DTU Library

Objective assessment of the impact of frame rate on video quality

Ukhanova, Ann; Korhonen, Jari; Forchhammer, Søren

Link back to DTU Orbit

OBJECTIVE ASSESSMENT OF THE IMPACT OF FRAME RATE ON VIDEO QUALITY

Anna Ukhanova, Jari Korhonen, and *Søren Forchhammer*

Department of Photonics Engineering (DTU Fotonik)
Technical University of Denmark (DTU), Kgs. Lyngby, Denmark

ABSTRACT

In this paper, we present a novel objective quality metric that takes the impact of frame rate into account. The proposed metric uses PSNR, frame rate and a content dependent parameter that can easily be obtained from spatial and temporal activity indices. The results have been validated on data from a subjective quality study, where the test subjects have been choosing the preferred path from the lowest quality to the best quality, at each step making a choice in favor of higher frame rate or lower distortion. A comparison with other relevant objective metrics shows that the proposed metric on average provides a more precise correlation with the subjective results.

Index Terms— Video quality, frame rate impact, objective metric

1. INTRODUCTION

The perceived video quality in today's video applications is a significant part of the Quality of Experience (QoE) for the end users. Therefore, it is very important to measure the possible quality degradations in the system in order to maintain and control the quality of the video data. The methods for quality measurement are divided into two categories: subjective and objective measurements. Subjective quality is often expressed in Mean Opinion Score (MOS), but there are also other possible methods not based on quality scoring, e.g. pairwise comparisons [1]. Whereas subjective quality evaluation provides the most reliable results as quality is estimated by human beings, who represent the end users, it requires more resources and is not suitable for some applications, e.g. in real-time quality monitoring. As the most popular objective metrics, such as peak signal-to-noise ratio (PSNR), do not correlate perfectly with the subjective quality [2], more accurate objective quality metrics are needed for many applications.

As video data also includes the temporal dimension, the quality assessment for video is more demanding than for images. Unfortunately, most of the established video quality metrics do not take frame rate into consideration. However, the impact of frame rate on perceptual video quality has been studied by many researchers. McCarthy, Sasse and

Miras studied the effects of quantization vs. frame rate for video sequences with sports content [3]. Their work concluded that on a small screen, high spatial quality is preferred over the frame rate. However, their study did not propose any objective method for quality measurement.

A metric QM based on peak signal-to-noise ratio (PSNR) considering both quantization and frame rate is described in [4]. In case of high frame rates, the metric is dominated by the quantization errors and is close to the PSNR. Otherwise, the PSNR value (measured on the temporally upsampled sequences in case of low frame rate) is compensated depending on the frame rate reduction and the motion speed. However, the compensation is not enough for very low frame rate, as in this case PSNR values for interpolated sequences do not vary significantly for different video quality, which results in almost equal values of the quality metric at low frame rates.

Another quality metric VQMTQ considering both frame rate and quantization artifacts is proposed in [5]. That metric provides a good fit to the measured MOS scores. The model uses sequence-dependent parameters; however, it is possible to predict them based on the characteristics of the video sequences.

Peng and Steinbach proposed a novel full-reference video quality metric STVQM based on PSNR, frame rate, and spatiotemporal activity measures [6]. Their experiments show that they both perform well and the difference between VQMTQ and STVQM metrics is not statistically significant. However, STVQM has some advantages, namely the use of only two standard video activity indicators that can easily be computed, compared to the four parameters with significantly more complex interpretation used in VQMTQ. Moreover, STVQM has no codec-dependent parameters, unlike VQMTQ.

In this paper we propose a novel quality metric for objective quality assessment of the video data taking the impact of the frame rate into account. Our experimental results show that the proposed method allows determining the quality of the video sequences close to the subjective human opinion, and can compete with other recently developed objective quality metrics, such as VQMTQ and STVQM, by offering more constant performance with different contents.

2. METHODOLOGY

As reference data for this study, we have used the data obtained from the study published in [7]. The purpose of the original experiment was to find the quality optimal path through the plane, where one dimension is the quality measured in terms of peak signal-to-noise ratio (PSNR), and one other dimension is the frame rate. Different combinations of frame rates and PSNRs are represented by nodes, arranged in a form of a two dimensional grid. At each step, test subjects choose between two sequences according to their preference: one with higher frame rate but lower PSNR (the node upwards), and another with lower frame rate but higher PSNR (the node to the right). With this method, it is possible to find the average preferred path from the lowest to the highest frame rate and PSNR, reflecting the perceived relative importance of frame rate and PSNR along the path. The details of the study are omitted due to the lack of space, but interested readers may refer to [7].

Unfortunately, the described method does not produce any subjective quality scores directly. However, we can measure the PSNR difference between two different PSNR levels, and then estimate the perceptual PSNR (PPSNR) difference between two frame rate levels from the relation between test subjects choosing the higher frame rate and those choosing the higher PSNR. The concept is illustrated in Fig. 1. In the illustration, nodes 0, a and b represent video sequences with different combinations of frame rates (FR_1, FR_2) and PSNR levels ($PSNR_1$, $PSNR_2$). Transition probabilities p_a and p_b denote the proportion of test subjects who prefer to move from node 0 to node a and b, respectively ($p_a + p_b = 1$). PSNR difference between nodes 0 and b is known to be $\Delta PSNR$, and when p_a and p_b are known, we can assume that PPSNR difference $\Delta PPSNR$ between nodes 0 and a can be estimated as:

$$\Delta PPSNR = \Delta PSNR \cdot p_a / p_b \qquad (1)$$

The estimate is most reliable when $p_a \sim p_b \sim 0.5$ and becomes less accurate when either p_a or p_b approaches zero. The method can be applied regardless of the sign of $\Delta PSNR$.
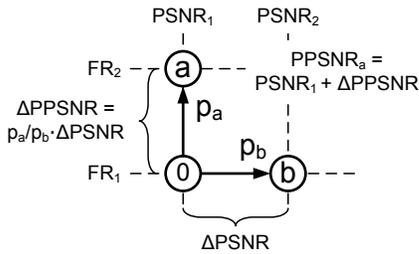

Figure 1. Method for deriving PPSNR values.

The data from [7] contains the preferred paths for five sequences judged by 25 test subjects, across planes with six frame rates and PSNR levels, ie. there are 6x6 nodes in each graph. Each node is assigned with a number of "occupants", ie. test subjects who have traversed through the node in question. For the practical application of the technique described above, we have first initialized the PPSNR values of the nodes with full frame rate with their measured respective PSNR values. Then, we have resolved the PPSNR values cumulatively for the lower frame rates by computing the proportion of "occupants" preferring higher frame rate or PSNR at each step. Naturally, some nodes have too few occupants to give reliable results; these nodes have been discarded from the final results. The most reliable PPSNR values are supposed to be found close to the most popular paths. Figure 2 shows the resulting PPSNR values for sequence "Coastguard".
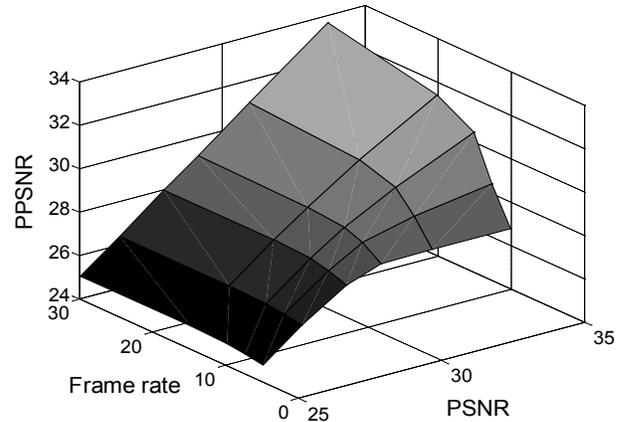

Figure 2. PPSNR values for "Coastguard".

The next challenge is to find an objective quality index matching with the experimental data obtained in the previous phase as accurately as possible. After attempting with several functions with potential resemblance to the surfaces such as in Fig. 2, we have identified that Eq. (2) gives the most promising results. The parameters of the quality function are *PSNR*, frame rate *FR*, and a content dependent parameter *c*.

$$PPSNR = PSNR \cdot \left(1 - \exp(1 - 10^8 \cdot FR \cdot PSNR^{-c})\right) \quad (2)$$

Since parameter *c* is related to the content type, we have used the spatial and temporal activity measures to predict *c*. In this work, we have used slightly modified definitions of the SI and TI indices in [8], denoted as *SA* and *TA*, as defined in [6]:

$$SA = mean_{time}(std_{space}[Sobel(F_n)]) \qquad (3)$$

$$TA = mean_{time}(std_{space}[F_n - F_{n-1}]) \qquad (4)$$

The parameter c is then predicted from SA and TA:

$$c = \exp(\alpha + \beta \cdot SA \cdot TA), \qquad (5)$$

where α and β are model parameters that can be solved by applying linear regression to the experimentally obtained values of $\ln(c)$. The logarithmic space is used instead of linear space, since we got slightly better results using the logarithmic space.

3. PERFORMANCE EVALUATION

In order to evaluate the performance of the proposed metric, we have compared our metric with VQMTQ and STVQM, as these are the most relevant comparison points known from the related studies. To make the comparison, we have created the preferred paths as in [7], derived from the quality values produced by the metrics named above. Based on the relative quality differences between adjacent nodes, we have computed the transition probabilities p_a and p_b for each node and then computed cumulatively the relative amount of occupants in each node, starting from the lowest frame rate and quality. The method for computing p_a and p_b is basically the reverse of the method shown in Fig. 1: we assume that p_a/p_b equals to $\Delta VQ_a / \Delta VQ_b$, where ΔVQ_a and ΔVQ_b are the differences in quality values between the source node and nodes a and b. Instead of $PPSNR$, VQ can be any quality metric with locally linear behavior. When the objective quality values are known, p_a and p_b can be computed as:

$$p_a = \frac{\Delta VQ_a}{\Delta VQ_a + \Delta VQ_b}, \; p_b = 1 - p_a \qquad (6)$$

As mentioned in Section 2, test subjects make their choice in several steps, starting with the lowest frame rate and PSNR. After the first step, the hypothetical test subjects are divided to nodes (1,2) and (2,1). After the second step, test subjects are distributed among nodes (1,3), (2,2) and (3,1), and so on. If we denote the relative number of occupants in node (x,y) as $s_{x,y}$, the average indices (x_i, y_i) after i steps can be computed as follows:

$$\begin{cases} x_i^{obj} = \sum_{x+y=i+2} s_{x,y} \cdot x \\ y_i^{obj} = \sum_{x+y=i+2} s_{x,y} \cdot y \end{cases} \qquad (7)$$

From [7], we have the preferred paths based on subjective evaluation for the following sequences in CIF (352x288) resolution: "Akiyo", "City", "Coastguard", "Football", and "Ice". The illustrations of the original subjective path for sequences "Coastguard" and "Football" are shown in Fig. 3, together with the paths resolved following the procedure

described above, applied to the objective quality indices from VQMTQ, STVQM and the proposed method.

Then, we have compared the subjective path against the objective paths by computing the mean Euclidean distance D between the average subjective position (x_i^{subj}, y_i^{subj}) and the position (x_i^{obj}, y_i^{obj}) computed from the objective data at each step i:

$$D = mean(\sqrt{(x_i^{obj} - x_i^{subj})^2 + (y_i^{obj} - y_i^{subj})^2}), \; (8)$$

where position is defined in terms of indices, since FR and PSNR scales are different.

Unfortunately, we only have the subjective paths available for the five abovementioned sequences, and such a small number of data does not allow us to create separate subsets for training and validation. To alleviate this a linear regression was performed for each sequence separately using the "leave-one-out" method, ie. parameters α and β required for predicting c were obtained using the known values of c for the other four sequences. The results for the distance comparisons between the subjective preferred path and the objective paths are listed in the Table I for each test sequence. Similar relative results can be obtained by using median instead of mean Euclidean distance.

Table I. Minimum distance comparison for preferred paths obtained by different quality metrics.

	akiyo	city	coastg.	footb.	ice	**avg.**
VQMTQ	0.65	0.34	0.27	1.24	0.37	**0.57**
STVQM	0.11	0.35	0.23	1.06	0.92	**0.53**
Proposed	0.11	0.69	0.54	0.20	0.49	**0.41**

The comparison with the QM metric [4] has not be included, since we have observed that increasing quality and bit rate with low frame rates does not always give increasing QM values, and this is why we have concluded that QM metric cannot predict the quality reliably at low frame rates.

As we can see, on average our proposed metric performs slightly better than VQMTQ and STVQM. However, the performance fluctuates between contents for all the metrics and none of them shows excellent performance on all sequences. For some contents, VQMTQ and STVQM perform better than the proposed metric, but the proposed metric achieves a more constant performance across different contents. The average performance of STVQM and VQMTQ is roughly similar, but VQMTQ has some disadvantages, such as the use of parameters dependent on the codec and content, which are not trivial to compute. We expect that the proposed method can be improved by using subjective data for more sequences, as the use of only four sequences in the "leave-one-out" method for parameter estimation can lead to overemphasis of outliers.
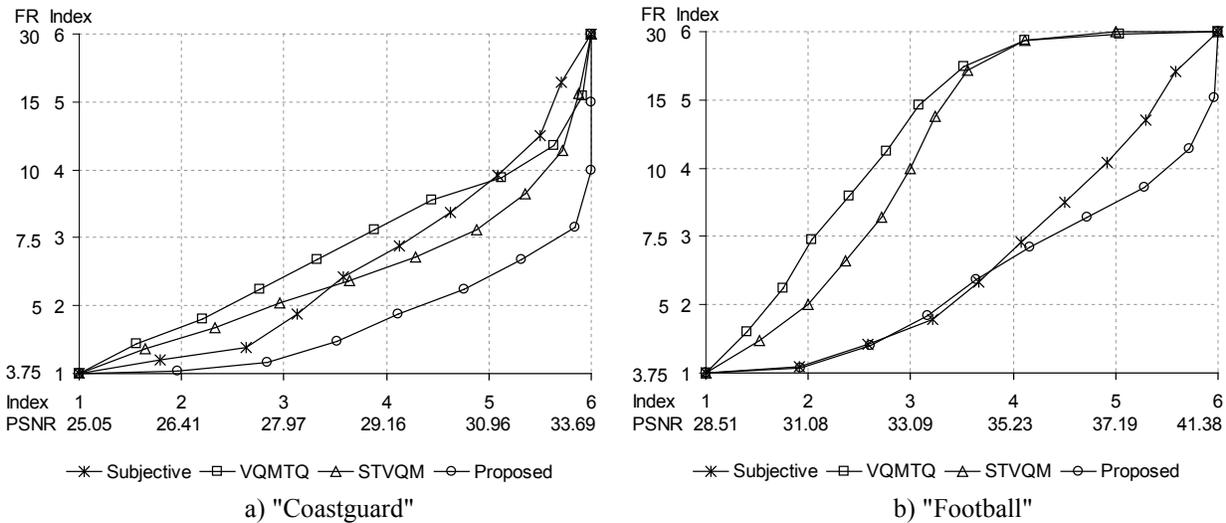
Figure 3. Preferred subjective and objective paths for sequences "Coastguard" and "Football".

4. CONCLUSIONS

Several studies about the impact of frame rate on subjective video quality conclude that the impact is highly dependent on the content. However, most of the established video quality metrics known from the literature do not take frame rate into consideration. In this paper, we have used data from a subjective study where the preferred path from the lowest quality to the best quality is determined by choosing between different frame rate and distortion levels at each hop. We have derived the relative subjective scores from the probability (frequency) of test subjects for choosing each of the two alternatives at each point.

We have proposed a simple metric for assessing the impact of frame rate on quality. In the proposed metric, the PSNR value of the video sequence is multiplied by a factor including frame rate and a content dependent parameter that can be computed from spatial and temporal activity indices SI and TI. We have compared the proposed metric against two other relevant metrics, VQMTQ and STVQM, and observed that in average, the proposed metric is capable of predicting the paths chosen by test subjects more accurately than the other metrics. However, the performance shows a significant dependency on the content. In the future, we plan to improve the proposed metric by involving more extensive subjective experiments, covering a larger variety of contents and higher resolutions.

5. ACKNOWLEDGEMENT

6. REFERENCES

[1] J.-S. Lee, F. D. Simone, T. Ebrahimi, "Subjective Quality Evaluation via Paired Comparison: Application to Scalable Video Coding," *IEEE Trans. Multimedia,* vol. 13, no.5, 2011.

[2] B. Girod, "What's Wrong with Mean-squared Error," in Digital Images and Human Vision, A. B. Watson, ed., pp. 207-220, MIT Press, 1993.

[3] J. McCarthy, M. A. Sasse, and D. Miras, "Sharp or Smooth: Comparing the Effects of Quantization vs. Frame Rate for Streamed Video," *Proc. of ACM CHI Hum. Factors Comput. Syst.*, pp. 535–542, Apr. 2004.

[4] R. Feghali, D. Wang, F. Speranza, and A. Vincent, "Video Quality Metric for Bit Rate Control via Joint Adjustment of Quantization and Frame Rate," *IEEE Trans. Broadcasting*, vol. 53, no. 1, pp. 441–446, Mar. 2007.

[5] Y.-F. Ou, Z. Ma, T. Liu and Y. Wang, "Perceptual Quality Assessment of Video Considering Both Frame Rate and Quantization Artifacts," *IEEE Trans. Circuits Syst. Video Techn.*, vol. 21, no. 3, 2011.

[6] Y. Peng and E. Steinbach, "A Novel Full-reference Video Quality Metric and Its Application to Wireless Video Transmission," *Proc. of ICIP'11*, Sep. 2011.

[7] J. Korhonen, U. Reiter, and J. You, "Subjective Comparison of Temporal and Quality Scalability," *Proc. of QoMEX'11*, Sep. 2011.

[8] ITU-T P.910, "Subjective Video Quality Assessment Methods for Multimedia Applications," International Telecommunication Union, Sep. 1999.