a video with a constant frame rate is perceptually more appealing than a video with variable frame rates. The results presented in Sec. 3 show that frame rate variation has statistically significant impact on MOS when the frame rate ratio follows a power law function of $f_l$ for fast motion, and that the improvement is larger when $f_l$ is low; and the quality degradation due to frame rate variation from $f_h$ to $f_l$ follows a power law function of the frame rate ratio $f_l/f_h$. We also conducted ANOVA to evaluate the statistical significance of the impact of frame rate variation and video content on the perceived quality.

This paper focuses on the impact of periodic frame rate changes on the perceptual quality under the same quantization level and the same variation frequency. Future studies will examine the impact of other factors including variation frequency, variation of QP, and joint impact of frame rate and QP variation.

### 6. REFERENCES


