Image Quantizer based on Contrast Band-Pass Filtering

Jaime Moreno[1,2], Oswaldo Morales[1], and Ricardo Tejeida[1]

[1] National Polytechnic Institute of Mexico,
IPN Avenue, Lindavista, Mexico City, 07738, Mexico.
[2] XLIM Laboratory, Signal, Image and Communications Department,
University of Poitiers, 86962 Futuroscope, France.
jmorenoe@ipn.mx

Abstract The aim of this work is to explain how to apply perceptual criteria in order to define a perceptual forward and inverse quantizer. We present its application to the H*i*-SET coder. Our approach consists in quantizing wavelet transform coefficients using some of the human visual system behavior properties. Taking in to account that noise is fatal to image compression performance, because it can be both annoying for the observer and consumes excessive bandwidth when the imagery is transmitted. Perceptual quantization reduces unperceivable details and thus improve both visual impression and transmission properties. The comparison between JPEG2000 coder and the combination of H*i*-SET with the proposed perceptual quantizer (χSET) shows that the latter is not favorable in PSNR than the former, but the recovered image is more compressed (less bit-rate) at the same or even better visual quality measured with well-know image quality metrics, such as MSSIM, UQI or VIF, for instance.

Keywords: Human Visual System, Contrast Sensitivity Function, Perceived Images, Wavelet Transform, Peak Signal-to-Noise Ratio, No-Reference Image Quality Assessment, JPEG2000.

1 Introduction

Digital image compression has been a research topic for many years and a number of image compression standards has been created for different applications. The JPEG2000 is intended to provide rate-distortion and subjective image quality performance superior to existing standards, as well as to supply functionality [1]. However, JPEG2000 does not provide the most relevant characteristics of the human visual system, since for removing information in order to compress the image mainly information theory criteria are applied. This information removal introduces artifacts to the image that are visible at high compression rates, because of many pixels with high perceptual significance have been discarded.

Hence, it is necessary an advanced model that removes information according to perceptual criteria, preserving the pixels with high perceptual relevance regardless of the numerical information. The Chromatic Induction Wavelet Model presents some perceptual concepts that can be suitable for it. Both CBPF and JPEG2000 use wavelet transform. CBPF uses it in order to generate an approximation to how every pixel is perceived from a certain distance taking into account the value of its neighboring pixels. By contrast, JPEG2000 applies a perceptual criteria for all coefficients in a certain spatial frequency independently of the values of its surrounding ones. In other words, JPEG2000 performs a global transformation of wavelet coefficients, while CBPF performs a local one.

CBPF attenuates the details that the human visual system is not able to perceive, enhances those that are perceptually relevant and produces an approximation of the image that the brain visual cortex perceives. At long distances the lack of information does not produce the well-known compression artifacts, rather it is presented as a softened version, where the details with high perceptual value remain (for example, some edges).

2 JPEG2000 Global Visual Frequency Weighting

In JPEG2000, only one set of weights is chosen and applied to wavelet coefficients according to a particular viewing condition (100, 200 or 400 dpi's) with fixed visual weighting[1, Annex J.8]. This viewing condition may be truncated depending on the stages of embedding, in other words at low bit rates, the quality of the compressed image is poor and the detailed features of the image are not available since at a relatively large distance the low frequencies are perceptually more important. The table 1 specifies a set of weights which was designed for the luminance component based on the CSF value at the mid-frequency of each spatial frequency. The viewing distance is supposed to be 4000 pixels, corresponding to 10 inches for 400 dpi print or display. The weight for LL is not included in the table, because it is always 1. Levels 1, 2, . . . , 5 denote the spatial frequency levels in low to high frequency order with three spatial orientations, *horizontal, vertical* and *diagonal*.

Table 1. Recommended JPEG2000 frequency (s) weighting for 400 dpi's ($s = 1$ is the lowest frequency wavelet plane).

s	*horizontal*	*vertical*	*diagonal*
1	1	1	1
2	1	1	0.731 668
3	0.564 344	0.564 344	0.285 968
4	0.179 609	0.179 609	0.043 903
5	0.014 774	0.014 774	0.000 573

3 Perceptual Forward Quantization

3.1 Methodology

Quantization is the only cause that introduces distortion into a compression process. Since each transform sample at the perceptual image $\mathcal{I}_\rho$ is mapped independently to a corresponding step size either Δ_s or Δ_n, thus $\mathcal{I}_\rho$ is associated with a specific interval on the real line. Then, the perceptually quantized coefficients $\mathcal{Q}$, from a known viewing distance d, are calculated as follows:

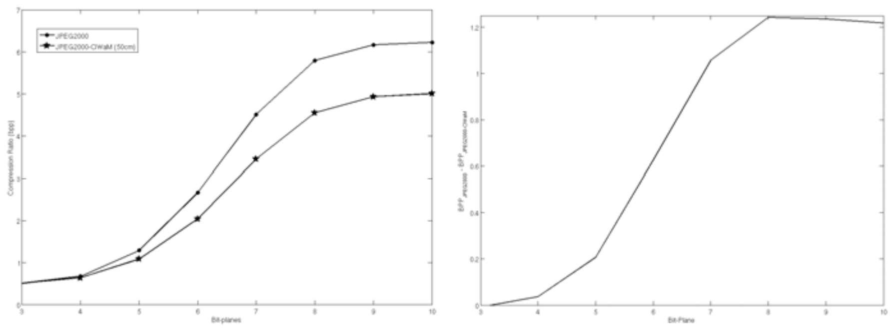
$$\mathcal{Q} = \sum_{s=1}^{n} \sum_{o=v,h,d} sign(\omega_{s,o}) \left\lfloor \frac{|\alpha(\nu, r) \cdot \omega_{s,o}|}{\Delta_s} \right\rfloor + \left\lfloor \frac{c_n}{\Delta_n} \right\rfloor \tag{1}$$

Unlike the classical techniques of Visual Frequency Weighting (VFW) on JPEG2000, which apply one CSF weight per sub-band [1, Annex J.8], Perceptual Quantization using CBPF (ρSQ) applies one CSF weight per coefficient over all wavelet planes $\omega_{s,o}$. In this section we only explain Forward Perceptual Quantization using CBPF (F-ρSQ). Thus, Equation 1 introduces the perceptual criteria of Perceptual Images to each quantized coefficient of Equation of Dead-zone Scalar Quantizer. A normalized quantization step size $\Delta = 1/128$ is used, namely the range between the minimal and maximal values at $\mathcal{I}_\rho$ is divided into 128 intervals. Finally, the perceptually quantized coefficients are entropy coded, before forming the output code stream or bitstream.

3.2 Experimental Results applied to JPEG2000

The Perceptual quantizer F-ρSQ in JPEG2000 is tested on all the color images of the *Miscellaneous volume* of the University of Southern California Image Data Base[2]. The data sets are eight 256×256 pixel images and eight 512×512 pixel images, but only visual results of the well-known images *Lena*, *F-16* and *Baboon* are depicted, which are 24-bit color images and 512×512 of resolution. The CBPF model is performed for a 19 inch monitor with 1280 pixels of horizontal resolution at 50 centimeters of viewing distance. The software used to obtain a JPEG2000 compression for the experiment is *JJ2000*[3]. Figure 1(a) shows the assessment results of the average performance of color image compression for each bit-plane using a Dead-zone Uniform Scalar Quantizer (SQ, function with heavy dots), and it also depicts the results obtained when applying F-ρSQ(function with heavy stars). Using CBPF as a method of forward quantization, achieves better compression ratios than SQ with the same threshold, obtaining better results at the highest bit-planes, since CBPF reduces unperceivable features. Figure 1(b) shows the contribution of F-ρSQ in the JPEG2000 compression ratio, for example, at the eighth bit-plane, CBPF reduces 1.2423 bits per pixel than the bit rate obtained by SQ, namely in a 512×512 pixel color image, CBPF estimates that 39.75KB of information is perceptually irrelevant at 50 centimeters.

Figure 2 depicts examples of recovered images compressed at 0.9 bits per pixel by means of JPEG2000 (a) without and (b) with F-ρSQ. Also these figures



(a) (b)

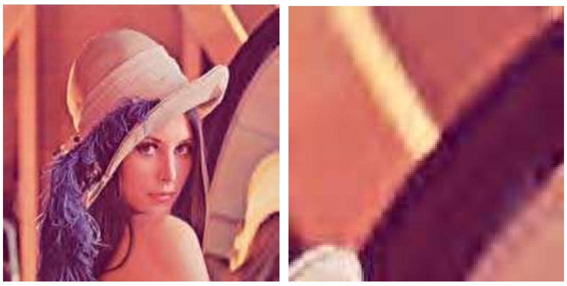
Figure 1. (a)JPEG2000 Compression ratio (bpp) as a function of Bit-plane. Function with heavy dots shows JPEG2000 only quantized by the dead-zone uniform scalar manner. While function with heavy stars shows JPEG2000 perceptually pre-quantized by F-ρSQ. (b)The bit-rate decrease by each Bit-plane after applying F-ρSQ on the JPEG2000 compression.

show that the perceptual quality of images forward quantized by ρSQ is better than the objective one. Also, figure 3 shows examples of recovered images of *Baboon* compressed at 0.59, and 0.45 bits per pixel by means of JPEG2000 (a) without and (b) with F-ρSQ. In Fig. 3(a) PSNR=26.18 dB and in Fig. 3(b) PSNR=26.15 dB but a perceptual metrics like WSNR [4], for example, assesses that it is equal to 34.08 dB. Therefore, the recovered image Forward quantized by ρSQ is perceptually better than the one only quantized by a SQ. Since the latter produces more compression artifacts, the ρSQ result at 0.45 bpp (Fig. 3(b)) contains less artifacts than SQ at 0.59 bpp. For example the *Baboon*'s eye is softer and better defined using F-ρSQ and it additionally saves 4.48 KB of information.

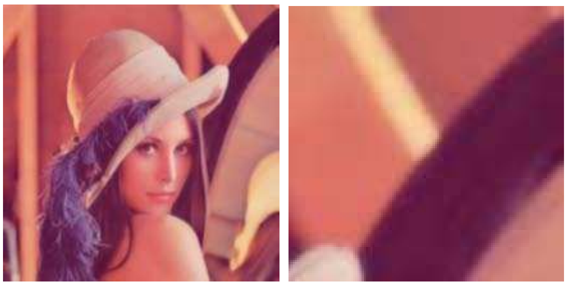
4 Perceptual Inverse Quantization

The proposed Perceptual Quantization is a generalized method, which can be applied to wavelet-transform-based image compression algorithms such as EZW, SPIHT, SPECK or JPEG2000. In this work, we introduce both forward (F-ρSQ) and inverse perceptual quantization (I-ρSQ) into the Hi-SET coder. This process is shown in the green blocks of Fig. 4. An advantage of introducing ρSQ is to maintain the embedded features not only of Hi-SET algorithm but also of any wavelet-based image coder. Thus, we call CBPF Perceptual Quantization + Hi-SET = cHi-SET or χSET.

Both JPEG2000 and χSET choose their VFWs according to a final viewing condition. When JPEG2000 modifies the quantization step size with a certain visual weight, it needs to explicitly specify the quantizer, which is not very suitable for embedded coding. While χSET neither needs to store the visual



(a) JPEG2000 PSNR=31.19 dB.



(b) JPEG2000-F-ρSQ PSNR=27.57 dB.

Figure 2. Examples of recovered images of Lenna compressed at 0.9 bpp.

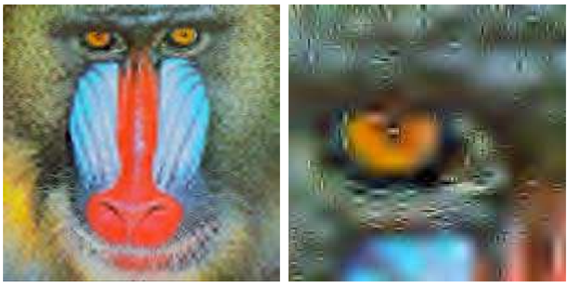
weights nor to necessarily specify a quantizer in order to keep its embedded coding properties.

The main challenge underlies in to recover not only a good approximation of coefficients $\mathcal{Q}$ but also the visual weight $\alpha(\nu, r)$(Eq. 1) that weighted them. A recovered approximation $\widehat{\mathcal{Q}}$ with a certain distortion Λ is decoded from the bitstream by the entropy decoding process. The VFWs were not encoded during the entropy encoding process, since it would increase the amount of stored data. A possible solution is to embed these weights $\alpha(\nu, r)$ into $\widehat{\mathcal{Q}}$. Thus, our goal is to recover the $\alpha(\nu, r)$ weights only using the information from the bitstream, namely from the Forward quantized coefficients $\widehat{\mathcal{Q}}$.

Therefore, our hypothesis is that an approximation $\widehat{\alpha}(\nu, r)$ of $\alpha(\nu, r)$ can be recovered applying CBPF to $\widehat{\mathcal{Q}}$, with the same viewing conditions used in $\mathcal{I}$. That is, $\widehat{\alpha}(\nu, r)$ is the recovered e-CSF. Thus, the perceptual inverse quantizer or the recovered $\widehat{\alpha}(\nu, r)$ introduces perceptual criteria to Inverse Scalar Quantizer and is given by:

$$\widehat{\mathcal{I}} = \begin{cases} \displaystyle\sum_{s=1}^{n} \sum_{o=v,h,d} sign(\widehat{\omega_{s,o}}) \, \frac{\Delta_s \cdot (|\widehat{\omega_{s,o}}| + \delta)}{\widehat{\alpha}(\nu, r)} + (\widehat{c_n} + \delta) \cdot \Delta_n & |\widehat{\omega_{s,o}}| > 0 \\ 0, & \widehat{\omega_{s,o}} = 0 \end{cases} \qquad (2)$$

For the sake of showing that the encoded VFWs are approximately equal to the decoded ones, that is $\alpha(\nu, r) \approx \widehat{\alpha}(\nu, r)$, we perform two experiments.



(a) JPEG2000 compressed at 0.59 bpp.

(b) JPEG2000-F-ρSQ compressed at 0.45 bpp.

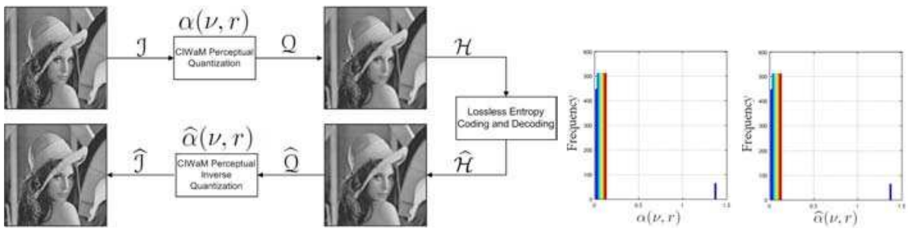
Figure 3. Examples of recovered images of Baboon.

Figure 4. The χSET image compression algorithm. Green blocks are the F-ρSQ and I-ρSQ procedures.

Experiment 1: Histogram of $\alpha(\nu, r)$ and $\widehat{\alpha}(\nu, r)$. The process of this short experiment is shown by Figure 5. Figure 5(a) depicts the process for obtaining losslessy both Encoded and Decoded visual weights for the 512×512 *Lena* image, channel Y at 10 meters. While Figures 5(b) and 5(c) shows the frequency histograms of $\alpha(\nu, r)$ and $\widehat{\alpha}(\nu, r)$, respectively. In both graphs, the horizontal axis represents the sort of VFW variations, whereas the vertical axis represents the number of repetitions in that particular VFW. The distribution in both histograms is similar and they have the same shape.

Experiment 2: Correlation analysis between $\alpha(\nu, r)$ and $\widehat{\alpha}(\nu, r)$. We employ the process shown in Fig. 5(a) for all the images of the CMU, CSIQ, and IVC Image Databases. In order to obtain $\widehat{\alpha}(\nu, r)$, we measure the lineal correlation between the original $\alpha(\nu, r)$ applied during the F-ρSQ process and the recovered $\widehat{\alpha}(\nu, r)$. Table 2 shows that there is a high similarity between the applied VFW and the recovered one, since their correlation is 0.9849, for gray-scale images, and 0.9840, for color images.

In this section, we only expose the results for the CMU image database.



(a) (b) (c)

Figure 5. (a) Graphical representation of a whole process of compression and decompression. Histograms of (b) $\alpha(\nu, r)$ and (c) $\widehat{\alpha}(\nu, r)$ visual frequency weights for the 512×512 image *Lenna*, channel Y at 10 meters.

Table 2. Correlation between $\alpha(\nu, r)$ and $\widehat{\alpha}(\nu, r)$ across CMU, CSIQ, and IVC Image Databases.

Image Database	8 bpp gray-scale	24 bpp color
CMU	0.9840	0.9857
CSIQ	0.9857	0.9851
IVC	0.9840	0.9840
Overall	**0.9849**	**0.9844**

Fig. 6 depicts the PSNR difference (dB) of each color image of the CMU database, that is, the gain in dB of image quality after applying $\widehat{\alpha}(\nu, r)$ at $d = 2000$ centimeters to the $\widehat{\mathcal{Q}}$ images. On average, this gain is about 15 dB. Visual examples of these results are shown by Fig. 7, where the right images are the original images, central images are perceptual quantized images after applying $\alpha(\nu, r)$ and left images are recovered images after applying $\widehat{\alpha}(\nu, r)$.

After applying $\widehat{\alpha}(\nu, r)$, a visual inspection of these sixteen recovered images show a perceptually lossless quality. We perform the same experiment experiment for gray-scale and color images with d = 20, 40, 60, 80, 100, 200, 400, 800, 1000 and 2000 centimeters, in addition to test their objective and subjective image quality by means of the PSNR and MSSIM metrics, respectively.

In Figs. 8 and 9, green functions denoted as F-ρSQ are the quality metrics of perceptual quantized images after applying $\alpha(\nu, r)$, while blue functions denoted as I-ρSQ are the quality metrics of recovered images after applying $\widehat{\alpha}(\nu, r)$. Thus, either for gray-scale or color images, both PSNR and MSSIM estimations of the quantized image $\mathcal{Q}$ decrease regarding d, the longer d the greater the image quality decline. When the image decoder recovers $\widehat{\mathcal{Q}}$ and it is perceptually inverse quantized, the quality barely varies and is close to perceptually lossless, no matter the distance.

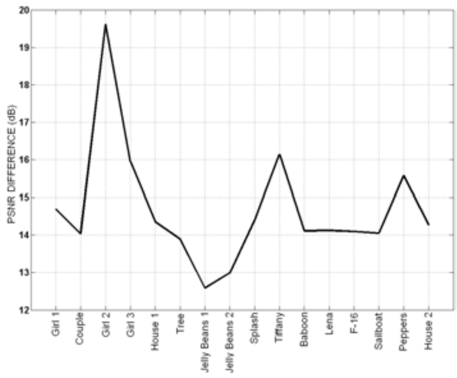


Figure 6. PSNR difference between $\widehat{\mathcal{Q}}$ image after applying $\alpha(\nu, r)$ and recovered $\widehat{\mathcal{I}}$ after applying $\widehat{\alpha}(\nu, r)$ for every color image of the CMU database.

5 Conclusions

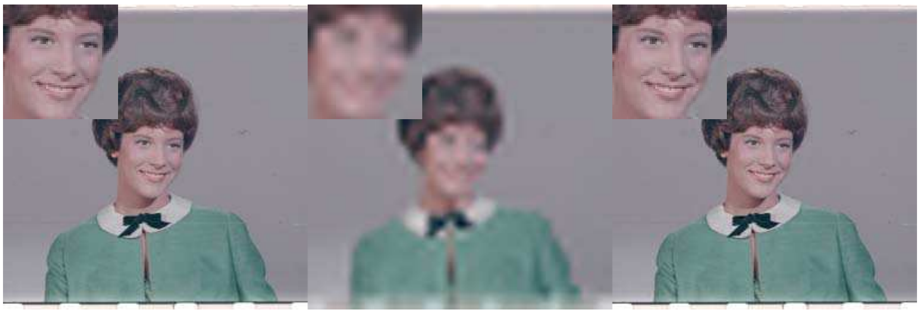
In this work, we defined both forward (F-ρSQ) and inverse (I-ρSQ) perceptual quantizer using CBPF. We incorporated it to Hi-SET, testing a perceptual image compression system χSET. In order to measure the effectiveness of the perceptual quantization, a performance analysis is done using thirteen assessments such as PSNR, MSSIM, VIF, WSNR or $\mathcal{NR}$PSNR, for instance, which measured the image quality between reconstructed and original images. The experimental results show that the solely usage of the Forward Perceptual Quantization improves the JPEG2000 compression and image perceptual quality. In addition, when both Forward and Inverse Quantization are applied into Hi-SET, it significatively improves the results regarding the JPEG2000 compression.

Acknowledgment

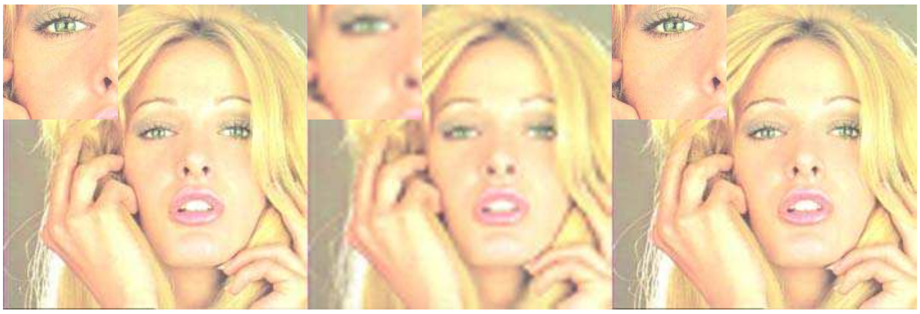
This work is supported by National Polytechnic Institute of Mexico by means of Project No. 20140096, the Academic Secretary and the Committee of Operation and Promotion of Academic Activities (COFAA), National Council of Science and Technology of Mexico by means of Project No. 204151/2013, and LABEX Σ-LIM France, Coimbra Group Scholarship Programme granted by University of Poitiers and Region of Poitou-Charentes, France.

References

1. M. Boliek, C. Christopoulos, and E. Majani, *Information Technology: JPEG2000 Image Coding System*, JPEG 2000 Part I final committee draft version 1.0 ed., ISO/IEC JTC1/SC29 WG1, JPEG 2000, April 2000.



(a) Girl 2

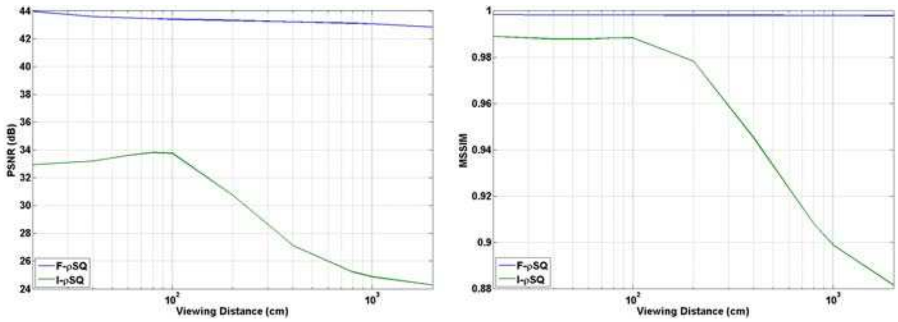


(b) Tiffany

(c) Peppers

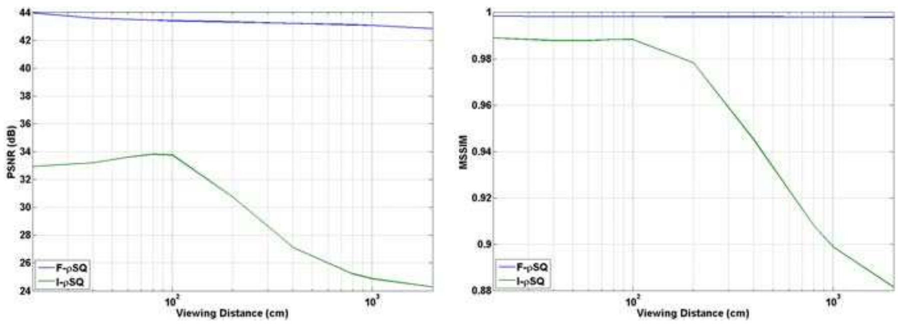
Figure 7. Visual examples of Perceptual Quantization. Left images are the original images, central images are forward perceptual quantized images (F-ρSQ) after applying $\alpha(\nu, r)$ at $d = 2000$ centimeters and right images are recovered I-ρSQ images after applying $\widehat{\alpha}(\nu, r)$.

2. S. I. P. I. of the University of Southern California. (1997) The USC-SIPI image database. Signal and Image Processing Institute of the University of Southern California. [Online]. Available: http://sipi.usc.edu/database/
3. C. Research, École Polytechnique Fédérale de Lausanne, and Ericsson. (2001) JJ2000 implementation in Java. Cannon Research, École Polytechnique Fédérale



(a) PSNR (b) MSSIM

Figure 8. PSNR and MSSIM assessments of compression of Gray-scale Images (Y Channel) of the CMU image database. Green functions denoted as F-ρSQ are the quality metrics of forward perceptual quantized images after applying $\alpha(\nu, r)$, while blue functions denoted as I-ρSQ are the quality metrics of recovered images after applying $\widehat{\alpha}(\nu, r)$.



(a) PSNR (b) MSSIM

Figure 9. PSNR and MSSIM assessments of compression of Color Images of the CMU image database. Green functions denoted as F-ρSQ are the quality metrics of forward perceptual quantized images after applying $\alpha(\nu, r)$, while blue functions denoted as I-ρSQ are the quality metrics of recovered images after applying $\widehat{\alpha}(\nu, r)$.

de Lausanne and Ericsson. [Online]. Available: http://jj2000.epfl.ch/

4. T. Mitsa and K. Varkur, "Evaluation of contrast sensitivity functions for formulation of quality measures incorporated in halftoning algorithms," *IEEE International Conference on Acustics, Speech and Signal Processing*, vol. 5, pp. 301–304, 1993.