14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

NO-REFERENCE PERCEPTUAL QUALITY ASSESSMENT OF COLOUR IMAGE

B. Bringier, N. Richard, M.-C. Larabi and C. Fernandez-Maloigne

SIC Laboratory, CNRS FRE 2731, University of Poitiers
Bat. SP2MI, Bvd Marie et PierreCurie, PB 30179
86962 Futuroscope cedex, France
phone: + 33 5 49 49 74 84, fax: + 33 5 49 49 65 70, email: {name}@sic.sp2mi.univ-poitiers.fr
web: www.sic.sp2mi.univ-poitiers.fr

ABSTRACT

Image quality assessment plays an important role in various image processing applications. In recent years, some objective image quality metrics correlated with perceived quality measurement have been developed. Two categories of metrics can be distinguished: with full-reference and no-reference. Full-reference looks at decrease in image quality from some reference of ideal. No-reference approach attempts to model the judgment of image quality without the reference. Unfortunately, the universal image quality model is not on the horizon and empirical models establishes on psychophysical experimentation are generally used. In this paper, we present a new algorithm for quality assessment of colour reproduction based on human visual system modeling. A local contrast definition is used to assign quality scores. Finally, a good correlation is obtained between human evaluations and our method.

1. INTRODUCTION

Image quality measurement is crucial for many image processing applications. So, the best way to assess the quality of an image is perhaps to look at it because human eyes are the ultimate receivers in most processing environments. However, this method is time consuming and expensive for practical usage. The most employed solution to measure the quality or the fidelity of an image is thus the use of parametric models. Great efforts have been done these last years to develop objective image quality metrics that correlate with perceived quality measurement. Unfortunately, only limited successes have been achieved. Indeed, in major cases, these metrics are limited to measure differences between image before and after processing. According to applications, the human visual system is incorporated in these models. In practice, such full-reference methods may not be applicable because the reference image is often not available. It is then necessary to measure directly on the processed image. Unfortunately, no-reference image quality is an extremely difficult task because many unquantifiable factors play a role in human assessment of quality, such as aesthetics, cognitive relevance, learning, context, etc. Consequently, most proposed no-reference quality metrics are designed for one or a set of predefined specific distortion types and are not generalized for evaluating images having other types of distortions. In image compression for example, no-reference quality metrics measure artifacts such as blockiness, blurness, ringing, etc. [9, 13]. The field of blind quality assessment has been thus largely unexplored.

In this paper, we present a new no-reference algorithm for quality assessment of colour reproduction based on human visual system modeling. For this application, the contrast is generally defined to be the most important quality parameter. So, image contrast is commonly defined in terms of tone reproduction curve. Unfortunately, two sets of images having very different white and black points may have very different perceptual contrasts. Image quality can not be established from the tone reproduction curve. Consequently, some empirical models based on psychophysical experimentation have been developed to compute the quality perceived regarding the contrast in an image. The more succeeded model uses a simple definition of Lightness-Contrast, Chroma-Contrast and Sharpness-Contrast [4, 3] in the Lab colour space. However, the parameter weights in this type of models depend on the set of images used in the human quality assessment. To solve this problem, we propose a new no-reference algorithm based on a modelisation of the human visual system. Initially, we compute the perceived information on a soft or a hard reproduction image. Then, a local contrast definition is used to assign quality scores. Finally, we validate our perceptual metric thanks to subjective experiments and analyze the correlation between the metric predictions and the observer ratings.

2. PERCEIVED INFORMATION ON A SOFT REPRODUCTION IMAGE

2.1 Colour reproduction

A lot of new display technologies appears in numerous applications. And, for example, LCD or plasma screen do not have the same restitution than conventional CRT screens. Some characteristics, like tone-reproduction curve or spectral and basic colorimetric characteristics can change the colour reproduction and, by the way, the perception of an image made by human. To model the luminous field emitted by the screen to estimate the perceived information by human visual system, traditional methods described in a number of scientific papers [1] can be use. In this work, we chose the S-Curve characterization [8] because it allows to approximate as well CRT than LCD curve types. In this model, the relationship between digital input values RGB (*In*) and luminous field (*Out*) emitted by the screen is defined as:

$$Out = A\frac{In^{\alpha}}{In^{\beta}+C} \tag{1}$$

2.2 Perceived information

The second step of our model is to calculate the perceived information form a displayed image. To integrate the notions of viewing distance and frequencies in images, we perform a filtering in Fourier domain by using colour contrast sensitivity functions, such as those shown in Figure 1.

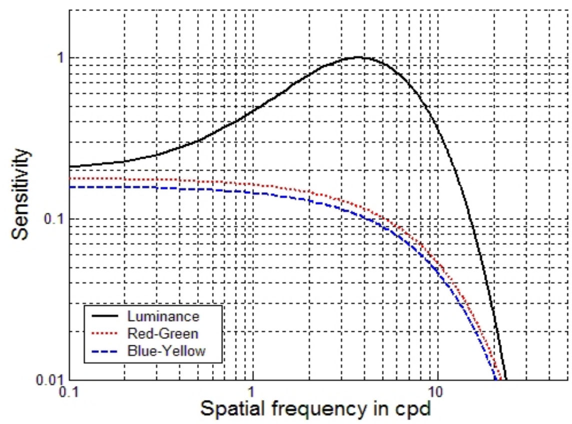


Figure 1: Typical shape of contrast sensitivity functions (CSF) measured with a threshold detection experiment.

The Contrast Sensitivity Function (CSF) describes the pattern sensitivity of the Human Visual System (HSV) as a function of contrast and spatial frequencies. Psychophysical experiments have Shown the high sensitivity of the HSV for low frequencies and low sensitivity for high frequencies [7]. The CSF is probably the most important stage in any HVS model. We can distinguish achromatic and chromatic sensitivity functions, representable by filters. These filters can be defined as [10]:

$$S_L(f) = a_1 f^2 \exp(b_1 f^{c_1}) + a_2 \exp(b_2 f^{c_2}) \qquad (2)$$

$$S_C(f) = a_3 \exp(b_3 f^{c_3}) \qquad (3)$$

where S_L indicates the sensitivity for the luminance channel and S_C the corresponding sensitivity for the chrominance channels. The parameters $a_1 \dots a_3$, $b_1 \dots b_3$ and $c_1 \dots c_3$ are subject to the parameter fit, f is the spatial frequency in cycle per degree.

This model assumes, for the achromatic CSF, a sum of two exponential functions to create a curve with band-pass characteristics and a single exponential function for the low-pass shape of the chromatic CSF. This filtering is performed in an opponent colour space, containing one luminance and two chrominance channels. The use of an opponent colour representation, based on psycho-physiological models, seem most appropriate in the context of image quality [12]. The opponent channels, AC1C2 are a linear transform from LMS.

3. QUALITY ASSESSMENT

The human perception is able to naturally define quality standards to classify a set of images. It is thus natural to use the human visual system to develop a quality model for colour reproduction. The new method suggested to predict quality uses contrast definition of Peli [11]. Figure 2 summarized the approach to compute Local Band-limited Contrast (LBC).

Indeed, the response of the human visual system depends much less on the absolute luminance than on the relation of its local variations to the surrounding luminance. This property is know as Weber's law.

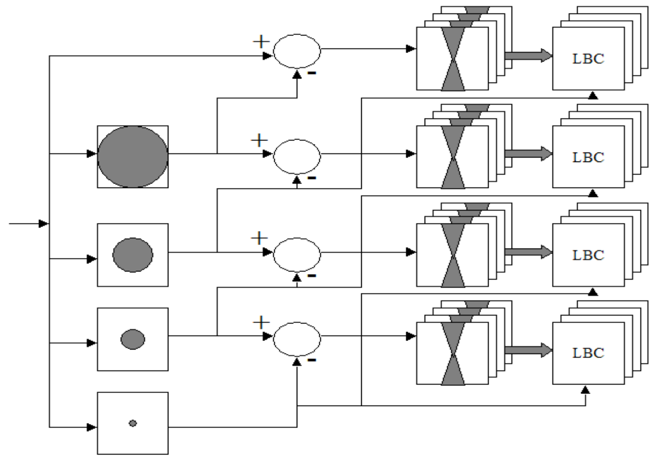


Figure 2: Local Band-limited Contrast method.

3.1 Frequency decomposition

To compute contrast, the image is filtered by a set of band-pass filters and fan filters like cortex transform [14]. Four spatial frequency bands and four orientations compose the frequency decomposition. With this filtering, the radial frequency selectivity and the orientational selectivity are modeled (cf. figure 3). The effects of these filters are cascaded to describe the combined radial and orientational selectivity of cortical neurons [5].

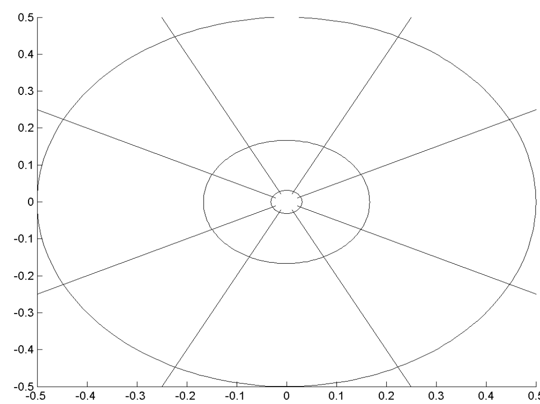


Figure 3: Frequency decomposition used in our model [5].

In the spatial domain, the filtered image $B_{k,l}$ can be represented as:

$$B_{k,l}(x,y) = L(x,y) * f_{k,l}(x,y) \qquad (4)$$

where $L(x,y)$ is the luminance distribution of the image and $f_{k,l}(x,y)$ is the filter with k frequency selectivity and l fan selectivity.

3.2 Local Band-limited Contrast calculation

Now, we have to convert the frequency bands $B_{k,l}$ into some measure of contrast perceived in image. Generally, Weber or Michelson contrast are used to compute simple stimuli contrast. Unfortunately, it is also obvious that none of these simple global definitions is appropriate for measuring contrast in natural images, because a few very bright or very dark points would determine the contrast of the whole image. To solve this problem, Peli proposed a local band-limited contrast measure:

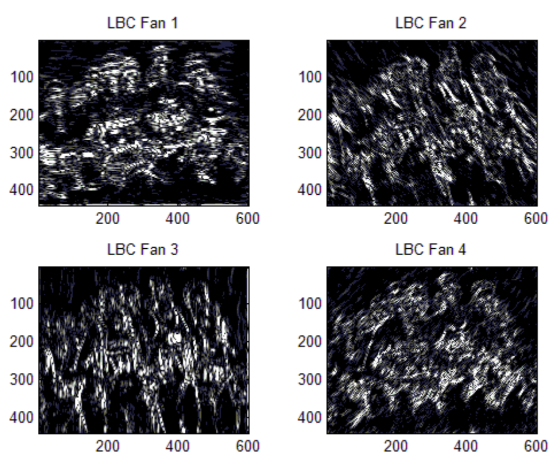
$$LBC_k(x,y) = \frac{B_k(x,y)}{\sum_{i=0}^{k-1} B_i} \tag{5}$$

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 [11]:

$$LBC_{k,l}(x,y) = \begin{cases} \frac{B_{k,l}(x,y)}{M_{k,l}+\sum_{i=0}^{k-1} B_{i,l}} & \forall \quad k = 2..K, l = 1..L \\ \frac{B_{k,l}(x,y)}{M_{k,l}+B_0} & \forall \quad k = 1, l = 1..L \end{cases} \tag{6}$$

where B_0 is the average of the image defined by the center of Daly frequency decomposition and $M_{k,l}$ is function to the average of the image and can be used to model the frequency and orientation sensitivity of the HVS. Figure 4 show an example of LBC calculation for average frequency in all orientation (corresponding to the second frequency bands in Daly decomposition).

(a) Image *Athlete*



(b) LBC of image *Athletes*

Figure 4: Local band-limited contrast image for average frequency and four orientations.

This contrast informations are now combined to provide a contrast assessment. Coefficients are used according to frequency and orientation decomposition. A parameter weight of value 2 is affected to achromatic channel compared to chromatic channel. The same parameter weight is affected to high frequency informations compared to average frequency informations. Thus, luminance contrast is considered more important than colour contrast. Indeed, in this study, contrast changes in images according to the tone reproduction curve appear primarily in luminance. Moreover, the sharpness depends primarily on the high frequencies, it is thus obvious to assign a high weight. It is also in coherence with neurophysiological functions.

4. EXPERIMENTS AND RESULTS

To confirm the perceptual relevance of our metric, we carried out a set of subjective experiments. We ask twenty observers to evaluate the quality perceived in a set of twelve test images representing the typical image used in multimedia application (Corel Photo).

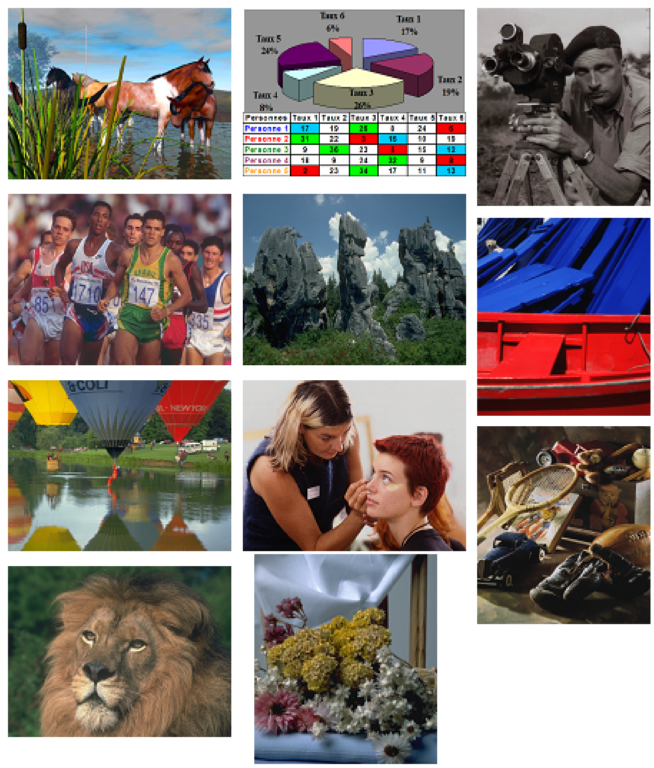


Figure 5: Test images for the quality experiment (from left to right and from top to bottom: *Synthesis*, *Graph*, *Camera*, *Athlete*, *Landscape*, *Boat*, *Transport*, *Model*, *Toys*, *Lion head* and *Fruit*.

With these images, our database was created by simulating nine tone reproduction curves that can be typically obtained in CRT and LCD screens, shown in figure 2. Thus we obtain a total of 126 test images. The study was conducted in one session. We screened the subjective ratings for outliers according to ITU-R Rec. BT.500.10 [2]. Each observer was shown the nine tone reproduction curve simulation on the same screen for all image. Observers were asked to provide their perception of quality to grade from the best to the worst the nine reproduction curves. For our analysis, the grade was then converted into discrete scales 1-9 linearly (1 for the worst and 9 for the best). We computed

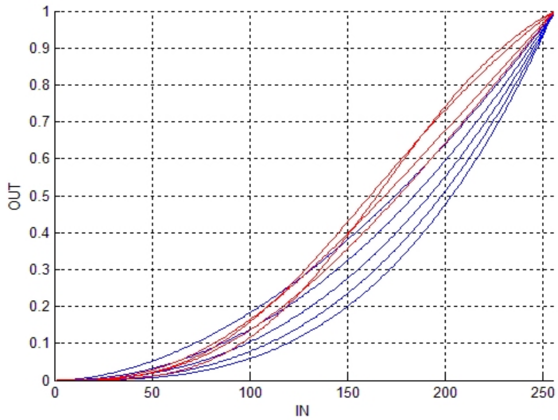


Figure 6: Tone reproduction curve use in subjective experiment.

the mean opinion scores (MOS) and the corresponding 95% confidence intervals.

4.1 Contrast prediction versus MOS

The model of perceived contrast prediction is applied to the whole images used during the subjective test. Figure 8 shows a part of the most significant results. We can see two types of curve. For *Transport* test image (figure 8(a)), we obtain a linear correlation between MOS and perceived contrast prediction. In this case, we can conclude that perceived contrast prediction can define directly the quality prediction. Thus, on the whole images which have the same results (seven images of database), we obtain a correlation between perceived quality and subjective results as high as 90%.

For other test images of the database (*Model*, *Synthesis*, *Boat* and *Toys*), we obtain slightly different results. Initially, the correlation between perceived contrast and subjective judgment is positive. Then, after a threshold, we obtain a negative correlation (figure 8(b)). However, a too important contrast in an image can decrease its visual quality as shows Janssen [6]. In his work, a psychophysical experimentation is used to determine the quality and the naturalness of an image according to contrast.

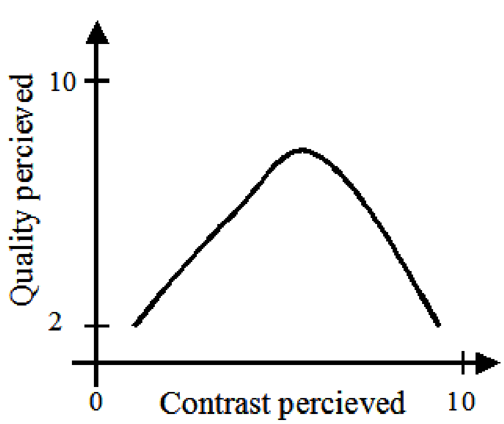
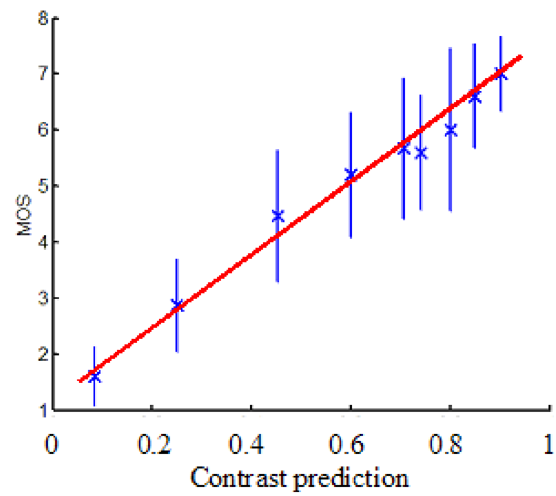
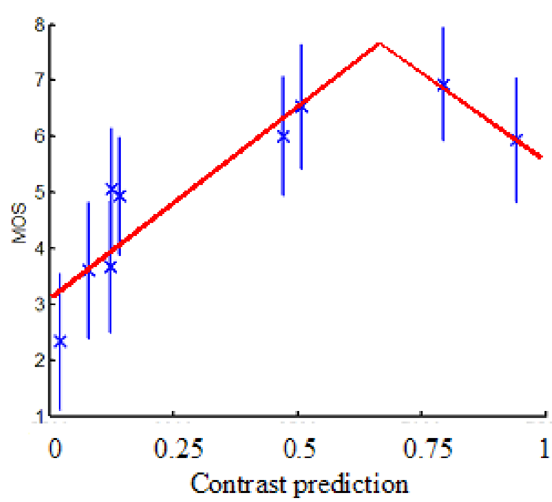


Figure 7: Quality judgments versus contrast judgments [6].

To this end, contrast of black and white image set (Kodak Photo CD) is modified by applying different values for Gamma tone reproduction curves. Three stages compose this study. In the first phase, the observers judge the contrast perceived in the image. Then, the observers must consider the naturalness and quality of the image. Finally, Janssen shows the relation between the perceived contrast and the naturalness or quality of an image. Figure 7 shows his results. As for our study, he obtains a threshold between perceived contrast and quality judgment. And around this value, the quality judgment of the observers decreases according to the contrast perceived. Consequently, the human visual system is able to identify a very high contrast in a global image. So, he do not allow a good quality. These observations are now used to obtain a quality assessment.



(a) Image *Transport*



(b) Image *Synthesis*

Figure 8: Error-bar plot with 95% confidence intervals of subjective ratings versus no-reference perceptual contrast measurement for two images of the database.

4.2 Quality prediction versus MOS

A threshold, function of contrast, is now applied to obtain a quality measurement. Figure 9 shows the results of quality prediction for the image *Synthesis*. Same 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%.

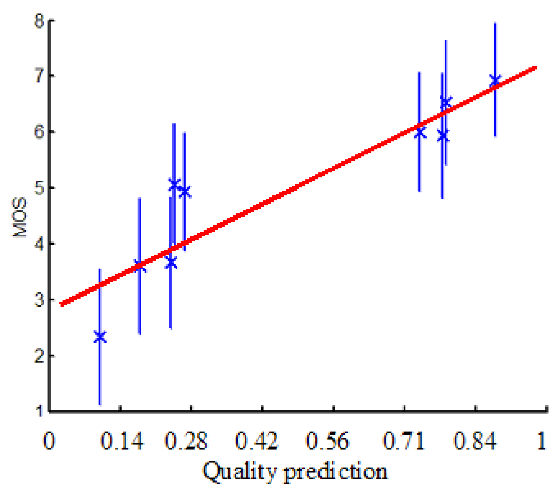


Figure 9: Error-bar plot with 95% confidence intervals of subjective ratings versus no-reference perceptual quality measurement for images *Synthese*.

The prediction which uses the human visual mechanisms is thus closer than a completely empirical method. Indeed, if we use the method exposed in [4], we obtain a correlation quality prediction and subjective experiments approximately equal to 83%. Moreover, the results are not homogenous from one image to another. With the new proposed method in this paper, the obtained results are similar for all images.

5. CONCLUSIONS

In this article, we presented a new No-Reference image quality assessment method based on a multiple channel HVS model for contrast definition. This type of quality prediction model is rare in the literature. Indeed, a full-reference model is generally used to evaluate the performance of image processing systems. Moreover, contrary to the models which evaluate a precise artifact like blockiness in compression, we use the human visual properties to develop our no-reference model. Consequently, we obtain a generic measurement which enables us to evaluate the quality of an image which does not have deformations like blockiness.

Finally, this model is validated with a subjective test based on the contrast change in an image. A correlation, between the results of such test and our quality prediction, higher than 90% is obtained. This result proves the efficiency of our model. Moreover, some model's coefficients can be refined to improve the prediction performance.

These results are encouraging and show a new way of work for image quality measurement without image reference. The use of human visual system modeling solves the dependency problem of the model parameters from a learning database and allows a generic formulation.

REFERENCES

[1] R.S. Berns, R.J. Motta, and M.E. Gorzynski. CRT colorimetry, part I : theory and practice. *Color research and application*, 18(5):299–314, Octobre 1993.

[2] ITU-R Recommendation BT.500-10. Methodology and photography. Technical report, ITU, Geneva, Switzerland, 2000.

[3] A.J. Calabria and M.D. Fairchild. Perceived image contrast and observer preference I: The effects of lightness, chroma, and sharpness manipulations on contrast perception. *Journal of imaging Science & Technology*, 47:479–493, 2003.

[4] 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.

[5] S. Daly. The visible differences predictor: An algorithm for the assessment of image fidelity. *A.B. Watson (ed.), MIT Press*, Digital images and human vision:179–206, 1993.

[6] T.J.W.N. Janssen. *Computational Image Quality*. PhD thesis, IPO, Center for User-System Interaction, Eindhoven, The Netherlands, 1999.

[7] D.H. Kelly. Spatiotemporal variation of chromatic and achromatic contrast thresholds. *Journal of optical society of america*, 73(6):742–750, Juin 1983.

[8] Y. Kwak, C. Li, and L. MacDonald. Controling color of liquid-cristal displays. *Journal of the SID*, 11(2):341–348, 2003.

[9] 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.

[10] M. Nadenau. *Integration of human color vision models into high quality image compression*. PhD thesis, Ecole polytechnique fédérale de Laussane, 2000.

[11] E. Peli. Contrast in complex images. *Journal of optical society of america*, 7(10), Octobre 1990.

[12] A. B. Poirson and B.A. Wandell. Pattern-color separable pathways predict sensitivity to simple colored patterns. *Vision research*, Mars 1996.

[13] H.R. Sheikh, Z. Wang, L. Cormack, and A.C. Bovik. Blind quality assessment for JPEG2000 compressed images. *IEEE Asilomar Conference on Signals Systems and Computers*, Pacific Grove, California, Novembre 2002.

[14] A.B. Watson. The cortex transform : Rapid computation of simulated neural images. *Computer vision, graphics, and image processing*, 39(3):311–327, 1987.