Compressed Image Colour Quality Assessment: Comparison between JPEG and JPEG2000

C. Fernández-Maloigne, M.C. Larabi
Laboratoire SIC , CNRS FRE 2731
Bat SP2MI, bvd Marie et Pierre Curie, BP 30179,
86962 Futuroscope Chasseneuil cedex (France)
{fernandez,larabi}@sic.univ-poitiers.fr

ABSTRACT

The aim of this work is to show the importance of psychophysical experiments on the one hand and to compare them with the results of some frequently-used mathematical criteria such as PSNR and the Lab colour distances on the other hand. This will be illustrated will by the evaluation of different colour compression techniques: the usual JPEG norm, and 2 different implementations of the new norm JPEG2000 with and without visual weighting.

1. INTRODUCTION

In recent years there has been a growing number of digital image and video application, which all imply displaying, storing and transmitting large volumes of image data. Obviously, image data compression is a must and various lossless or lossy compression have been developed, notably the JPEG-family standards. The evaluation of the compression quality is a simple and straightforward task in the case of lossless compression: standard criteria (compression ratio, execution time, etc.) can be used. In the case of lossy compression the main difficulty arises in describing the type and the amount of degradation induced in the reconstructed image. The need of evaluating image quality in a human-assessment-consistent way has led to several approaches in image quality evaluation. There are two major types of image quality criteria: objective criteria and subjective (human judgement-based) ones. Since subjective image quality measure exhibit some inherent drawbacks (the use of a normalized evaluation room, a large panel of human observers, etc.), there has been a great deal of interest in developing quantitative measures, either in numerical or graphical form, that can be consistently used as a substitute [1-3]. This paper is organized as follows: Section 2 will present a short overview of the JPEG and JPEG2000 image compression standards used for this paper. Section 3 will present some quantitative metrics for objective colour quality assessments. Section 4 will deal with subjective quality assessments and will present ordering tests. Section 5 will present the experimental conditions and procedures used for this work and the evaluation with both objective and subjective methods. The paper ends with some conclusions to show the importance and the exactness of these experiences judgment on the one hand and to compare them with the results of some frequently-used mathematical criteria such as PSNR and the Lab colour distances on the other hand.

2. CODERS PRESENTATION

JPEG2000 is the most recent addition to the family of international standards developed by the group JPEG (Joint Photographic Expert Group).[4] It is seen as the successor of the JPEG standard. The scheme of the two compression standards JPEG and JPEG2000 are given by the figure 1.[5] This one includes a DC (Discret Cosinus) or DW (Discret Wavelet) Transform, a quantification bloc followed by the coding one.

So, we have made comparative studies between JPEG, using Discrete Cosinus Transform (DCT), and JPEG2000, using Discrete Wavelet Transform (DWT) without weighting and a version with fixed visual frequency weightings [6-7], by modification of embedded coding order[8].

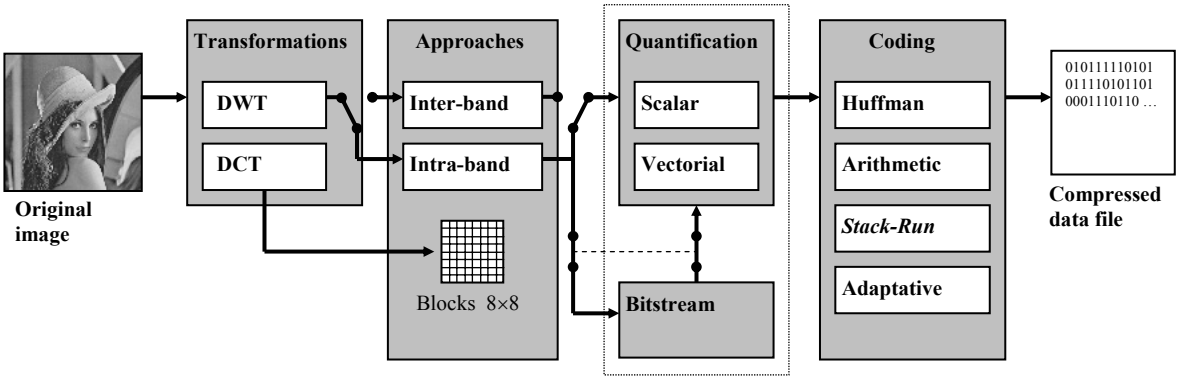


Figure 1: JPEG and JPEG2000 schemas

3. QUALITY ASSESSMENT METRICS

In order to measure the fidelity of the processed image (output image) with regard to the original one (input image), some simple mathematical metrics are used. These allow a first order measure of the damages introduced by the algorithm in the image. These criteria are described in the classic works of image processing[9-10]. The most frequently-used criterion is the PSNR (Peak Signal to Noise Ratio):

$$PSNR = 10\,Log_{10}\frac{\left(\text{max value of signal}\right)^2}{MSE} \tag{1}$$

Typically, each pixel of a grey-level image is coded in 8 bits (i.e. 255 levels). PSNR can also be estimated for colour images, by computing it on each component separately. This configuration does not take into account either the correlations between the different components, or neighbourhood of a pixel. So we can compute different colour distances such as Lab and Luv distances :

$$\Delta E^*_{ab} = \sqrt{\left(\Delta L^*\right)^2 + \left(\Delta a^*\right)^2 + \left(\Delta b^*\right)^2} \tag{2}$$

$$\Delta E^*_{uv} = \sqrt{\left(\Delta L^*\right)^2 + \left(\Delta u^*\right)^2 + \left(\Delta v^*\right)^2} \tag{3}$$

Recently, The CIE (Commission Internationale de l'Eclairage) has normalized a new colour distance: CIEDE2000.

$$\Delta E_{DE\,00} = \sqrt{\left(\frac{\Delta L^*_{ab}}{K_L S_L}\right)^2 + \left(\frac{\Delta C'_{ab}}{K_C S_C}\right)^2 + \left(\frac{\Delta H'_{ab}}{K_H S_H}\right)^2 + R_T\left(\frac{\Delta C'_{ab}}{K_C S_C}\right)\left(\frac{\Delta H'_{ab}}{K_H S_H}\right)} \tag{4}$$

$$a' = \left(1+G\right)a^* \quad \text{and} \quad G = 0.5\left(1-\sqrt{\frac{\overline{C^*_{ab}}^7}{\overline{C^*_{ab}}^7 + 25^7}}\right) \tag{5}$$

$$C'_{ab} = \sqrt{a'^2 + b^{*2}} \quad \text{with} \quad \Delta C'_{ab} = C'_{orig} - C'_{repro} \quad \text{and} \quad \Delta L^*_{ab} = L^*_{orig} - L^*_{repro} \tag{6}$$

$$\Delta H'_{ab} = 2\sin\left(\Delta h'/2\right)\sqrt{C'_{orig}\,C'_{repro}} \quad , \Delta h' = h'_{orig} - h'_{repro} \quad \text{and} \quad h' = \tan^{-1}\left(b^*/a'\right) \tag{7}$$

$$S_L = 1 + \frac{0.015\left(\overline{L^*_{ab}} - 50\right)^2}{\sqrt{20 + \left(\overline{L^*_{ab}} - 50\right)^2}}, \quad S_C = 1 + 0.045\,\overline{C'_{ab}}\,, \; S_H = 1 + 0.015\,\overline{C'_{ab}}\,T \tag{8}$$

K_L, K_C and K_H are parameters linked to observation conditions.

4. PSYCHOPHYSICAL EXPERIMENTS

In spite of the research efforts, the understanding of the complex interaction existing between the image models and the HVS models still insufficient to quantify the visual quality. Therefore psychophysical experiences stay the only viable means for the evaluation of the performances of

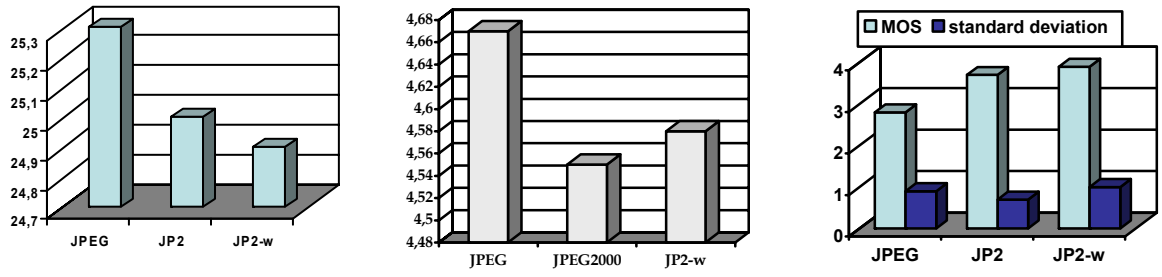
image processing tools such as compression, watermarking, etc. The psychophysical experiences are characterized by the participation of human observers to exploit the HVS capabilities to make a quality assessment of an objet.[11] It can involve experiences where the observer is called to choose between two images or to make a classification of a set of images. During the psychophysical experiences, it is necessary to make sure that the human observer has on one hand, a normal vision of colours and on the other hand, a correct visual acuteness. In order to verify these conditions, several tests as the famous Ishihara one are made. The other precautions should to be taken with regard to the used material and the environment in which the psychophysical experiences take place. This environment normalization[12] regroups several criteria as the distance between the screen and the observer, the reflection of the room surfaces and the gamma correction of monitors, as supplied by the ITU organization[13]. For our quality assessment experiments, the used protocol is the ordering test. The purpose of this test is to ask the observer to make a classification of images, from the worst (in qualitative terms) to the best, with regard to a reference image. Furthermore, the images are randomly placed around the original image. This avoids influencing the observer choice. In order to minimize the errors made by the observers, the same series of images are proposed to them twice but in a different disposal. At the end, we can compute the Mean Opinion Score:

MOS = $\overline{u}_{jkr} = \frac{1}{N}\sum_{i=1}^{N} u_{ijkr}$, u_{ijkr} is the score of observer *i* for degradation *j* of image *k* and iteration *r*, and *N* is the number of observers.

5. OUR EXPERIMENTS

In the framework of our experiments, we have focused on comparisons between JPEG and JPEG2000 compression. Both objective and subjective tests were performed on 12 images from the Kodak database (768x512 pixels, 24 bpp RGB colour images). Each image was compressed at various bit-rates, ranging from 4 bpp to 0.0062 bpp (4 bpp, 3 bpp, 2 bpp, 1 bpp, 0.5 bpp, 0.375 bpp, 0.25 bpp, 0.187 bpp, 0.125 bpp, 0.062 bpp) and decompressed according to the JPEG/JPEG 2000 standards. Three compression algorithms were used, named JPEG, JPEG2000 and JP2-w, with visual weighting. The subjective tests were performed with a panel of 30 observers, with different image processing backgrounds. The observers were asked to evaluate the photographic samples of the degraded images, as displayed on a 24"""" Sony high-resolution computer monitor, with a Trinitron tube, in a special room, which was built according to the international standards ITU-R500-10 and ISO3664[17]. We have used an ordering test where the original image is surrounded by 3 images representing each coder.

The figures 2-a, 2-b and 2-c show respectively, the PSNR, the CIEDE2000 and the MOS (Mean Opinion Score) associated to the standard deviation related to the MOS for a same bit-rate. We can remark that the PSNR gives the JPEG coder as the best approximately for all images. These results are contradicted by the MOS where the JPEG2000-w coder is in majority considered as the best by the observers. Regarding the colour distances, the different computed results obtained with the different Lab metrics have shown that, on the average, the CIEDE2000 is the most correlated with the psychophysical experiments, but this correlation is content image dependent. Therefore, we can see that the coder with weightings sometimes offers a greatest distance than the coder JPEG2000 without visual weightings. That could be explained by the fact that if JPEG2000 better preserve the original colour palettes than JPEG, the visual weightings, computed on frequency bands, can disturb the global colour distance, computed spatially, pixel by pixel.



-a- -b- -c-

Figure 2. Evaluation results : a- *average PSNR, b- average CIED2000, c- MOS and standard deviation*

So, in spite of the important visual difference between JPEG and JPEG2000, as we can see on figure 3 and between JPEG2000 without and with weightings, as seen on figure 4, the PSNR has given the first coder as the best for all the images, and the last as the worst. These figures allow us to confirm the inappropriateness of the PSNR as a metric for quality assessment.

Figure 3: JPG and JP2 compressed images, 34.9 Ko

Figure 4: from left to right : original image, same JP2 compression without and with compression.

4. CONCLUSIONS

This paper presented the quality assessment work done both for colour compression JPEG and JPEG2000. In terms of subjective quality, the JPEG2000 implementation using the visual frequency weightings offers generally better results. That is in total contradiction with the PSNR measure and that's not always confirmed by CIED2000 difference equations which is nevertheless the best colour distance regarding the human perception. Globally, the results show that the mathematical metrics are not enough consistent to evaluate quality of colour image processes. Our actual research directions concern the definition of a metric based on the Human Visual System, better correlated to psychophysical experiences[14].

Note: These works have been financed by the French RNRT program with Thalès Communication.

REFERENCES

1. S. Daly, """"The Visible Differences Predictor : an Algorithm for the Assessment of Image Fidelity""", in *Digital Images and Human Vision*, edited by A. B. Watson, MIT press, pp. 197-206, 1993.
2. M. Esckicioglu and P. Fisher, """"Image Quality Measures and their performances""", *IEEE Transactions on Communications.* Vol. COM-43, No. 12, pp. 2959-2965, Dec., 1995.
3. P. Carrai, I. Heynderickx, P. Gastaldo, R. Zunino, "Image quality assessment by using neural networks", in IEEE International Symposium on Circuits and Systems, ISCAS 2002, Volume 5, pp. 253-256, 2002.
4. D. S. Taubman and M. Marcellin,"JPEG2000: Image compression fundamentals, Standards and practice", Kluwer cademic Publishing, 2002.
5. M. Rabbani and R. Joshi, "An overview of the JPEG2000 still image compression standard," Signal Processing: Image Communication 17, pp. 3–48, january 2002.

6 .A. Skodras, C. Christopoulos, and T. Ebrahimi, The JPEG2000 still image compression standard, *IEEE Signal. Processing Mag.,* vol. 18, pp 36-58, Sept. 2001.

7. "Information Technology – JPEG 2000 Image Coding System," ISO/IEC FDIS15444-1: 2000 (Aug. 2000).
8. A . Stoica, "Apport de la perception couleur à l'estimation de la qualité des images numériques – Application à JPEG2000 », Thèse, Université de Poitiers, 6 décembre 2004.
9. .R. C. Gonzalez and P. Wintz, Digital Image Processing, Addison-Wesley Publisking Comp., 97.
10. M. Kunt, Traitement numérique des images, Presses Politech et Universitaires Romandes, 1993.
11. J. O. Limb, "Distorsion criteria of the human viewer," IEEE Trans on Systems, Man and Cybern.9(12), pp. 778–793, 79.
12. C. Charrier, Vers l'optimisation statistique et perceptuelle de la qualité pour la compression des images couleur par quantification vectorielle. Thèse, Université Jean Monnet Saint-Etienne, 1998.
13. ITU-R, "Methodology for the subjective assessment of the quality", ITU-R BT.500-10, 2000.
14. C. Fernandez-Maloigne, M.-C. Larabi, B. Bringier, N. Richard,"Spatial-temporal characteristics of the visual brain and their effects on colour quality evaluation",AIC05, Grenada, Spain 8-13 May 2005.