FRAME DROPPING EFFECTS ON USER QUALITY PERCEPTION

Ricardo R. Pastrana-Vidal°, Jean-Charles Gicquel*, Catherine Colomes*, Hocine Cherifi°*

*France Telecom R&D, DIH/EQS/MAI, 4 rue de Clos Courtel, 35512 – Cesson Sévigné, France
{ricardo.pastrana, jeancharles.gicquel, catherine.colomes}@francetelecom.fr

°LIRSIA, Université de Bourgogne, Faculté de Science Mirande, 21011 Dijon cedex, France
cherifi@crid-u.bourgogne.fr

ABSTRACT

Severe motion discontinuities are one of the most common degradations in video streaming. Frame dropping may be the source of this temporal discontinuity. When considering quality, it is essential to quantify user perception of the received sequence. We present a set of psychovisual experiments performed to characterize the effects of frame dropping. First, a previous study focusing on the impact of sporadically dropped pictures is briefly reviewed. Then, we present two experiments performed to quantify the effect of these temporal discontinuities for high temporal densities and distributions. We have found that quality impairment mainly depends on the duration and temporal density of discontinuities (burst of dropped frames). Finally, we propose a model that predicts the quality assessment from a group of assessors when a video is impaired by frame dropping. The model predictions show a high correlation with the observer's ratings.

1. INTRODUCTION

There is an increasing interest in real time video services over packet networks. When considering quality, it is essential to quantify user perception of the received sequence. Severe motion discontinuities are one of the most common degradations in video streaming. The end-user perceives a jerky motion when the discontinuities are uniformly distributed over time and an instantaneous fluidity break is perceived when the motion loss is isolated or irregularly distributed. Bit rate adaptation techniques, cell loss in the packet networks or restitution strategy could be the origin of this perceived jerkiness. At the source coding stage, temporal down-sampling is one of the most widely used techniques for bit rate adaptation; the sequence undergoes a frame dropping operation that affects the motion information. Packet loss or jitter could cause a sporadic or non-uniform frame discarding at the decoding process because of the buffering time limit [1]. The last picture that was received is then displayed until a new image has been reconstructed. The user then perceives a frozen image followed by an abrupt displacement of the objects. In the following text, we will use the term *temporal discontinuity* as a perceptual synonym of a dropped picture burst.

Nowadays, psychovisual experiments are the only recognized way to characterize the perceived quality. Recently, a quality perception study conducted by the authors showed that subjects had a significant negative reaction to this temporal discontinuity [2]. Our previous study was mainly focused on the impact of sporadically dropped pictures; i.e., for a reduced number of discontinuities. However, several questions were left open in relation to the influence of these temporal discontinuities for high temporal densities and distributions. We have therefore been studying the effect of frame dropping at a constant rate either alone or combined with a burst of discarded pictures for different durations on perceived quality; the influence of different temporal distribution profiles was also studied. The corresponding psychovisual experiments and results are reported here. Furthermore, we propose a model to calculate the effect on perceived quality of several frame dropping conditions.

2. PREVIOUS STUDY

We have organized the first study in two main parts: the visibility of frame dropping and its effect on the user quality.

2.1. Visibility findings

We have found that the discontinuity caused by one discarded picture (80ms) is visible. We have also reported that the unequivocal detection (100%) of a burst of dropped frames is attained when the number of discarded pictures is equivalent to 200ms. Furthermore, this finding is valid for all selected contents.

2.2. Quality at low density frame dropping

First, the impact of a single burst of discarded pictures was evaluated as a function of burst duration. Second, we have examined how quality is affected by a reduced number of discontinuities.

Fig. 1 shows the mean opinion scores (MOS) as a function of burst duration. We have found that a short dropping of 80ms has a reduced impact on quality. The perceived quality is still considered as excellent. However, subjects showed significant negative reactions for discontinuities of 160 and 280ms. Quality function exhibits a fall in ratings of 1 MOS category. After this range, quality function shows a slower decreasing rate.

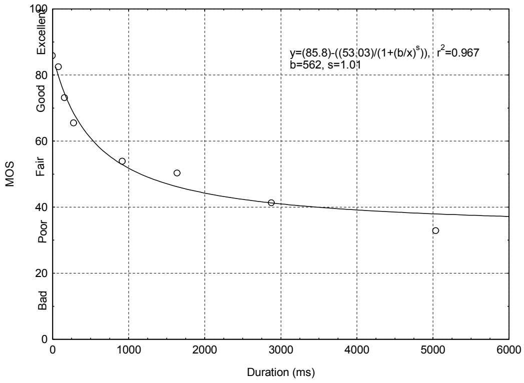

Fig. 1. Mean Opinion Scores (MOS) as a function of discontinuity duration. The temporal discontinuity is generated by a single burst of consecutively dropped pictures.

With regard to effect of a reduced number of dropped picture burst on the means opinion scores, the assessment results exhibited that quality impairment mainly depends on the duration and temporal density of discontinuities in a non linear manner. Subjects seem to be more annoyed by several short discontinuities than by a single burst of long duration. On the other hand, the relation between quality impairment and motion is content dependent.

3. EXPERIMENT 1

The main goal of this experiment was to characterize the effect of frame dropping at a constant rate either alone or combined with a burst of discarded pictures for different durations on perceived quality.

3.1. Apparatus

The test was carried out in a luminance-controlled environment. The background and screen luminance conditions were based on the ITU-R BT.500 [3] and ITU-R BT.814 [4] recommendations. All videos were stored and displayed on the PC station. The computer CRT was a NEC Multisync FE750+, 17 inches. The distance between subject and computer screen was approximately 60cm. This is a distance frequently found in home and professional computer viewing situations. The SEOVQ [5] (Subjective Evaluation and Optimization of Video Quality) software tool for quality evaluation of multimedia images was used to perform the test and gather participant's answers by means of an evaluation interface.

3.2. Stimuli

Four original video contents were selected: sport, entertainment, news, and a classic test sequence (Mobile & Calendar). The sequences were all 10 sec long to avoid the forgiveness effect [6]. The image format was CIF, RGB 24 bits and 25fps. The impaired sequences present discarded pictures placed at different motion contexts. The duration of the impairments were selected with enough perceptual distance. Spatial degradation was not introduced in order to avoid tradeoff between acuity and fluidity in the quality assessment.

The discontinuity caused by the frame dropping process was inserted in the reference video sequences by discarding n pictures from the temporal index p and repeating n times the picture corresponding to the temporal index $p-1$.

This experiment is focused on the effect of a single burst of discarded pictures when the nominal frame rate is reduced. Three frames rates were employed: 12.5, 8.33 and 5.0fps. respectively corresponding to a temporal decimation by a factor of 2, 3 and 5 because the original frame rate was 25fps. We have also combined every frame rate with a discontinuity (burst of dropped pictures) of different duration. In order to cover a broad range of possible situations we have selected values from 160 to 2880 ms.

3.3. Observers and method

The quality ratings from 20 subjects were gathered using the SAMVIQ method (Subjective Assessment Methodology for Video Quality) developed by EQS laboratory, France Telecom R&D. This method was submitted to the ITU-R study group 6Q (Quality assessment) and it was adopted by The European Broadcasting Union (EBU) [7]. The SAMVIQ was derived from the DSCQS [3] (Double Stimulus using a Continuous Quality Scale) method but it has the advantage of being more appropriate for the assessment of a wide range of qualities. A numerical scale (0-100) for rating overall quality is used. This scale is related to five quality categories (bad, poor, fair, good, excellent) that are uniformly distributed.

3.4. Results of the experiment 1

The data of the experiments consists of a mean opinion score (MOS) over all contents. Fig. 2. plots the results of experiment 1.

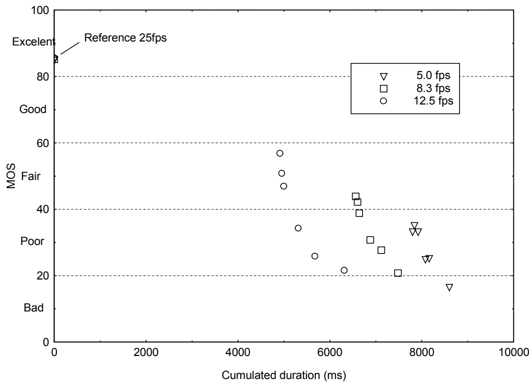

Fig. 2. The effects of regular frame dropping (frame rate reduction) either alone or combined with a burst of discarded pictures. The MOS are plotted as a function of the cumulated duration of all dropped pictures.

The upper point of each set corresponds to the frame rates alone. The other points are the result of the combination with a burst of discarded pictures. We can see that the assessors were less annoyed by a regular frame dropping (frame rate reduction alone) than by an irregular discarding.

The plotted data in Fig. 2 shows that all discontinuities caused by frame dropping (regular or irregular) contribute to the overall quality degradation but their degree of contribution depends on their duration and density. The quality of the reference sequences is related to a cumulated duration of zero because we have assumed that for a frame rate of 25fps, i. e. image period of 40ms, the temporal discontinuities are not existent (perceptible).

4. EXPERIMENT 2

The test apparatus, observers and the subjective method were the same as the elements described in experiment 1. The goal of this experiment was double. The first goal was to characterize the influence of low, medium and high burst densities on quality. This could also be seen as frame rate reduction during a short, medium or long period.

4.1. Stimuli

The burst strengths were: 80ms (2 images), 160ms (4 images) and 280ms (7 images) corresponding to a frame rate of 12.5, 6.25 and 3.57 fps. The first profile have 3 burst of dropped frames. The second profile corresponds to a reduction of frame rates during 4000ms and the last profile to a frame rate decimation on the overall sequence.

4.2. Results of experiment 2

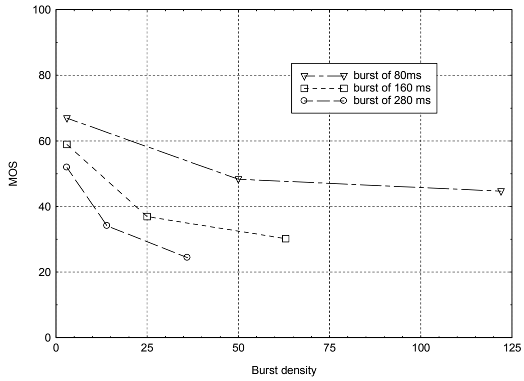

Fig. 3. MOS as a function of burst density.

Quality as a function of burst density exhibits a more dramatic fall in ratings when the burst duration is more significant (Fig. 3). It is interesting to note that for 80ms of discontinuity duration the quality reduction between regular and high density is not significant.

5. PROPOSED MODEL

We have found that quality impairment mainly depends on the duration and temporal density of discontinuities (burst of dropped frames) in a non linear manner. These findings led us to propose a model to calculate the effect on quality of several frame dropping conditions: regular and non regular discarding process; sporadically dropped pictures of different burst durations, distribution profiles and densities.

Our model combines the quality function for a single burst of dropped frames, the density of discontinuities and a power function depending on the burst density:

$$\hat{Q} = Q_{ref} - d_{total} ,$$

$$d_{total} = \left[\sum_{t=\min}^{\max} d_t \right]^{1/2} ,$$

$$d_t = n(t) \times \left[\hat{e}(t) \right]^{p(n(t))} ,$$

$$\hat{e}(t) = Q_{ref} - \hat{q}(t) ,$$

$$\hat{q}(t) = \mathrm{m}_{\max} - \left((\mathrm{m}_{\max} - \mathrm{m}_{\min}) / \left(1 + (\mathrm{b}/t)^{\mathrm{s}}\right) \right),$$

$$p(n(t)) = \mathrm{p}_{\max} - \left((\mathrm{p}_{\max} - \mathrm{p}_{\min}) / \left(1 + (\mathrm{c}/n(t))^{\mathrm{r}}\right) \right),$$

where Q_{ref} is the MOS computed over all reference sequence (unimpaired), d_{total} is the overall degradation in the 10sec sequence, t is the burst duration, d_t is the calculated contribution from all bursts having a duration of t, the term $n(t)$ corresponds to the distribution of burst duration. The expression $\hat{q}(t)$ is the quality function for an isolated burst of dropped frames having a duration of t, the plot of this function can be found in Fig. 1. The constants $\mathrm{m}_{\max}$ and $\mathrm{m}_{\min}$ are the extreme quality values found in the experiment results. The $p(n(t))$ is the power or exponent function that depends on the distribution of burst duration. This function was obtained by fitting the data of optimized exponents for several burst densities of a constant duration.

The power function accounts for the fact that subjects are more sensitive to dropping variations than to a uniform discarding. By means of this variable exponent, the contribution of each burst is less significant when the number of bursts of the same duration is high. The exponent value could vary from 1 to 2.

This model expresses our main hypothesis: the overall degradation caused by a certain frame dropping distribution may be calculated as an integration (summation) of the individual effects.

The term $q(t)$ takes into account the non linear negative reaction to a single discontinuity (burst of dropped frames) found in our previous subjective experiment (Fig. 1).

The model predictions show a high correlation with the observer's ratings (Pearson r = 0.9436). Fig. 4 shows the predicted scores versus quality assessment results from the experiments over 74 test conditions. For the tuning model, 21 frame dropping distributions were used.

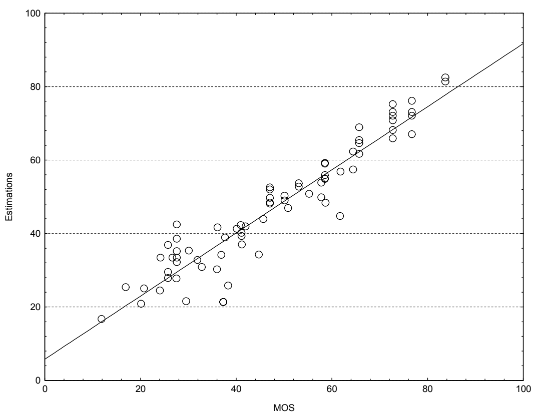

Fig. 4. Predicted scores and experimental MOS

6. CONCLUSION

Subjects were less annoyed by a regular frame dropping (frame rate reduction alone) than by an irregular discarding or sporadic discarding.

We introduced a model that predicts the quality assessment from a group of users when video is impaired by frame dropping. This model is able to calculate the effect on quality under several frame dropping conditions: regular and non regular discarding processes; sporadically dropped pictures for different burst durations, distribution profiles and densities. The model predictions show a high correlation with the observer's ratings (r = 0.94).

In addition, this study has outlined a promising methodology for a perceptual characterization and modeling of temporal discontinuities that is useful for video quality metrics design, perceptual frame discarding algorithms and for a better understanding of time-varying quality.

7. REFERENCES

[1] M. Kalman, E. Steinbach, B. Girod, "R-D optimized media streaming enhanced with adaptive media playout," *ICME*, Vol. 1, IEEE, Lausanne, pp. 869 - 872, August 2002.

[2] R.R. Pastrana-Vidal, J.C. Gicquel, C. Colomes, and H. Cherifi, "Sporadic Frame Dropping Impact on Quality Perception," *Human Vision and Electronic Imaging IX*, SPIE, vol. 5292, San José CA, January 2004.

[3] ITU-R BT.500, *Methodology for the Subjective Assessment of the Quality of Television Pictures*, 1998.

[4] ITU-R BT.814, *Specifications and Alignment Procedures for Setting of Brightness and Contrast of Displays*, 1994.

[5] J. Blin, *SEOVQ Software Tool for Quality, Preference and Acceptability Evaluation of Multimedia Images*, France Telecom R&D report FT.BD.FTR&D/DIH/EQS/462/02/JLB, 2002.

[6] R. Aldrige, J. Davidoff, M. Ghanbari, D. Hands, and D. Pearson, "Recency Effect in the Subjective Assessment of Digital Coded Television Pictures," *Image processing and its applications*, issue 410, IEE, Edinburgh, pp. 336-339, July 1995.

[7] *SAMVIQ – Subjective Assessment Methodology for Video Quality*, EBU Project Group B/VIM Video In Multimedia, BPN 056 report, May 2003.