IMAGE QUALITY CRITERION BASED ON THE CANCELLATION OF THE MASKED NOISE

S. Comes , O. Bruyndonckx , and B. Macq

Laboratoire de Télécommunications et Télédétection
Université Catholique de Louvain
2 Place du Levant, B-1348 Louvain-la-Neuve, Belgium
E-mail : bruyndon@tele.ucl.ac.be

ABSTRACT

This paper investigates the development of a new image quality criterion based on the psychovisual model of tuned channels and more particularly the phenomenon of masking. The masking model parameters have been evaluated by psychovisual tests assuming a logarithmic relationship between the visibility threshold and the contrast value of the background for a given perceptual channel. The masking has the consequence that only a part of the noise of a noisy image is really visible and affects the visual quality of the image. The idea of the proposed criterion is to evaluate image quality after the cancellation of the masked noise defined as the invisible noise. The criterion has been used in order to compare different coders and different post-processings of noisy images.

1. INTRODUCTION

When working with images affected by coding noise or post-processed by a restoration algorithm it is very useful to estimate numerically the level of quality of the images. Subjective quality experiments are so awkward and expensive that the development of objective criteria is necessary to fastly evaluate image quality. The most popular quality measure is the mean squared error (MSE) or the signal to noise ratio (SNR). Unfortunately this raw error measure only fits well for certain kind of impairments such as additive noise. In the other cases it is necessary to take human visual properties into account. Since many years several perceptual models have been proposed to estimate image quality [1, 2, 3].

S. Comes is Research Assistant at Telecommunications and Remote Sensing Laboratory, Université Catholique de Louvain, Belgium

O. Bruyndonckx is Research Assistant of the National Fund for Scientific Research, Belgium (FNRS) thanks to a grant of the Belgian Public Telecommunications Company (BELGACOM) at Telecommunications and Remote Sensing Laboratory, Université Catholique de Louvain, Belgium

B. Macq is Research Associate of the National Fund for Scientific Research, Belgium (FNRS) at Telecommunications and Remote Sensing Laboratory, Université Catholique de Louvain, Belgium

This paper investigates the development of a new image quality criterion based on the *psychovisual model of tuned channels* which has been used previously in some image quality criteria [4, 5, 6]. The image error is split up into perceptual components according to the position in the visual field, the spatial frequency and the orientation. These components are weighted by a masking function taking the background content into account. Our method is original for two reasons : (i) invisible or masked noise is removed from the computation of the image quality; (ii) the perceptual decomposition performs perfect reconstruction. So after weighting the remaining perceptual components are compounded in order to provide a perceptual error image from which the MSE is evaluated. Instead of a criterion suitable to any image our criterion rather provides a comparison tool for images quality taking more account of the perceptual aspects of the human visual system and so easier to validate than an absolute criterion. Note our quality criterion is restricted to still monochrome pictures.

In section 2 the psychovisual model used in the criterion is presented. More particularly the phenomenon of *masking* is developed. Section 3 describes the image quality criterion based on the cancellation of the masked noise. Some results of the application of the criterion to noisy images are in section 4.

2. THE PSYCHOVISUAL MODEL OF TUNED CHANNELS

According to the psychovisual model of tuned channels the visual information is transmitted from the retina to the cortex through *perceptual channels* which perform a decomposition with respect to the location in the visual field, the orientation and the spatial frequency of the stimuli. So the human visual system decomposes the image into different characteristics called *perceptual components.*

Each perceptual channel is characterized by its own visibility threshold. A stimulus tuned to a given channel is not perceived if its contrast value is below the visibility threshold corresponding to this channel. The visibility threshold for a perceptual channel depends on the content of the background image in the corresponding tuned channel. This phenomenon called *masking* is very important as part of

0-7803-2431-5/95 $4.00

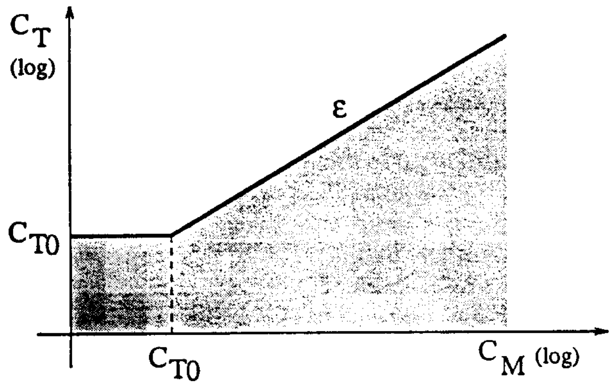

Figure 1: Masking model : logarithmic relation between the visibility threshold C_T and the masker contrast C_M

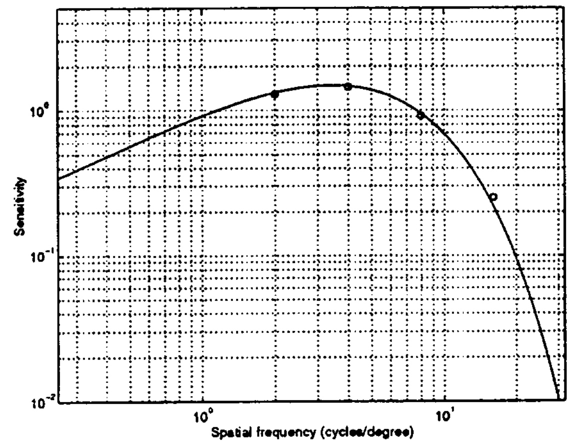

Figure 2: Sensitivity function derived from psychovisual experiments

image coding. Coders with loss generate coding noise which corrupts the original signal. The visibility of a noise pattern is not the same whether its background is uniform or textured.

In order to modelize the phenomenon of masking we assume [6] a logarithmic relationship between the visibility threshold C_T and the contrast value of the background C_M for a given perceptual channel, as illustrated in figure 1. Without any background in the channel ($C_M = 0$), the visibility is determined by the band-pass filtering effect of the human vision, especially the visibility threshold without masking C_{T_o}. Furthermore the detection threshold C_T of a stimulus increases as the masker contrast C_M increases following a slope ϵ.

Both parameters of the masking model have been evaluated by detection psychovisual tests following the assumptions that the slope ϵ is the same in each perceptual channel and the detection threshold without masking C_{T_o} only depends on the spatial frequency. Unlike classical masking tests using gratings, these stimuli were replaced by patterns of additive white gaussian noise (AWGN) passed through a given perceptual channel. Such kind of stimulus better fits coding noise characteristics. The detection test procedure was a version of a two alternatives forced-choice staircase method driven by a PEST procedure [7]. The experiments were carried out in our laboratory with 6 subjects placed at a viewing distance of 6 times the height of the TV screen.

About the parameter C_{T_o} the test-stimulus was a disk (diameter of 64 pixels) of filtered AWGN with uniform background. Experiments provided detection thresholds C_{T_o} for 4 spatial frequencies (2,4,8 and 16 c/d). The sensitivity curve (defined by the inverse of C_{T_o}) proposed by Mannos and Sakrison [1] was optimized to fit these data and is depicted in figure 2.

In order to compute the parameter ϵ the previous test-stimulus superimposed on a larger AWGN masker disk belonging to the same perceptual channel. Detection thresholds C_T were evaluated for 4 masker contrasts C_M ($C_M =$ 2.5,5,10 and 20). A linear regression provided a estimation of ϵ equal to 0.6.

3. IMAGE QUALITY CRITERION

In order to evaluate the quality of a noisy image the knowledge of the noise level is necessary. The masking has the consequence that only a part of the noise of a noisy image is really visible and affects the visual quality of the image. Criteria such as SNR consider the entire noise present in the image (sometimes with a weighting bound to the visibility level of the noise). The idea of the proposed criterion is to evaluate image quality *after the cancellation of the masked noise* defined as the invisible noise. The masked noise is the set of perceptual components of the noise whom contrast belongs to the dashed area of figure 1.

The global scheme of the procedure is shown in figure 3. The noise and the original picture are split up into several perceptual channels by means of a multiresolution filters bank performing a polar decomposition with perfect reconstruction [8]. More precisely the perceptual decomposition produces 21 perceptual pictures tuned to 5 frequencies (2, 4, 8, 16 c/d and more) and 4 orientations (0, 45, 90, 135°), plus the low-pass component. A *perceptual component* is defined as each pixel of the 21 perceptual pictures. Each perceptual component of the noise is compared to both the contrast of the corresponding perceptual component in the original image and the detection threshold without masking. The detection threshold without masking is considered as constant in a given perceptual channel and computed as the mean value of the function C_{T_o} (the inverse of the sensitivity function shown in figure 2) in the bandwidth of the channel. For high-frequency bands which are unlimited the value of C_{t_o} for $f = 32$ c/d is taken. The perceptual component of the noise for a given channel is removed in case of its contrast belongs to the dashed area of figure 1.

Image	$PSNR$ (dB)	$MPSNR$ (dB)	Q
Fruit JPEG	28.57	34.656	0.951
Fruit ASD	28.62	40.997	2.513
Fruit AFB	29.62	37.379	1.526
Lena JPEG	27.31	34.326	0.893
Lena PP1	29.92	44.269	3.411
Lena PP2	27.85	38.966	1.938

Table 1: Comparison of $PSNR$, $MPSNR$ and Q between several kind of noisy pictures shown in figures 4 and 5

This means our criterion is based on the *cancellation of the masked noise*. The unmasked noise is then reconstructed with the remaining perceptual channels [8].

The MSE P_B of the remaining noise is given by:

$$P_B = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} B^2(m,n) \quad (1)$$

where $B(m,n)$ represents the value of the unmasked noise at the location (m,n) of the $M \times N$ picture.

Using P_B to estimate the quality two approaches have been followed: the first one consists to measure the *Peak-SNR (PSNR)* considering P_B as the power of the noise. We define then the *Masked-PSNR (MPSNR)* as:

$$MPSNR = \frac{(255)^2}{P_B} \quad (2)$$

The second approach gives a measure on a five-grade quality scale going from 1 (very annoying noise) to 5 (imperceptible noise). The *quality factor* Q is then evaluated by:

$$Q = \frac{5}{1 + KP_B} \quad (3)$$

where K is a normalization factor to give Q between 1 and 5.

4. APPLICATION OF THE CRITERION

The criterion has been applied to different kind of noisy images. More precisely, the criterion has been used in order to compare different coders (JPEG, Adaptive Subband Decomposition (ASD) [9] and Asymmetric Filter Bank (AFB) [10]) and different post-processings of noisy images (Post-Processing by cancellation of the unmasked noise [8] with the noise assumed as known (PP1) and with the estimated noise (PP2)), as illustrated in table 1 and figures 4 and 5. Unlike the $PSNR$ both the $MPSNR$ and Q enable to classify the images according to their visual quality.

5. FURTHER WORKS

The quality criterion described in this paper is only a first attempt in order to reach a universal visual quality criterion. Some aspects of the scheme should be improved or optimized : the cancellation of the masked noise could be smoothed by a masked weighting, quality measure could be restricted to worst areas, ... These suggestions will have to be validated by global quality subjective tests.

6. CONCLUSIONS

In this paper the development of a new image quality criterion has been presented. The method evaluates the image quality after the cancellation of the masked noise. This scheme required the use of a perceptual decomposition transform and the evaluation of the masking model parameters. Once again let us emphasize that this criterion is more adequate when used to compare two noisy versions of an image than as an absolute criterion.

References

[1] J.L. Mannos and D.J. Sakrison. The effects of a visual fidelity criterion on the encoding of images. *IEEE Transactions on Information Theory*, 20(4):525–536, July 1974.

[2] F.X.J. Lukas and Z.L. Budrikis. Picture quality prediction based on a visual model. *IEEE Transactions on Communications*, 30(7):1679–1692, July 1982.

[3] M. Miyahara. Quality assessments for visual service. *IEEE Communications Magazine*, 26(10):51–60, October 1988.

[4] D.J. Sakrison. On the role of the observer and a distorsion measure in image transmission. *IEEE Transactions on Communications*, 28(11):1251–1267, November 1977.

[5] C. Zetzsche and G. Hauske. Multiple channel model for the prediction of subjective image quality. In *Human Vision, Visual Processing, and Digital Display*, volume 1077, pages 209–216, Los Angeles (California), 18-20 January 1989. SPIE.

[6] S. Comes and B. Macq. Human visual quality criterion. In *Visual Communications and Image Processing*, pages 1–12, Lausanne, October 1990. SPIE, M. Kunt.

[7] J.L. Hall. Hybrid adaptative procedure for estimation of psychometric functions. *Journal of Optical Society of America*, 69(6):1763–1769, June 1981.

[8] B. Macq, M. Mattavelli, O. Van Calster, E. van der Plancke, S. Comes, and W. Li. Image visual quality restoration by cancellation of the unmasked noise. In *Int. Conf. on Acoustics, Speech and Signal Processing*, pages V53–V56, Adelaïde, April 1994. IEEE.

[9] O. Egger, W. Li, and M. Kunt. Very low bit rate image compression using an adaptative morphological subband decomposition. *to be published in Proccedings of the IEEE, Special Issue on Advances in Image and Video Compression*, Early 1995.

[10] O. Egger and W Li. Subband coding of images using asymmetrical filter banks. *to be published in IEEE Trans. on Image Processing*, April 1995.

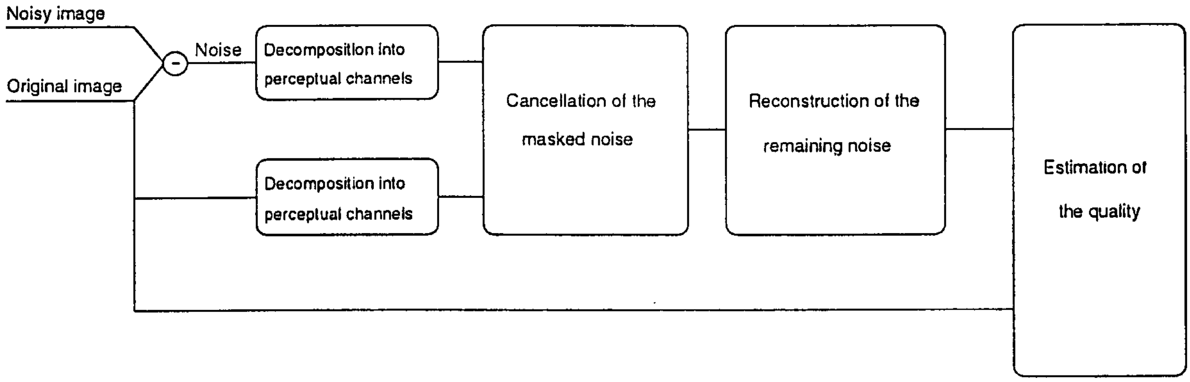

Figure 3: Global scheme of the quality estimation

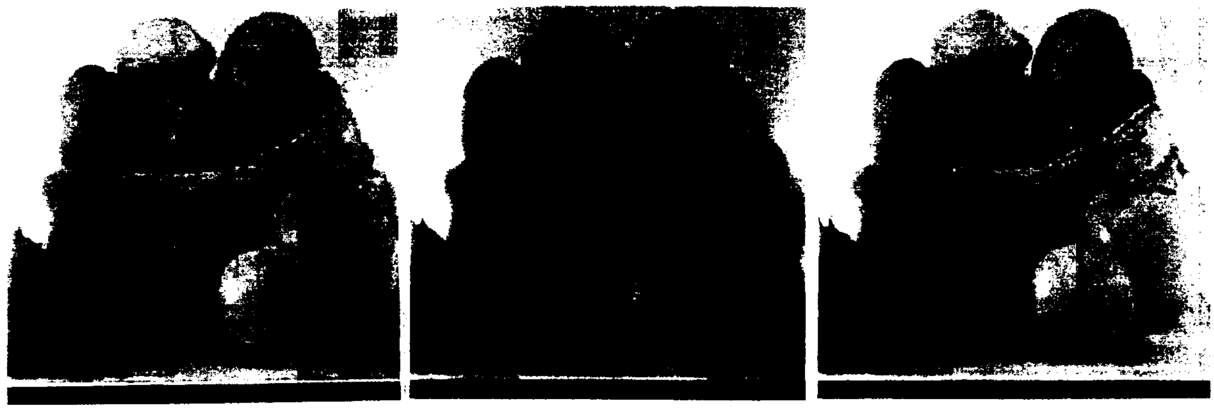

Figure 4: Comparison of different coders : (left) JPEG, (middle) ASD, (right) AFB.

Figure 5: Comparison of different post-processing schemes : (left) JPEG, (middle) PP1, (right) PP2.