A PSYCHOVISUAL COLOR IMAGE QUALITY METRIC INTEGRATING BOTH INTRA AND INTER CHANNEL MASKING EFFECT

Christophe Charrier[a] *and Thierry Eude*[bc]

[a] University of Caen - Basse Normandie
LUSAC EA 2607, Vision and Image Analysis Group
120 route de l'exode, 50000 - Saint-Lô, France
phone: +33 233 77 11 66, fax: +33 233 77 11 67
email: c.charrier@chbg.unicaen.fr

[b] Université de Bourgogne
Le2i, UMR CNRS 5158,
BP 47870, 21078 Dijon Cedex, France

[c] Laval University
LSVN, cité universitaire
Québec, QC, Canada, G1K 7P4
phone: (418) 656-5669, fax: (418) 656-2324
email: thierry.eude@ift.ulaval.ca

ABSTRACT

Any process applied to digital images has to be validated by a performance measure. In the compression area, this measure of performance provides a quality measure of the reconstructed images. The use of psychophysical tests to measure the quality is quite time consuming. Therefore, many quality metrics have been defined in order to reach a high correlation degree with the human judgment. In this paper, a perceptually tuned metric based on a wavelet transform and a measure of the intra- and inter-channel visual masking effect is developed. A performance measure is then computed in terms of correlation and robustness to the type of image.

1. INTRODUCTION

Constraints of record, distribution and presentation of information require compression algorithms reducing the size of the original message. During the compression process, a quality metric has to be used to evaluate the perceptual distance between the original image and the reconstructed one, to optimize the compression rate with respect to the required quality. A quality metric has to evaluate the perception threshold of degradation for applications where differences between the original image and the reconstructed one are not allowed. Furthermore, this quality metric has to provide a scale of perceptible degradation for applications that do not need high quality (quality on demand).

Many quality metrics have been specifically developed to evaluate the quality of reconstructed images. The most recent integrate Human Visual System (HVS) models. Thus, they are able to take into account known phenomena such as the color representation, the contrast sensitivity as well as masking effects. To develop a quality metric, the widely used scheme consists in performing 1) a color space transformation to obtain decorrelated color coordinates, 2) a decomposition of these new coordinates from perceptual channels. Then an error is estimated for each one of these channels. The final quality measure is obtained from a weighted sum of these errors. Nevertheless, the process of the intra-channel errors alone does not allow a consideration of the masking effects due to existing interactions between different channels. Actually, FOLEY [1] has demonstrated that interaction between different channels exists. WATSON and SALOMON proposed a model integrating the existing interactions between luminance channels and chromatic channels [2]. ROSS and SPEED [3], introduced a model based on an adaptive term depending on one or more channels corresponding to the spatial frequencies and the orientation.

In this paper, a new quality metric based on a vision model integrating both intra and inter-channel masking effects is presented.

2. THE MEASURE OF THE QUALITY

Fig. 1 shows the used perceptual model integrated into the quality metric. A transformation of the (R,G,B) coordinated to the YC_bC_r color space [4] is performed. Then, from the obtained coordinates (Y, C_b, C_r), a wavelet decomposition is applied to obtain a multichannel decomposition for different frequencies and orientations. In this way, the decomposition performed by the HVS is respected as closely as possible. This transformation is performed using a 9/7 Daubechies filter [5]. The error is then determined for each subband, and for each coefficient w located in (i, j) from

$$e_b(i,j) = |w_b(i,j) - \hat{w}_b(i,j)| \quad (1)$$

where b is the considered subband.

The distortion measure $d_b(i,j)$ is then computed with respect to a masking coefficient $m_b(i,j)$

$$d_b(i,j) = \log\left(\frac{\alpha \cdot e_b(i,j)}{m_b(i,j)}\right) \quad (2)$$

where α is the parameter determined from the color contrast sensitivity. $m_b(i,j)$ enables a consideration of intra and inter-channel masking effects as follows :

$$m_b(i,j) = C_{\text{intra},b}(i,j).C_{\text{inter}}(i,j) \quad (3)$$

where $C_{\text{intra},b}(i,j)$ represents a visibility threshold applied to the coefficient located at (i,j) within the considered subband b. $C_{\text{inter}}(i,j)$ is a measure of inter-channel masking effects.

Thus, for each subband, a distortion measure is computed. The greater the masking effect is, the better the quality of (equ. 2).

Finally, a global score is obtained combining all of the computed distortion measures, using the Minkowski sum

$$E = \left(\sum_b \sum_{i,j} |d_b(i,j)|^\beta\right)^{1/\beta} \quad \text{with} \quad 1 \le \beta \le 4. \quad (4)$$

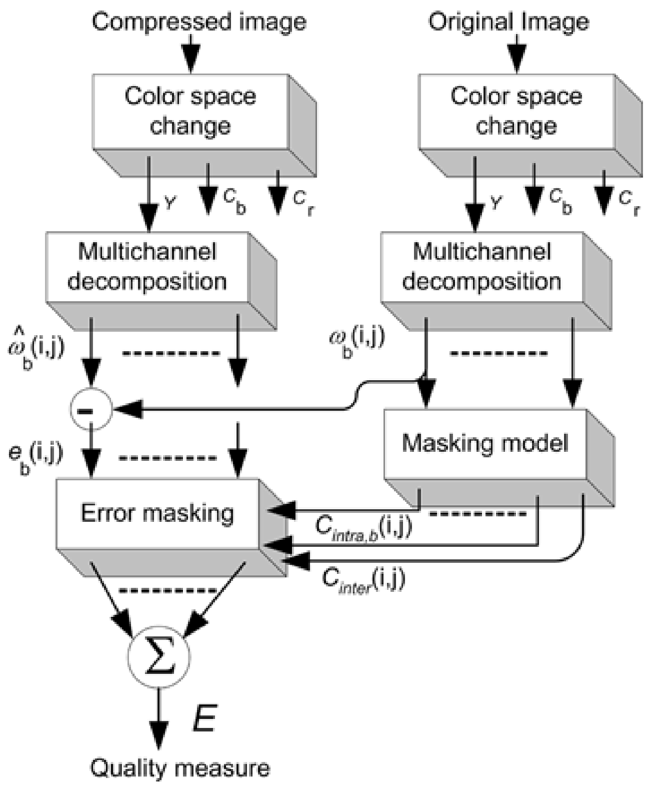

Figure 1: Perceptual model used.

2.1 RGB Components Transformation

A decorrelating transformation must be applied to the first three components of an image. Two goals have been achieved by this transformation, namely, color decorrelation for efficient compression and reasonable color space with respect to the human visual system for quantization. We have chosen to use the YC_bC_r color space. This space is widely used for compression of color images since one can easily reach high compression rates when the conversion is applied to a color image before compression schemes. The linear transformation of the R, G and B gamma corrected coordinates is

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix} + \begin{pmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{pmatrix} \begin{pmatrix} R' \\ V' \\ B' \end{pmatrix} \tag{5}$$

2.2 Masking model

Masking is an important visual phenomena that describes existing interactions between stimuli. The term of masking is used when the perception of a stimulus S1 is masked by the existence of a second stimulus S2. Since masking effects are present within the spatial frequencies subband, one has to take into account the masking effects within one subband (intra-band masking effect) and the masking effects due to the existing interactions between different frequencies subbands (inter-band masking effects), as described in equ. 3.

2.2.1 Intra-channel masking

A visible stimulus can be hidden by another one. In other words, the masking effect increases the visibility threshold with respect to the mask contrast. A simple intra-band masking model must clarify how masking can be parametrized. The model describes basically, how the contrast threshold , at which the signal becomes visible, varies with respect to the presence of a masker contrast. The non linearity of the function of the sensitivity threshold $C_{\text{intra},b}$ can be approximated by two piece-wise linear functions

$$C_{\text{intra},b}(i,j) = \max(1, w_b(i,j)^{\varepsilon}) \tag{6}$$

where ε is the slope-parameter of the curve present in Figure 2. A typical value of ε from psycho-visual experiments is $\varepsilon = 0.75$ [6].

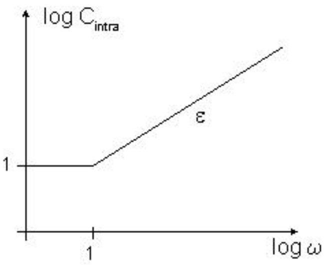

Figure 2: Threshold elevation characterized by slope-parameter ε.

2.2.2 Inter-channel masking

In order to model the increased masking provided by large coefficients in other subbands, the model adopted needs to take into account interband interactions over the three orientations and the spatial support. In other words, a coefficient predicting the masking effect over the three orientations is introduced. This coefficient, labeled as $h_{i,j}$ allows the computation of the inter-band masking effect located at (i,j) within the band u. The technique consists in using a set of locally neighboring wavelet coefficients of size $M \times M$ centered around position i,j over the three orientations. A frequential Gaussian mask ω_f is then applied on each neighborhood, in order to filter high frequencies, as carried out by the HVS. In each of the three subbands (corresponding to the three orientations) a region centered around the position i,j is weighted by a Gaussian kernel and summed. Then a spatial Gaussian mask ω_s is used to weight each of these sum. Hence the final result $h_{i,j}$ is given by

$$h_{i,j} = \sum_{(i,j)\in V_{i,j}} \omega_s(i,j) \sum_{u\in[1,2,3]} \omega_f(u)\ \hat{w}^2_{i,j,u}. \tag{7}$$

where u is the channel number that runs over the three orientations, and $V_{i,j}$ represents the neighborhood $M \times M$ centered around the position i,j. Here, the neighborhood is limited to a region of 7×7. This is due to the fact that:

1. for values lower than 7, the size of the neighborhood is too small to allow a correct and a whole evaluation of existing interactions between the three subbands,

2. for values greater than 7, the size of the neighborhood can take into account all existing interactions between the three subbands. Nevertheless, using a size of the mask greater than 7, the obtained precision gain induces a drastic growth of the complexity in terms of data access and mathematical operations.

In Equ. (7), the weighting coefficients are normalized by

$$\sum_{u\in[1,2,3]} \omega_f(u) \sum_{(i,j)\in V_{i,j}} \omega_s(i,j) = 1, \tag{8}$$

where ω_f and ω_s are taken from a Gaussian distribution

$$\omega(R) = \frac{1}{(2\pi)^{\frac{2}{3}}\sqrt{\|\Lambda\|}} e^{[-\frac{1}{2}(R^T)\Lambda^{-1}(R)]} \tag{9}$$

where $\Lambda = \mathrm{E}[(R-m)(R^T - m^T)]$ represents the variance-covariance matrix of the weighted mask.

The modeling of the inter-band masking can be formulated as follows

$$C_{\text{inter}}(i,j) = \max\left(1, h_{i,j}^{\gamma}\right) \tag{10}$$

where γ is a slope parameter.

3. EVALUATION OF PERFORMANCES

3.1 Methodology

There are a number of attributes that characterize a visual quality metric in terms of the ability of its estimation performance to respect the subjective ratings. These attributes are *accuracy*, *monotonicity*, and *consistency* according to processed type image (low-resolution scenery or strong contrasting textures).

Accuracy a_P is the ability of a metric to predict subjective ratings with a minimum average error and can be determined by means of the PEARSON linear correlation coefficient; for a set of N data pairs (x_i, y_i), it is defined as follows:

$$a_P = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}} \tag{11}$$

where $\bar{x}$ and $\bar{y}$ are the means of the respective data sets.

Monotonicity m_S is another important attribute as it measures if increases in one variable are associated with increases in the other variable. Ideally, changes of a metric is rating between different sequences should always have the same sign as the changes of the corresponding subjective ratings. The degree of monotonicity can be quantified by the Spearman rank-order correlation coefficient, which is defined as follows:

$$\begin{aligned} m_S &= \frac{\sum(\chi_i - \bar{\chi})(\psi_i - \bar{\psi})}{\sqrt{\sum(\chi_i - \bar{\chi})^2}\sqrt{\sum(\psi_i - \bar{\psi})^2}} \\ &= 1 - \frac{6(\psi_i - \psi)^2}{N(N^2-1)} \end{aligned} \tag{12}$$

where χ_i is the rank of x_i and ψ_i is the rank of y_i, and $\bar{\chi}$ and $\bar{\psi}$ are the respective midranks. The Spearman rank-order correlation is non-parametric, *i.e.* it makes no assumptions about the form of the relationship between the x_i and y_i. Then, to evaluate the performances of the proposed metric, we used on the one hand the results of our previous work [7] which consisted in a comparative evaluation of metrics proposed in the literature and on the other hand a recently developed metric [8].

So, in [7], it was shown that the metric proposed by KARUNASEKERA *et al.* offers a good compromise between accuracy and robustness regarding image type (from homogeneous to strong textured). In [8], LAI *et al.* developed a quality metric based on Haar wavelet decomposition. thus, they model some properties of the low level human vision (contrast sensitivity, intraband masking effect, etc.). This metric is correlated with rating obtained during psychophysic test of quality.

The evaluation of the proposed metric consists in comparing with the two metrics described previously, according to the three mentioned comparison criterions.

3.2 Results

The images used have been selected from the "LIVE Quality Assessment Database" [9] color image database, Figure 3 showing a sample. The database is composed of 29 original images (*i.e.*, noncompressed). Their complexity ranges from faintly textured to strongly textured.

(a) lighthouse (b) caps

(c) rapids (d) plane

Figure 3: LIVE database sample.

They are coded with 8 bits per channel and were compressed with JPEG2000 standard. The compression rates are chosen in order to obtain a reconstructed images quality rating distribution which is relatively uniform on the rating scale defined in the recommendation UIT-R BT.500-10 [10]. So, the database used totals 198 images.

In order to measure existing correlation between rating obtained by means of the proposed metric and average quality rating (*Mean Opinion Score*) obtained by means of an observers group, the evaluation quality results provided with the LIVE image database were used. In this way a highly boring observers rating process is avoided.

Table 1 presents results obtained for proposed, KARUNASEKERA *et al.* and LAI *et al.* metrics.

Metric	Accuracy a_p	Monotonicity m_s
Proposed	0.972	0.964
KARUNASEKERA	0.927	0.921
LAI	0.951	0.944

Table 1: Accuracy and monotonicity measured for each one of the three metrics.

The results reveal a greater average accuracy at 97%, as well as an average monotonicity at 96%. This denotes a strong corelation between subjectively obtained values and objectively obtained values (with the proposed metric).

Figure 4 presents quality values obtained by means of the proposed metric versus quality rating provided with the database. One notes that the metric remains correlated with the average observers rating. It denotes a robustness of the metric against the type of the images which are evaluated (homogeneous or textured).

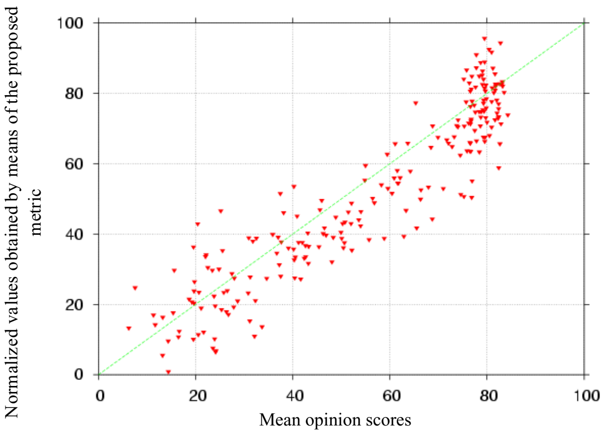

Figure 4: Average values of the observers opinions versus quality evaluated by the proposed metric.

4. CONCLUSION

A color compressed image quality metric integrating a masking effect (intra and interband) modeling was developed.

A change of spatio-colorimetric referential followed by a wavelet decomposition of the image are first processed on the original and compressed image. The intraband modeling used enables a quantification of the masking on only one band, whereas the developed interband modeling affects each of the three obtained orientations. Thus, the modeling provides a global measure of the masking effect.

The use of such modeling in the masking effect measure provides a metric which is strongly correlated with human perception.

REFERENCES

[1] J. M. Foley, “Human luminance pattern-vision mechanisms: masking experiments require a new model,” *Journal of the Optical Society of America*, vol. 11, pp. 1710–1719, June 1994.

[2] A. B. Watson and J. Solomon, “Model of visual constrast gain control and pattern masking,” *Journal of the Optical Society of America*, vol. 14, pp. 2379–2391, Sept. 1997.

[3] J. Ross and H. D. Speed, “Spatial and temporal contrast sensitivity function of the visual system,” *Proceedings of the Royal Society of London*, vol. 246, pp. 61–70, 1991.

[4] M. Nadenau and J. Reichel, “Opponent color, human vision and wavelets for image compression,” in *Proc. of the Seventh Color Imaging Conference, IS&T*, (Scottsdale, Arizona), pp. 237–242, Nov. 1999.

[5] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, “Image coding using wavelet tranform,” *IEEE Transactions on Image Processing*, vol. 1, pp. 205–220, 1992.

[6] M. Nadenau and J. Reichel, “Image compression related contrast masking measurements,” in *SPIE, Human Vision and Electronic Imaging*, (San-Jose, CA), pp. 188–199, Jan. 2000.

[7] A. Mayache and T. Eude, “An evaluation of quality metrics for compressed images based on human visual sensitivity,” in *Fourth International Conference on Signal Processing (ICSP’98)*, (Beijing, China), pp. 779–782, Oct. 1998.

[8] Y.-K. Lai and C.-C. J. Kuo, “A haar wavelet approach to compressed image quaity assessment,” *Journal of Visual Communication and Image Representation*, vol. 11, pp. 17–40, Mar. 2000.

[9] Laboratory for Image & Video Engineering, University of Texas (Austin), “LIVE Image Quality Assessment Database,” *http://live.ece.utexas.edu/research/Quality*, 2002.

[10] ITU-R Recommendation BT.500-10, “Méthodologie d’évaluation subjective de la qualité des images de télévision,” tech. rep., ITU, Geneva, Switzerland, 2000.