

Quality Assessment for CRT and LCD Color Reproduction Using a Blind Metric

B. Bringier*, L. Quintard*⁺ and M.-C. Larabi*

* *Université de Poitiers, Laboratory SIC, Bat. SP2MI Bvd Marie et Pierre Curie, BP 30179, 86962 Futuroscope cedex, France*

⁺ *Laboratoire National de métrologie et d'Essais, 29 Avenue Roger Hennequin, 78197 Trappes Cedex, France*

Received 10th November 2007; revised 17th April 2008; accepted 10th June 2008

Abstract

This paper deals with image quality assessment that is capturing the focus of several research teams from academic and industrial parts. This field has an important role in various applications related to image from acquisition to projection. A large numbers of objective image quality metrics have been developed during the last decade. These metrics are more or less correlated to end-user feedback and can be separated in three categories: 1) Full Reference (FR) trying to evaluate the impairment in comparison to the reference image, 2) Reduced Reference (RR) using some features extracted from an image to represent it and compare it with the distorted one and 3) No Reference (NR) measures known as distortions such as blockiness, blurriness, . . . without the use of a reference. Unfortunately, the quality assessment community have not achieved a universal image quality model and only empirical models established on psychophysical experimentation are generally used. In this paper, we focus only on the third category to evaluate the quality of CRT (Cathode Ray Tube) and LCD (Liquid Crystal Display) color reproduction where a blind metric is, based on modeling a part of the human visual system behavior. The objective results are validated by single-media and cross-media subjective tests. This allows to study the ability of simulating displays on a reference one.

Key Words: No Reference metric, quality assessment, Display technologies ,color reproduction, single-media and cross-media validation.

1 INTRODUCTION

Image quality assessment is a very important activity for many image applications. The best way to assess the quality of an image is to ask observers to look at it as the Human Visual System (HVS) is the end-receiver in most processing environments. However, this approach is tedious, time consuming and expensive for practical usage. Moreover, it requires a normalized environment ensuring the best conditions for the targeted application. The most frequently used method for quality and/or fidelity measurement is the computation of parametric models that include characteristics of either the image or the application. Great efforts have been made over the last few years to develop objective image quality metrics that correlate with the perceived quality measurement with unfortunately relatively limited success. Indeed, in most cases, these metrics are limited to measuring differences between images before and after processing. According to the application, models including some

Correspondence to: <quintard@sic.sp2mi.univ-poitiers.fr>

Recommended for acceptance by David Fofi and Ralph Seulin

ELCVIA ISSN:1577-5097

Published by Computer Vision Center

HVS properties are integrated in the developed metrics. However, display device characteristics are not often taken into account in these models.

Quality assessment metrics could be classified into three categories: Full Reference (FR), Reduced Reference (RR) and No Reference (NR) metrics. In practice, FR methods may not be usable since the reference image is often unavailable. Hence, it is necessary to assess the quality by using the impaired image only. Unfortunately, NR quality measurement is an extremely difficult task because many unquantifiable factors play a role in the way the observers assess quality, such as aesthetics, cognitive relevance, learning, context, user background and so on. Consequently, most proposed NR quality metrics are designed for one or a set of predefined specific distortion types and cannot be extrapolated to images with other distortions. In image compression applications, for instance, NR quality metrics measure structural artifacts such as blockiness, blurriness, ringing due to coding algorithms [14, 17]. Because of the complexity of the blind quality assessment field it is largely unexplored. RR metrics are preferred in certain cases because they allow to embed some characteristics of the image in order to be compared to those extracted after the process (transmission for instance).

In this paper, we focus on a blind quality assessment using a NR metric for color reproduction applications dedicated to displays. The development of new display technologies such as LCD (Liquid Crystal Display) or OLED (Organic Light Emitting Diode) do not simplify the task of quality assessment of image reproduction. This could be explained by the large variation in gamut, tone reproduction curves, ... that can exist between LCD, CRT or any other technology. Thus, metrological approaches defined by standard organizations such as VESA [18] or CIE [6], that measure the physical performance of displays, are somewhat limited when the measurement of the subjective quality of rendering is needed. It is thus natural to ask observers to take part in the evaluation in order to measure the defects of the various monitors from a subjective point of view. Unfortunately, this stage is tedious and, time and money consuming. Novel studies, like the one presented in this paper, are necessary to integrate the human observer in the assessment loop. *Watson*[20] proposed a method of observer integration to determine the motion blur and local contrast defects. *Bringier et al.*[3] described a method to measure the differences, in terms of perception, between various display technologies. These methods do not allow to directly predict the subjective quality of a color reproduction. They are limited to the detection of defects and/or a comparison measurement without subjective quality information. This is why it is necessary to provide new approaches and methodologies to answer the problem.

For subjective quality applications, the contrast is generally considered to be one of the most important quality parameters [22, 7]. It is commonly defined in terms of tone reproduction curves for color reproduction applications. Unfortunately, two sets of images having very different white and black points may have very different perceptual contrasts. So, image quality cannot be established from the tone reproduction curves. Consequently, empirical models based on psychophysical experiments have been developed to compute the perceived quality with regards to the contrast of the image. The most successful model uses a simple definition of Lightness-Contrast, Chroma-Contrast and Sharpness-Contrast [5] in the CIE $L^*a^*b^*$ color space. However, the parameter weights in this type of models depend highly on the set of images used in the subjective experiments.

To solve this problem, we propose a new no-reference algorithm based on an HVS modeling. Initially, we compute the perceived information by using a hardware reproduction performed by the selected displays and a software reproduction performed by a simulation of the display characteristics. Then, a color local contrast definition is used to assign quality scores. Finally, we validate our quality measurement thanks to subjective experiments and analyze the correlation between the metric predictions and the observer ratings. The subjective experiments are performed in two ways : using a single-media validation and using a cross-media validation. The latter allows the explorations of a field where no assessment standard exists.

The remainder of this paper is organized as follows : Section 2 is dedicated to color reproduction issues. In section 3, we describe the proposed no-reference metric. Section 4 presents the experimental methodology and the results are given in section 5. This paper ends with a conclusion and some future directions.

2 Color reproduction

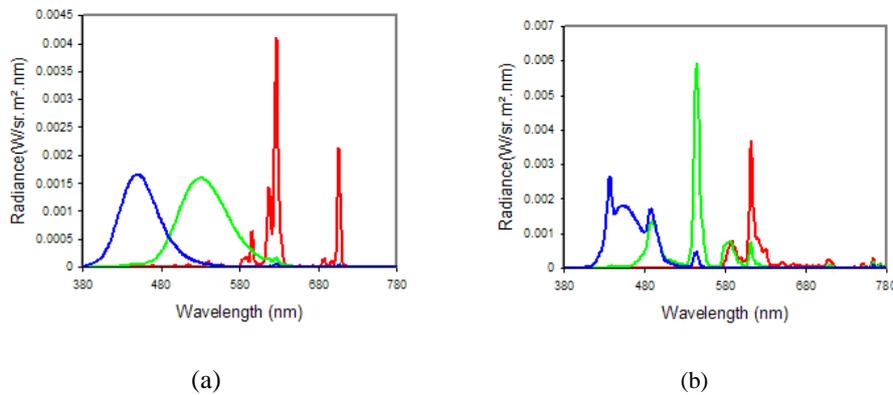


Figure 1: Example of spectral radiance distributions for full color, red, green and blue for CRT (a) and LCD (b)

A large numbers of new display technologies are introduced for the needs of emergent applications which generate different color reproductions of the same material. For instance, LCD or plasma displays do not have the same reproduction as conventional CRT. Some characteristics like tone-reproduction curve or spectral and basic colorimetric characteristics can change the color reproduction and, hence, the image perception of an observer. Figure 1 shows spectral differences between CRT and LCD displays.

To model the luminous field emitted by the display for estimating the perceived information by the HVS, traditional methods described in literature [6] are adequate and do not need improvements for the needs of this work. So, for the purpose of this work, we selected the S-Curve characterization [12] because it allows approximating as well as CRT and LCD displays.

3 Proposed no-reference metric

The human perception is able to naturally define quality standards to classify a set of images. It is thus natural to try to use the HVS behavior to develop a quality model for color reproduction. The proposed method for predicting quality uses the contrast definition of *Peli* [15] dedicated to color images. Figure 2 summarizes the approach of Local Band-limited Contrast (LBC) computation.

The LBC is performed in the AC_1C_2 color space [16] after the CSF (Contrast Sensitivity Function) filtering stage. The CSF is one of the most used models of the HVS and allows to represent its sensitivity with regards to spatial frequencies [13]. Several CSF models exist and follow different strategies which means that the choice can have an impact on the performance of the developed metric. To compute color contrast, the image is filtered by a set of band-pass filters and fan filters like a cortex transform [19]. Four spatial frequency bands and four orientations allow the frequency decomposition. With this filtering, the radial frequency selectivity and the orientation selectivity are modeled. The effects of these filters are cascaded to describe the combined radial and orientation selectivity of cortical neurons [7].

Then, the frequency bands are converted into a measure of perceived contrast in the image for both chromatic and achromatic channels. Generally, *Weber* (eq. 1) or *Michelson* (eq. 2) contrasts are used to compute simple stimuli contrast.

$$C_{Weber} = \frac{\Delta L}{L}. \quad (1)$$

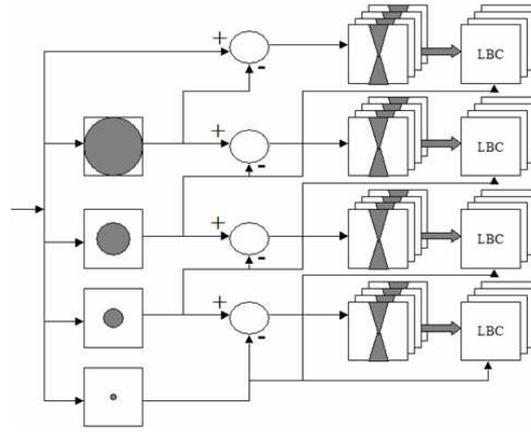


Figure 2: Local Band-limited Contrast method.

$$C_{Michelson} = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}. \quad (2)$$

Unfortunately, it is also obvious that none of these simple global definitions is appropriate for measuring the contrast of natural images, because a few points that are very bright or very dark would determine the contrast of the entire image. To solve this problem, *Peli*[15] proposed a local band-limited contrast definition given by the following equation:

$$LBC_k(x, y) = \frac{B_k(x, y)}{\sum_{i=0}^{k-1} B_i}, \quad (3)$$

where $B_k(x, y)$ is the band-pass filtered image of the k^{th} band, and $\sum_{i=0}^{k-1} B_i$ contains the energy below this band. In our model, we used a modified version of *Peli*'s contrast definition [21] given by this equation:

$$LBC_{k,l}(x, y) = \begin{cases} \frac{B_{k,l}(x,y)}{M_{k,l} + \sum_{i=0}^{k-1} B_{i,l}} & \forall k = 2 \dots K, l = 1 \dots L \\ \frac{B_{k,l}(x,y)}{M_{k,l} + B_0} & \forall k = 1, l = 1 \dots L \end{cases} \quad (4)$$

where B_0 is the average of the image defined by the center of *Daly* frequency decomposition and $M_{k,l}$ is the function that can be used to introduce the average of the image and to model the frequency and orientation sensitivity of the HVS. $M_{k,l}$ gives the flexibility to take into account the influence of chromatic and achromatic local average and the range of the display device. Moreover, $M_{k,l}$ avoids a division by zero in the original *Peli* equation.

At this stage, the color contrast information is combined to provide a global contrast assessment. Coefficients are used according to frequency and orientation decompositions. A factor of 2 is affected to achromatic channels compared to chromatic channels according to the color coding in the human visual system [1, 4]. The same weighting factor is affected to high frequency information compared to average frequency one. Thus, luminance contrast is considered more important than chromatic contrast. Indeed, in this study, the changes of contrast in images according to the tone reproduction curves appear primarily in luminance. Moreover, the sharpness depends primarily on the high frequencies; It is thus obvious to assign a high factor to them. It is also consistent with the neurophysiologically-based encoding functions.

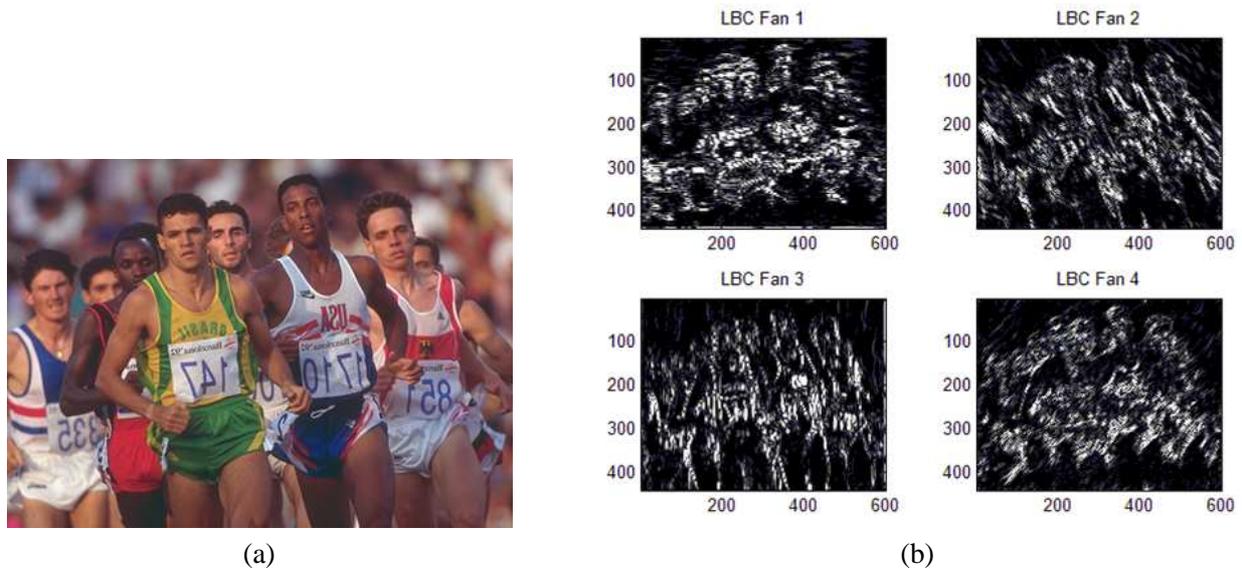


Figure 3: a- Original image, b- Local band-limited contrast image for average frequency and four orientations.

4 Subjective assessment methodology

Subjective tests provide the foundations for building vision models[11]. At the same time, they are the only true benchmark for evaluating the performance of the various perception-based image processing. Unfortunately, perceptual responses cannot be represented by exact figures; due to their inherent subjectivity, they can only be statistically described. Even in psychophysical threshold experiments, where the observer's task is just to give a yes/no answer, a significant variation between the observers could exist. Although the subjective experiment is time-consuming, it is considered as the main way to validate the metric results by studying the correlation existing between them.

4.1 Experimental conditions

In this section, we describe the most important conditions and the values that we have chosen to set up the experiments.

4.1.1 Psychophysical test-room

This room is constructed in compliance with the recommendations of the ITU [8]. It is lighted by neon tubes D50 for an ambient illumination of 62 ± 4 lux approximately. The walls of the psychophysical test-room are painted in neutral gray in order to minimize the stray reflections on the display (see figure 4).

4.1.2 Displays

The monitors used for the assessment have a diagonal of 17". The calibration of the CRT and LCD displays was carried out using a spectro-colorimeter PR-650. The white of the monitor is around $6500 \pm 100K$ for a brightness of 80 cd/m².

4.1.3 Observers

The observer panel consists of a coherent set of non-expert people representing different age, gender and socio-professional categories. Before each test sequence, this panel undergoes a visual acuity test (Snellen) and color blindness test (Ishihara).

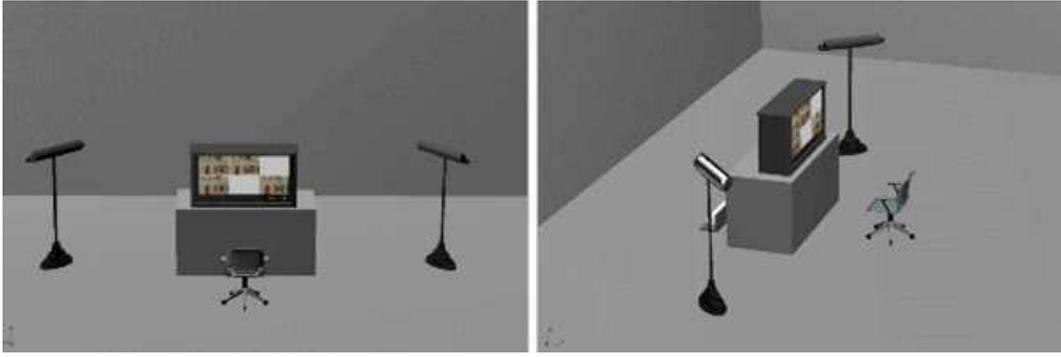


Figure 4: psychophysical test-room setup

4.2 Subjective assessment protocols

The aim of this section is to validate our metric by using subjective assessment. To obtain significant subjective data, the ITU [8] gives recommendations concerning the methodology, the environment and the data analysis. Nevertheless the assessment methodology is based on a single device procedure.

In order to avoid the bias introduced by the simulation of tone reproduction curves of the displays on the same and unique device, we managed a cross-media validation based on the assessment of different displays at the same time.

The two tests were performed in the same environment in order to preserve the reliability of the results. The analysis of the assessment results is decomposed in two stages. First, we compute the MOS (Mean Opinion Score) which is the average of the observers results and the 95% confidence interval inside which the results are considered as the most reliable. Because the reference case is a normal distribution, we need to run the kurtosis test to reject outliers observers and/or observations.

The kurtosis coefficient β_2 (i.e. the ratio between the fourth-order moment and the square of the second-order moment) is given by [10]:

$$\beta_2 = \frac{m_4}{(m_2)^2}, \quad (5)$$

$$\beta_{2jkr} = \frac{\frac{1}{N} \sum_{i=1}^N (\bar{u}_{jkr} - u_{ijkr})^4}{\left(\frac{1}{N} \sum_{i=1}^N (\bar{u}_{jkr} - u_{ijkr})^2 \right)^2}, \quad (6)$$

where u_{ijkr} is the score of the observer i for the degradation j of the image k and the r^{th} iteration. N represents the number of observers. If β_{2jkr} is between 2 and 4, we can consider that the distribution is normal. The different steps of the algorithm are summarized below :

Algorithm 1: Steps for outliers rejection

```

if ( $2 \leq \beta_{2jkr} \leq 4$ ) /* (normal distribution) */ then

```

```

  if ( $u_{ijkr} \geq \bar{u}_{ijkr} + 2\sigma_{jkr}$ ) then
    |  $P_i = P_i + 1$ ;
  endif

```

```

  if ( $u_{ijkr} \leq \bar{u}_{ijkr} - 2\sigma_{jkr}$ ) then
    |  $Q_i = Q_i + 1$ ;
  endif

```

```

endif

```

```

else

```

```

  if ( $u_{ijkr} \geq \bar{u}_{ijkr} + \sqrt{20}\sigma_{jkr}$ ) then
    |  $P_i = P_i + 1$ ;
  endif

```

```

  if ( $u_{ijkr} \leq \bar{u}_{ijkr} - \sqrt{20}\sigma_{jkr}$ ) then
    |  $Q_i = Q_i + 1$ ;
  endif

```

```

endif

```

```

/* Finally, we can carry out the following eliminatory test : */

```

```

if ( $\frac{P_i+Q_i}{J.K.R} > 0.05$ ) and ( $|\frac{P_i-Q_i}{P_i+Q_i}| < 0.3$ ) then
  | Eliminate scores of observer  $i$ ;
endif

```

```

/* Where  $J$  is the total number of degradations,  $K$  is the total number of images and  $R$  is the total number of iterations. */

```

4.2.1 Single device validation

The aim of this test is to obtain a classification of the various monitors. The assessment of the modified images is performed on a single device. The image modification lies in a variation function of the tone reproduction curves it represents. Figure 5 details the synoptic of the followed methodology for studying the correlation between objective and subjective measurements.

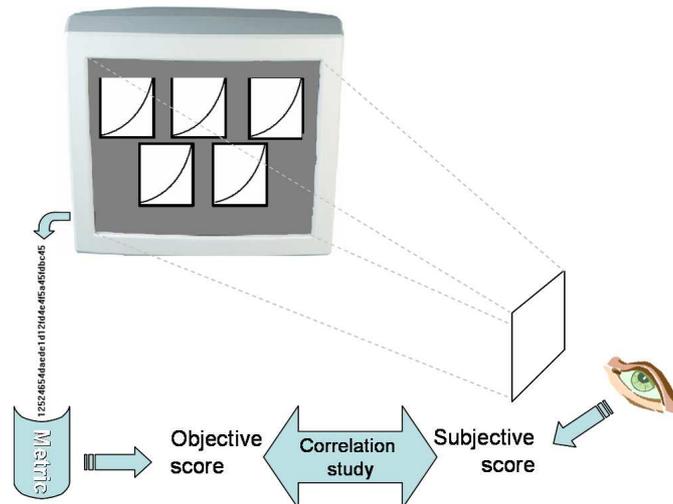


Figure 5: General Synoptic of single-media validation.

In this test, we have nine versions of the same image by using nine different tone reproduction curves, each representing results obtained in the characterization stage. These versions are displayed simultaneously on the

same monitor described above. The observer is asked to answer the following question : *The same image reproduced in different ways will be displayed. You are asked to select the worst representation by clicking on it. The image you have selected will be masked and you are requested to do this again eight times*

4.2.2 Cross-media device validation

During the last subjective experiment, the variations in the monitors' gamut were not taken into account. This is due to the use of a single monitor to simulate the entire set of monitors. The aim of the current test is to assess the color reproduction quality by using the monitors themselves. The image modification is not done by using a software but by the inherent characteristics of the monitor. What differs from the previous test, in order to avoid a too important variation of all monitor parameters, is the use of only LCD technology displays. Note that cross-media evaluation is very tedious and no standard exists. So, this limitation to LCD only allows the reduction of the number of parameters to be tuned. Figure 6 shows the synoptic cross-media validation flow-chart of the cross-media methodology. Thirty non-expert observers performed the test.

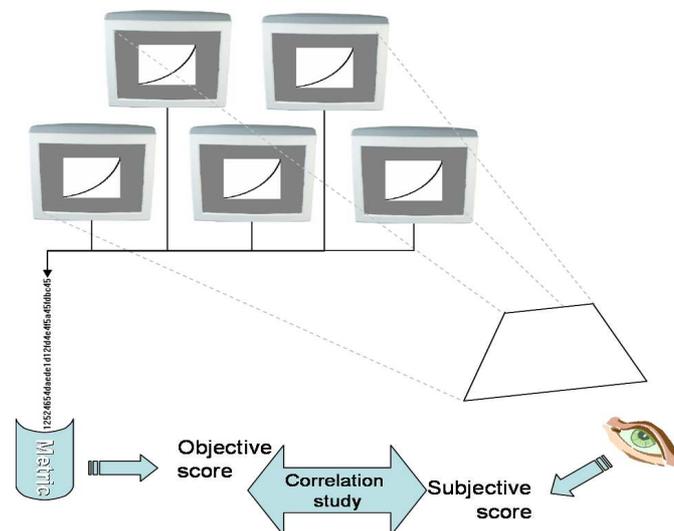


Figure 6: General synoptic of cross-media validation.

For this test, the image was displayed on the five monitors at the same time. The observer was asked to rank the color reproduction from the worst to the best. In this case, the image on the monitor was not masked allowing the observer to recheck his choice. The question he was asked to answer is quite similar to that of the single-device test.

5 Experimental results

To confirm the perceptual relevance of our metric, we carried out two sets of subjective experiments. First, a database, with twelve test images representing typical images used in multimedia applications (Corel Photo), was created by simulating nine tone reproduction curves that can be typically obtained in CRT and LCD displays. This experiment is used to validate the performance of the no-reference metric. Then, another experiment is performed on 5 LCD monitors to validate our method through a cross-media test. Figure 7 shows an example of the images used for the subjective experiments.



Figure 7: Example of test images for the quality assessment experiments (-a- Synthesis, -b- Transport).

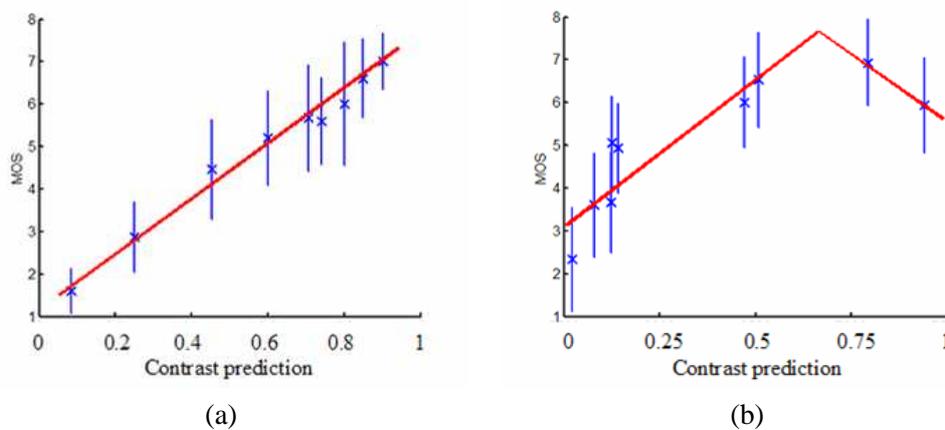


Figure 8: Error-bar plot with 95% confidence intervals of subjective ratings versus no-reference perceptual contrast measurement for two images of the database (a- Transport, b- Synthesis).

5.1 Single media results

The prediction model of perceived color contrast is applied to the entire set of images used during the subjective test. Figure 8 shows a part of the most significant results (complete results could be found in [2]). One can notice two types of curves: for "Transport" test image (see figure 7), we obtain a linear correlation between MOS and perceived contrast prediction. In this case, we can conclude that perceived contrast prediction could directly define the quality prediction. Thus, for the entire set of images that have the same results (seven images from the database), we obtain a correlation between perceived quality and subjective results as high as 90%.

For other test images from the database, we obtain slightly different results. Initially, the correlation between perceived contrast and subjective judgment is positive. Then, after a given threshold, we obtain a negative correlation (figure 8). However, too much contrast in an image can decrease its visual quality as explained by *Janssen* [9]. In the framework of his work, a psychophysical experimentation was used to determine the quality and the naturalness of an image based on the contrast.

The global color contrast value is not sufficient to determine quality. A threshold, function of contrast, must be applied to obtain a quality measurement. Figure 9 shows the results of quality prediction for "Synthese" image. Similar results are obtained for the three other images whose have the same subjective results. Thus, on the whole image set, we obtain a correlation between quality perceived and subjective results as high as 90% as previously mentioned.

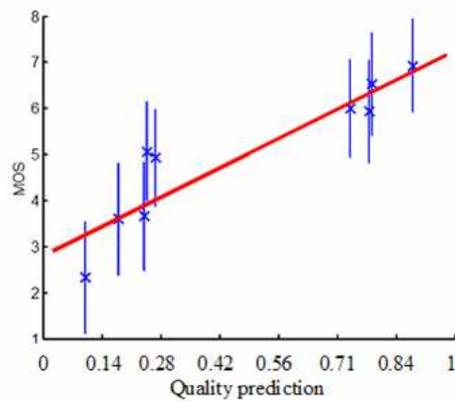


Figure 9: Error-bar plot with 95% confidence intervals of subjective ratings versus no-reference perceptual quality measurement for "Synthese" image.

5.2 Cross-media results

In this section, we present the results of the cross-media validation previously described. This approach is quite new because it allows the assessment of several displays at the same time by exploiting their own characteristics. As mentioned in the dedicated section, this methodology is tedious and time consuming. That is why it is not always possible to repeat it. However, the extracted results are very interesting for the metric validation and for future works on objective assessment.

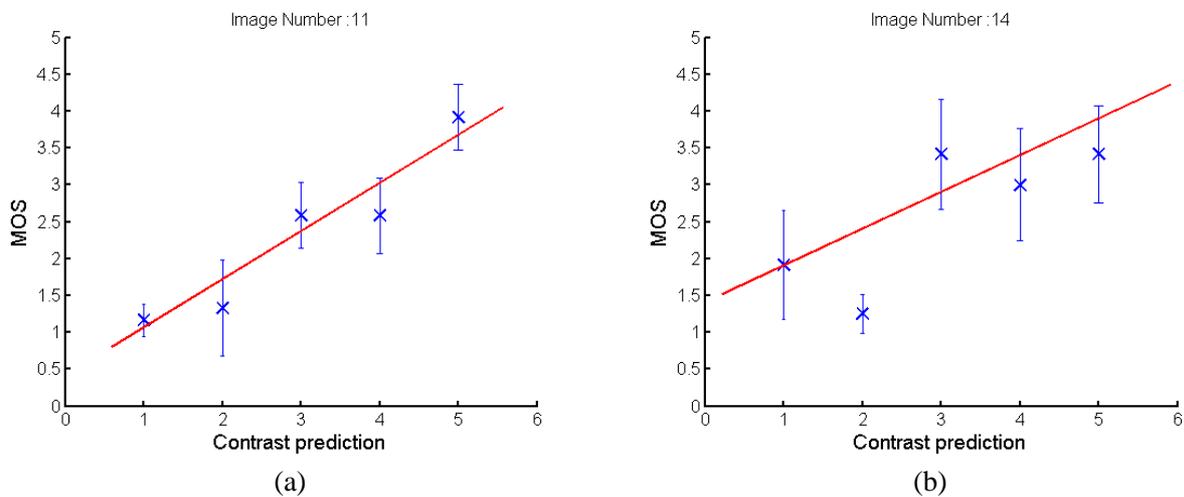


Figure 10: Error-bar plot with 95% confidence intervals of subjective ratings versus no-reference perceptual contrast measurement for two images of the database (a- Image "11", b- Image "14").

The metric is applied to a cross-media evaluation to verify the results obtained in the first experimentation. Figure 10 shows the results for the five LCD displays for two pictures (a and b). The X-coordinate represents the prediction of contrast and the Y-coordinate the subjective scale. The error-bar plot with 95% confidence intervals represents the data obtained for one display. The observer's preferences are classified to the display characteristics. The red bar shows the correlation between the subjective and objective data. We can observe a good correlation between them for picture 11 and a not as good correlation for picture 14.



Figure 11: Example of test images : -a- Image 11, -b- Image 14

To compute the correlation ratio, we used two well known coefficients : Pearson r_p and Spearman r_s . Pearson coefficient is given by the following equation :

$$r_p = \frac{n \sum_{i=0}^n X_i Y_i - (\sum_{i=0}^n X_i)(\sum_{i=0}^n Y_i)}{\sqrt{[n \sum_{i=0}^n X_i^2 - (\sum_{i=0}^n X_i)^2] [n \sum_{i=0}^n Y_i^2 - (\sum_{i=0}^n Y_i)^2]}} \quad (7)$$

where n is the number of pairs of scores. The degree of freedom is $df = n - 2$. Spearman coefficient is given as follows :

$$r_s = 1 - \frac{6 \sum_{i=0}^n d_i^2}{n(n^2 - 1)} \quad (8)$$

The Pearson correlation between the subjective and the objective results is globally equal to 0.78. It indicates a good relation between the variables. The Spearman coefficient is equal to 0.90 with a p-value < 0.05 . We obtain thus a good correlation between the subjective and the objective tests.

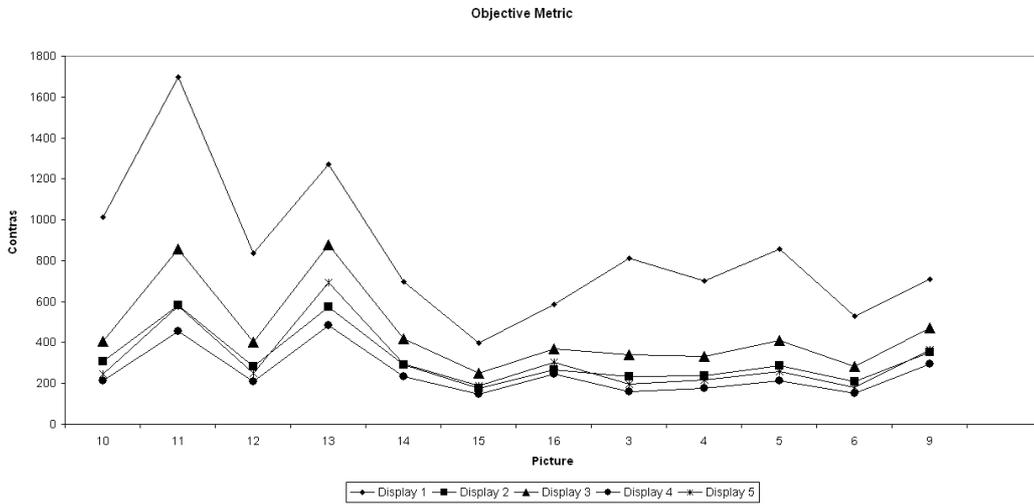


Figure 12: Contrast measurement for each display and each image.

Table 1 compiles the results for each image and allows to analyze the results. Figure 12 shows the predicted contrast for each image. We can observe a high variability and we obtained a poor correlation for image 16 only. For the other images the results are relatively acceptable. The best result is obtained for image 11 (figure 11-a).

Since during the subjective test the observer was asked to classify the different reproductions, the results must be analyzed by using the Spearman coefficient which exploits the notion of rank order. Table 2 gives the different Spearman coefficients. As we can notice the correlation is satisfactory in all cases.

Picture label	13	16	3	4	5	6	15	14	9	10	11	12
Coeff. Cor	0.73	0.39	0.75	0.79	0.85	0.74	0.62	0.54	0.87	0.75	0.89	0.64

Table 1: Pearson correlation for each image

Picture label	13	16	3	4	5	6	15	14	9	10	11	12
Coeff. Cor	0.70	0.50	0.90**	0.80	0.90**	0.90**	0.70	0.35	0.70	0.9**	1.00	0.56

Table 2: Spearman coefficient correlation for each image; ** p-value < 0.05

6 Conclusion and future works

In this paper, we presented a no-reference metric for quality assessment of color reproduction devices (displays). This type of quality prediction application is rare in literature. Indeed, a full reference model is generally used to evaluate the performance of image processing systems and no-reference models to evaluate a precise artifact like blockiness in compression. Moreover, we used the human visual properties to develop our no-reference model. Consequently, we obtained a generic measurement that enables to evaluate the quality of an image that does not have deformations like blockiness. Finally, the model is validated with two different subjective tests. The first one is based on the contrast change in an image. A correlation between the results of this test and our quality prediction higher than coefficients can be refined to improve the prediction performance. Then, another subjective test is performed to validate our approach for soft copy quality evaluation. The results give a good correlation between our method and the subjective data. Nevertheless, we want to achieve a better correlation in the future. For this we need to incorporate more criteria into the metric.

References

- [1] J. J. Atick, Z. Li, and A.N. Redlich. Color coding and its interaction with spatiotemporal processing in the retina. Technical report, Princeton's Institute of Advanced study, Princeton, New Jersey, 1990.
- [2] B. Bringier. *Elaboration de modèles d'évaluation de la qualité du rendu couleur*. PhD thesis, Université de Poitiers, 2005.
- [3] B. Bringier, N. Richard, D. Renoux, and C. Fernandez-Maloigne. Human vision and screen technology modelization for colour image quality assessment. In *10th congress of the international colour association AIC colour 05*, Granada, Spain, Mai 2005.
- [4] G. Buchsbaum and A. Gottschalk. Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proceeding of Royal Society of London*, 220:89–113, 1983.
- [5] A.J. Calabria and M.D. Fairchild. Perceived image contrast and observer preference II: Empirical modeling of perceived image contrast and observer preference data. *Journal of Imaging Science & Technology*, 47:494–508, 2003.
- [6] CIE. The relationship between digital and colorimetric data for computer-controlled CRT displays. Technical Report 122, CIE, 1996.
- [7] S. Daly. The visible differences predictor: An algorithm for the assessment of image fidelity. In A.B. Watson, editor, *Digital images and human vision*. MIT Press, 1993.
- [8] ITU-R. Methodology for the subjective assessment of the quality of television picture. Technical Report Recommendation BT.500-11, International Communication Union, 2002.

- [9] T.J.W.M. Janssen. *Computational Image Quality*. PhD thesis, Eindhoven, the netherlands, 1999.
- [10] D.N. Joanes and C.A. Gill. Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society*, 47:183, 1998.
- [11] B. W. Keelan. *Handbook of Image Quality: Characterization and Prediction*. Marcel Dekker, New York, NY, 2002.
- [12] Y. Kwak, C. Li, and L. MacDonald. Controlling color of liquid-cristal displays. *Journal of the SID*, 11(2):314–348, 2003.
- [13] M.-C. Larabi, V. Brodbeck, and C. Fernandez-Maloigne. A novel approach for constructing an achromatic contrast sensitivity function by matching. In *IEEE International Conference on Image Processing (ICIP)*, pages 441–444, Atlanta, GA, October 2006.
- [14] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi. Perceptual blur and ringing metrics: application to JPEG2000. *Signal processing: Image Communication*, 19:163–172, 2004.
- [15] E. Peli. Contrast in complex images. *Journal of optical society of America*, 7, 1990.
- [16] A.B. Poirson and B.A. Wandell. Pattern-color separable pathways predict sensitivity to simple colored patterns. *Vision research*, 36(4):515–526, 1996.
- [17] H. R. Sheikh, A. C. Bovik, and L. Cormack. No-reference quality assessment using natural scene statistics: JPEG2000. In *IEEE Asilomar conference on Signals, Systems and Computers*, California, 2003.
- [18] VESA. Flat panel display measurements standard, Version 2.0. Technical report, Video Electronics Standards Association Display Metrology Committee, Juin 2001.
- [19] A.B. Watson. The cortex transform: rapid computation of simulated neural images. *computer vision, graphics an image processing*, 39(3):311–327, 1987.
- [20] A.B. Watson. The spatial standard observer: a human vision model for display inspection. *SID Symposium Digest of Technical Papers*, 37:1312–1315, 2006.
- [21] S.J.P. Westen, R.L. Legendijk, and J. Biemond. Perceptual image quality based on a multiple channel hvs model. *IEEE International conference on Acoustics, Speech and Signal Processing*, 4:2351–2354, 1995.
- [22] S. Winkler. *Digital video quality - Vision models and metrics*. John Willey & Sons, 2005.