DISPLAY AWARENESS IN SUBJECTIVE AND OBJECTIVE VIDEO QUALITY EVALUATION

Sylvain Tourancheau[1], *Patrick Le Callet*[1], *Kjell Brunnström*[2] *and Dominique Barba*[1]

(1) Université de Nantes, IRCCyN laboratory
rue Christian Pauc, 44306 Nantes, France
phone: +33 (0)240 683 047
patrick.lecallet@univ-nantes.fr

(2) Acreo AB, Video and Display Quality
Electrum 236, SE-164 40 Kista, Sweden
phone: +46 8 632 77 00
kjell.brunnstrom@acreo.se

ABSTRACT

Most of studies on video quality assessment are focused on the impact of coding distortion or transmission error. In this paper, display is considered. Regarding technology, some subjective experiments suggest that there are differences in term of quality between LCD and CRT. CRT provides a better quality when viewing HD video content while it is LCD when viewing still colour images. One explanation of this behaviour is explained through the LCD motion blur. From a motion blur perception model, an efficient metric of the quality loss due to this effect on LCD is proposed. Finally, the results of subjective experiments using SD video content are consistent with a motion-blur-based model and point out that the effect of display technology is linked with the video resolution.

1. INTRODUCTION

The incoming of the high-definition new visual experience at home have boosted the new display technologies, since they enable the increase of the screen size necessary to sense immersion, impact and immediacy as in a movie theatre [1]. For these reasons, these new displays, and particularly the liquid crystal displays (LCD), will soon replace the old mature CRT technology.

The liquid crystal displays have many differences with the CRT displays. Some subjective preference tests between these two types of displays have highlighted a high preference for the CRT displays concerning moving pictures [2]. Many defects have been counted by the viewers, such as colour differences, degradations in dark areas and deinterlacing artifacts for the interlaced sequences. But among all these defects, the motion blur seems to be the most annoying one, particularly in the sequences with significant movements. On the other hand, the CRT displays suffer from several shortcomings too. The flickering can be annoying in certain conditions, and the small luminance range can led to flat pictures with dirty colours.

In the recent years, the subjective and objective quality assessment becomes a research topic of interest. The activities of the Video Quality Experts Group (VQEG) are a good example of this interest. Previous works [3] and work in progress [4] are mainly related to coding or transmitting purpose at a given resolution, e.g. coding artifacts and transmission errors. Considering the whole chain, the quality assessment should be able to manage the dependency to other technology issues.

In this paper, display is studied. As a consequence, only the high part of the quality range is considered, using high definition (HD) video sequences and still pictures with no (uncompressed) or very few coding distortions. In order to know the impact of the display distortions, the subjective quality assessment is performed both on LCD and CRT. A loss of quality is observed on LCD for the moving pictures but not for the still ones. Assuming that the LCD motion blur is the most annoying artifact when displaying moving pictures on LCD, its perception is described and its magnitude measured. This leads to the design of an objective metric which enables the prediction of the loss of perceived quality on LCD with respect to the CRT one. Such a metric could be used in order to evaluate the LCD improvements introduced by the manufacturers to reduce technology artifacts.

In the last section, the impact of the LCD technology on the perceived quality regarding the video resolution is explored with some subjective experiments on standard definition (SD) sequences. As expected, the results are consistent with the objective metric based on the LCD motion blur, i.e. the influence of display issues in video quality increase with the video resolution.

2. SUBJECTIVE QUALITY ASSESSMENT TESTS

2.1 Tests conditions and equipment

The subjective quality assessment tests have been performed in a specific showroom, with lighting conditions and display parameters precisely measured and adjusted according to the ITU recommendations BT.500-11 [5] and BT.710-4 ITU [6]. The same tests have been conducted twice during two different sessions, the first time using a HDTV CRT display JVC DT-V 1910CG and the second time on a HDTV LCD Philips T370 HW01. The surrounding conditions and display parameters for each session are presented in Table 1. The viewing distances have been set to $3H$, where H is the height of the displayed pictures.

2.2 Protocol

These tests have been performed with sequences and pictures of fair-to-excellent quality. As a consequence, the used protocol should enable the quality discrimination. A well known stable method for this purpose is the SAMVIQ protocol [7], developed by France Telecom R&D and standardised by the EBU and the ITU.

SAMVIQ is a multi stimuli continuous quality scale protocol. With this procedure, the observers can compare some processed sequences (resp. pictures) both between them and with an explicit reference sequence (resp. picture). This leads to a precise and reliable measurement of the quality

CRT display (JVC DT-V 1910CG)

Background luminance of the testroom	7 cd/m^2
Background chromaticity	D_{65}
Screen diagonal	16.5 in
Picture height (H)	20.5 cm
Viewing distance	61.5 cm (3H)
Display black luminance	0.53 cd/m^2
Display peak luminance	70.9 cd/m^2

LCD display (Philips T370 HW01)

Background luminance of the testroom	35 cd/m^2
Chromaticity of background	D_{65}
Screen diagonal	37 in
Pictures height (H)	46 cm
Viewing distance	138 cm (3H)
Display black luminance	0.64 cd/m^2
Display peak luminance	471 cd/m^2

Table 1: Viewing conditions and displays parameters for the two sessions.

[8]. The notation scale is continuous, each score can take a value between 0 and 100.

2.3 Observers

The observers were mainly (about 80%) students between 20 and 25 and the gender parity was almost respected (about 2/3 of male). All were familiar with standard television and cinema but not with HDTV. The acuity and the colour perception of each observer have been checked, respectively with Monoyer's plates and Ishihara's test for colour blindness. The observers with at least on error in Ishihara's test or with an acuity less than 9/10 was rejected.

After the tests have been completed by all the observers, a rejection technique from the EBU [7] is applied. This process verifies the consistency of the scores of one observer according to the mean score of all the observers. Following the application of this rejection process, 15 valid subjects should be retained at minimum.

3. VIDEO QUALITY ASSESSMENT

3.1 Material

In order to measure the difference of quality between the two types of displays for moving pictures, nine 1080i[1] sequences with significant movements have been chosen. These videos have been supplied by the European broadcasters SVT and Euro1080.

Each of them contains 250 frames which corresponds to a 10-second duration. Each reference (uncompressed) has been coded with the H.264 reference coder at seven different bit-rates in order to cover a range of quality from fair to excellent (according to authors' judgement). These distorted sequences and the reference one are then submitted to the observers' judgement through the SAMVIQ protocol. In addition to the explicit tagged reference sequence, a hidden reference sequence is placed among the distorted sequences.

The sequences are received in 1080i format by the two displays. They're displayed in interlaced format on the CRT but not on the LCD which de-interlaces them since the flat panel matrices work in a progressive mode.

[1] 1080i format: 1920×1080 resolution in interlaced mode

3.2 Results

The mean opinion scores (MOS) of the observers for the nine reference sequences and on the two types of displays are presented in Table 2. ΔMOS is the difference of the MOS from CRT and LCD:

$$\Delta MOS = MOS\ CRT - MOS\ LCD \quad (1)$$

Sequences	MOS CRT	MOS LCD	ΔMOS
VOILE	83.9	77.7	6.2
FOOT	82.8	76.3	6.5
CONCERT	84.5	73.8	10.7
SHOW	82.9	75.3	7.6
CREDITS	83.1	79.1	4.0
MOBCAL	81.4	81.0	0.4
PARKRUN	87.6	80.2	7.4
SHIELDS	86.7	78.2	8.5
STOCKHOLM	86.1	82.3	3.8

Table 2: MOS for the nine sequences on the two displays.

The perceived quality of the moving pictures displayed on LCD is globally lower than the perceived quality of the moving pictures displayed on CRT. It's interesting to notice that this loss of quality is quite important for the sequences with quick movements such as *Concert*, *Parkrun*, *Foot* and *Voile*.

This loss of quality on LCD seems to be related to the quantity and/or the fastness of the movements in the sequence. To validate this hypothesis, it has been decided to conduct the same experiment with still pictures.

4. STILL PICTURES QUALITY ASSESSMENT

4.1 Material

Five images have been chosen in order to measure the difference of quality between the two types of displays for still pictures. They contain specific contents such as natural textures, flesh colours, oriented contours, water reflection, written characters, etc. Each of them has been distorted with two types of process: JPEG compression to have an anchor with a fair quality, and down-scaling/up-scaling filtering to simulate the resolution adaptation from SD to HD. Here again, these distorted pictures and the reference one are submitted to the observers' judgement. A hidden reference picture is added to the set of sequences to assess.

The pictures are displayed in interlaced format on CRT and in progressive format on LCD in order to repeat exactly the same conditions as those of the video quality assessment. The same group of observers has been used for the two displays. The group has been split in two equal parts: the observers of the first part have passed the test on CRT first, the observers of the second part have passed the test on LCD first.

4.2 Results

The MOS of the observers for the five hidden reference pictures on the two displays are presented in Table 3.

Pictures	MOS CRT	MOS LCD	ΔMOS
FOOTBALL	66.3	79.4	-13.1
HAND	73.6	80.3	-6.7
HOUSE	51.8	81.8	-30.0
LANDSCAPE	73.5	78.7	-5.2
MAP	51.4	84.4	-33.0

Table 3: MOS for the five pictures on the two displays.

It can be observed that for still pictures the quality on LCD is globally preferred. For the pictures *House* and *Map* the difference between the two types of displays is largely in favour of LCD (with a ΔMOS of about a third of the quality scale). This can be explained by the presence of fine horizontally-oriented contours which, associated to the interlacing, make the flickering of the CRT more noticeable.

Overall, the shortcomings of CRT displays such as flickering and the limited range of luminance seems to lead to a lower feeling of natural and a lower sense of immersion. The LCD is brighter, vivid and colourful and the perceived quality of still pictures is clearly higher on it ($\Delta MOS_{mean} = -17.6$). However, with exactly the same viewing conditions and displays parameters, the perceived quality of moving pictures is higher on CRT ($\Delta MOS_{mean} = 6.4$). It's assumed that this difference must be due to the moving artifacts such as LCD motion blur and de-interlacing distortions which are not present on still pictures.

5. LCD MOTION BLUR

The results described in the previous part lead to the statement that the excellent perceived quality on LCD with still pictures is strongly reduced with moving pictures. Moving artifacts due to the LCD technology, and particularly the LCD motion blur, seem to be responsible for this loss of quality in video.

In this part, the LCD motion blur is described. The perception of this motion blur is then measured and a model of perception is proposed. Finally, this perception model is used to design an objective metric which enables the prediction of the loss of quality on LCD with respect to the perceived quality on CRT.

5.1 Description

The LCD motion blur has been widely studied in recent works [9, 10, 11]. It's mainly caused by the hold-type LCD's displaying method: the light intensity is maintained on the screen for the duration of the frame, whereas on CRT light intensity is a pulse which fades over the frame duration.

The main difference happens when the eyes of the observer are tracking a moving object on the LCD screen: for a given frame, the picture is sustained on the screen while the eyes are still moving slightly anticipating the movement of the object. The edges of this object are displaced on the retina resulting in a blur [12].

5.2 Motion blur perception

In order to measure the relation between the motion velocity and the magnitude of the perceived blur, psychophysics measurements have been designed [13]. The results of these experiments are presented in Figure 1, they lead to the following linear relations :

$$W = aV, \tag{2}$$

$$W = avT. \tag{3}$$

The width W (in pixels) of motion blur that appears on the edges of a moving object is proportional to its velocity V (in pixels per frame) as depicted in Equation 2. This can be expressed as a function of the video period T, with v the velocity of the movement in pixels per second (Equation 3).

Pan *et al.* have developed a theoretical model of LCD motion blur perception [10] and obtain the same relation. Their model permits to identify the parameter a, which depends on the temporal function of the display.

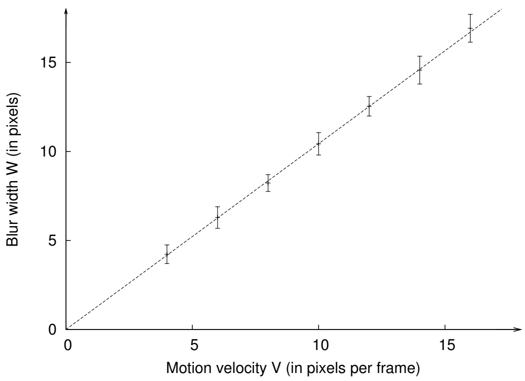

Figure 1: Perceived blur width W as a function of motion velocity V.

5.3 Prediction of ΔMOS based on the LCD motion blur

Using the LCD motion blur perception model, an objective metric is designed in order to predict the loss of quality ΔMOS between CRT and LCD highlighted by the subjective assessment on HD video. This metric is made in several steps. First, a spatio-temporal classification is done in two passes. First pass is corresponding to a block based motion estimation that leads to the construction of tubes which are the sets of blocks positions along the direction of motion. Second pass is the classification of each tube according to its spatial content. Since motion blur is only visible at sufficient contrast [14], only tubes categorised as textures and edges are selected. An average motion vector is computed from all the vectors of the remaining tubes. Norm of this global vector is used to compute the width of perceived motion blur according to Equation 2. This value W is an indicator of the average magnitude of perceived blur along the sequence. Finally, the prediction of the loss of quality ΔMOS_p is computed from a function of W. This function is non linear since there is no influence on perceived quality below a threshold of W, and the quality difference saturates for high values (cf. Figure 2).

An estimation of the subjective quality scores on LCD from the subjective quality scores on CRT can be made using the following relation:

$$\text{MOS LCD}_{est} = \text{MOS CRT} - \Delta\text{MOS}_p. \quad (4)$$

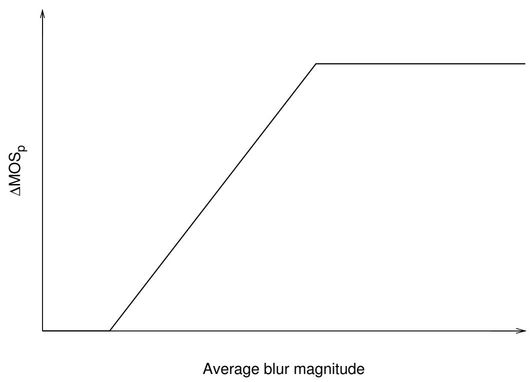

Figure 2: Prediction of the loss of quality ΔMOS_p from the average blur magnitude.

The quality of the model can be measured by the linear correlation coefficient (CC) and the root mean square error (RMSE) between the estimated LCD scores and the actual LCD scores. Values of 0.958 for CC and 1.30 for RMSE are obtained. These performances can be compared with those obtain comparing results between the CRT scores and the LCD ones: 0.241 for CC and 6.75 for RMSE.

Therefore, the proposed prediction model is able to evaluate some LCD improvements designed by manufacturers to reduce motion blur.

6. IMPACT OF LCD MOTION BLUR WITH RESPECT TO THE DISPLAY RESOLUTION

As it has been shown in the previous section, a high correlation exists between the magnitude of the perceived blur in a sequence and the loss of quality observed on LCD when displaying this sequence. Furthermore, the magnitude of the perceived blur depends on the display characteristics and resolution. More precisely in Equation 2, the motion velocity V (in pixels per frame) is proportional to the resolution. If the resolution is reduced by a factor N, the motion velocity would be reduced too by the same factor. As a consequence, the perceived blur would be smaller and should be less annoying as the resolution decreases.

6.1 Standard definition video quality assessment

In order to assess the impact of the LCD motion blur relatively to the resolution, the same video quality assessment tests have been realised with sequences at a standard definition (SD). Four sequences of the previous experiment have been chosen and reduced to SD resolution by computing the HD versions through a half-band filtering followed by a down-sampling by a factor of 2 (both along horizontal and vertical directions). This processing is performed on each field of the interlaced 1080i sequences. The obtained resolution (540i) does not match exactly to actual SD resolution (570i), but a down-sampling factor of 2 has the benefit to not necessitate any interpolation.

The tests have been led both on the CRT display and then on the LCD. Viewing conditions and display parameters are the same as those described in Section 2.1, except for the viewing distances which have been set to six times the pictures' height ($6H$). The SD sequences have been displayed inserted in a HD resolution grey level sequence in order to suit the displays native HD resolution.

Each reference (uncompressed) sequence has been distorted with the H.264 reference coder at seven different bitrates. The set of SD sequences to assess is constituted by these seven distorted sequences and by the explicit and hidden reference sequences.

6.2 Results

In this part, not only the scores of the hidden reference sequences are taken in account. In order to have a significant number of sequences, some distorted ones are considered too. As it's the impact of technology on the perceived quality which is assessed, only the sequences with very few coding distortions are taken in account. The mean of good-to-excellent MOS (above 65 on LCD) is computed for each content, same sequences are considered on each display. Table 4 shows these average MOS for good-to-excellent quality coded version of the four HD sequences, on the two displays. Same results for SD sequences is shown in Table 5. As expected, the loss of quality on LCD is strongly smaller in SD resolution relatively to HD resolution. Moreover, the loss of quality on SD sequences is not significant with regards to intervals of confidence.

As explained before, the perception of LCD motion blur closely depends on the display resolution. The quantity of perceived blur is proportional to the velocity of motion which is twice smaller in SD than in HD (since the resolution is divided by two). As a result, the perceived motion blur should be less annoying in SD and the advantages of LCD such as colourfulness and a larger luminance range seem to tower over this artifact, leading to a better global perceived quality on LCD display.

Sequences	MOS CRT	MOS LCD	ΔMOS
MOBCAL	79.5	71.9	7.6
PARKRUN	83.0	70.7	12.3
SHIELDS	81.4	67.9	13.5
STOCKHOLM	81.6	73.0	8.7

Table 4: Mean of the MOS for good-to-excellent quality sequences in HD resolution, on CRT and LCD.

Sequences	MOS CRT	MOS LCD	ΔMOS
MOBCAL	71.6	68.0	3.6
PARKRUN	77.6	72.9	4.7
SHIELDS	75.5	73.5	2.0
STOCKHOLM	75.7	73.2	2.5

Table 5: Mean of the MOS for good-to-excellent quality sequences in SD resolution, on CRT and LCD.

7. CONCLUSION

The subjective quality evaluation of HD moving pictures have shown that the perceived quality is better on CRT display than on LCD. This loss of quality ΔMOS seems to be due to the flat panel technology. Actually, some new artifacts such as motion blur are very annoying with quick movements. The benefits of LCD, for instance colourfulness and larger luminance range, have been highlighted with subjective assessment of still pictures: they lead to a better perceived quality on LCD than on CRT. However, in video they don't achieve to compensate the loss of quality due to the moving artifacts.

The LCD motion blur have been studied and a mathematical model is used to measure its magnitude as a function of the quantity of movements. A high correlation has been highlighted between the motion blur magnitude and the loss of quality on LCD which enables the prediction of the loss of quality ΔMOS between CRT and LCD.

Furthermore, the magnitude of the perceived blur depends on the velocity of moving objects and this velocity is linked to the display resolution. Consequently, the loss of quality ΔMOS would depend on the display resolution too and should be weaker with some lower resolutions than HD. Similar video quality assessment tests with SD sequences have confirmed this. At a lower resolution, the moving artifacts due to LCD technology are less annoying: the qualities of flat panels (colourfulness, a larger luminance range, etc.) seem to tower over these defects since the perceived quality is better on LCD than on CRT in SD resolution.

In this paper, it's shown that the new LCD technology leads to new shortcomings when displaying moving pictures. These new artifacts are not significant at low resolutions, which validate the use of LCD for subjective video quality assessment in the Multimedia Testplan of the VQEG [4]. However, when increasing the resolution, the artifacts due to the flat panel technology become more annoying and have an important impact on the perceived quality. It comes that the subjective video quality assessment at a high resolution (HDTV for example) should be led very carefully on LCD since a significant part of the perceived distortions could be due to the display.

8. ACKNOWLEDGMENT

This study is supported by a student grant from the French région Pays de la Loire, it falls within the framework of the French national research project Scalim@ges. The authors would like to thank Arnaud Tirel and Romuald Pépion for their assistance in performing the experiments described in the paper.

REFERENCES

[1] Takashi Fujio, "Future broadcasting and high definition television," NHK technical monograph, NHK, June 1982.

[2] ITU, "Report on results of comparative subjective picture quality assessment test between CRT and LCD," Questions ITU-R 95/6, 102/6, International Telecommunication Union - Radiocommunication Study Groups, August 2005.

[3] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," Tech. Rep., VQEG, 2003, http://www.its.bldrdoc.gov/vqeg/projects/frtv_phaseII/downloads/VQEGII_Final_Report.pdf.

[4] VQEG, "Multimedia group test plan," Tech. Rep. Draft version 1.16, VQEG, February 2007, http://www.its.bldrdoc.gov/vqeg/projects/multimedia/MM_new_testplan_v1.16_changes_accepted.doc.

[5] ITU, "Methodology for the subjective assessment of the quality of television pictures," Recommendation ITU-R BT.500-11, June 2002.

[6] ITU, "Subjective assessment methods for image quality in high-definition television," Recommendation ITU-R BT.710-4, November 1998.

[7] EBU, "SAMVIQ – Subjective assessment methodology for video quality," Tech. Rep., European Broadcasting Union, May 2003.

[8] Jean-Louis Blin, "New quality evaluation method suited to multimedia context: Samviq," in *Proceedings of the Second International Workshop on Video Processing and Quality Metrics, VPQM'06*, Scottsdale, January 2006.

[9] Michiel A. Klompenhouwer, "The temporal MTF of displays and related video signal processing," in *IEEE International Conference on Image Processing, 2005. ICIP 2005.*, September 2005, vol. 2, pp. 13–16.

[10] Hao Pan, Xiao-Fan Feng, and Scott Daly, "LCD motion blur modeling and analysis," in *IEEE International Conference on Image Processing, 2005. ICIP 2005.*, September 2005, vol. 2, pp. 21–24.

[11] Xiao-Fan Feng, "LCD motion-blur analysis, perception, and reduction using synchronized backlight flashing," in *Proceedings of the SPIE Conf. Human Vision and Electronic Imaging XI*. Electronic Imaging 2006, Janvier 2006, vol. 6057.

[12] Taiichiro Kurita, "Moving picture quality improvement for hold-type AM-LCDs," *SID Symposium Digest of Technical Papers*, vol. 32, no. 1, pp. 986–989, June 2001.

[13] Stéphane Péchard, Sylvain Tourancheau, Patrick Le Callet, Mathieu Carnec, and Dominique Barba, "Towards video quality metrics for HDTV," in *Proceedings of the Second International Workshop on Video Processing and Quality Metrics, VPQM'06*, Scottsdale, January 2006.

[14] Justin Laird, Mitchell Rosen, Jeff Pelz, Ethan Montag, and Scott Daly, "Spatio-velocity CSF as a function of retinal velocity using unstabilized stimuli," in *Proceedings of the SPIE Conf. Human Vision and Electronic Imaging XI*. Electronic Imaging 2006, Janvier 2006, vol. 6057.