

## Gain control for the chromatic channels in JPEG2000

Y. Navarro, J. Rovira, J. Gutiérrez and J. Malo

Dept. d'Òptica, Facultat de Física, Dept. d'Informàtica, Escola Superior d'Enginyeries, Universitat de València. Dr. Moliner 50. 46100 Burjassot, València, (SPAIN)

Corresponding author: Jesús Malo (jesus.malo@uv.es)

### ABSTRACT

The variety of perception models included in JPEG2000 are the key for its rate-distortion performance. However, the most powerful perceptual model (masking using non-linear divisive normalization) had to be simplified to make it invertible<sup>1</sup>. Recently, a robust and fast algorithm to invert the exact expression of divisive normalization has been derived and successfully used in the achromatic DCT context<sup>2</sup>. In this work we extend the above inversion method<sup>2</sup> to color images in wavelet representations to avoid the limitations of the simplified models included in JPEG2000.

### 1. INTRODUCTION

The aim of the image representation in transform coding is to represent the signal in a domain where its components are as independent as possible in order to allow scalar quantization<sup>3</sup>. Images that are intended to be judged by a human viewer contain two kinds of redundancies: statistical, due to the spatial relations of the tristimulus values in neighboring pixels; and perceptual, because the subjective distortion of color in a pixel depends on the surround. To achieve the desired independency, both redundancies have to be removed<sup>2,4,5</sup>. It has been argued that the goal of the early stages of human vision is to remove the statistical dependencies between the samples of the signal<sup>6</sup>. Therefore, accurate models of human image representation are appropriate for image coding.

The current models of human visual processing involve two stages: a linear wavelet-like transform of each opponent color channel, followed by a non-linear mapping, known as *divisive normalization*<sup>7</sup>. Accordingly, the coding scheme should use this representation and the decoding should invert it after the quantization. The first attempts to use this procedure were too simple. The original JPEG standard<sup>8</sup> used the block-DCT as the first stage and a simple linear approximation of the second one (using achromatic and chromatic CSFs). It has been reported that more accurate approximations of the second non-linear stage significantly improve the JPEG results<sup>5,9-11</sup>. The new standard JPEG2000 uses wavelets in the first linear stage<sup>1</sup>; and incorporates more complex expressions, *yet not exact*, for the non-linear stage. In particular, in addition of the simplest linear model, it allows point-wise non-linearities and simplified versions of the divisive normalization<sup>1</sup>. The major problem found in using the most accurate version of divisive normalization is the mathematical complexity of its inversion. This is why the standard just uses simplified versions of the non-linearity.

However, recently, a robust and fast algorithm to invert the exact expression of divisive normalization has been derived and used in the DCT context<sup>2</sup>. Accordingly, significant improvements in rate distortion performance have been found for achromatic images. In this work we extend the above method<sup>2</sup> to color images in wavelet representations as an alternative to JPEG2000.

### 2. QUANTIZATION IN THE DIVISIVE NORMALIZED DOMAIN

Assuming an opponent color space decomposition of the image, let  $a_{il}$  be the  $i$ -th (out of  $N_i$ ) wavelet coefficient of the  $l$  color channel. The current model of gain control (or masking) in the achromatic channel, and in each opponent color channel is<sup>6,7,12</sup>:

$$r_{il} = \frac{|\alpha_{il} \cdot a_{il}|^\gamma}{\beta_{il} + \sum_{j=1}^{N_i} h_{ij} \cdot |\alpha_{jl} \cdot a_{jl}|^\gamma} \quad (1)$$

where,  $\alpha_{il}$  stands for the CSF of the corresponding color channel<sup>13</sup>, and  $\beta$  and  $h$  are the parameters needed to fit the psychophysical contrast incremental thresholds of the basis functions<sup>2,7,12</sup>. The matrix,  $h$ , defines the *gaussian* neighborhoods (in space, scale and orientation) that describe the masking interactions between *all* the coefficients of the vectors  $a_l$ .

In JPEG2000, uniform scalar quantization is applied to the above perceptual representation<sup>1</sup>, thus restricting the Maximum Perceptual Error in each coefficient of each color channel<sup>2,5,11</sup>. However, the eventual advantages of the above scheme are not fully exploited in JPEG2000 because the masking neighborhood for each coefficient  $i$ ,  $h_{ij}$ , is restricted to be causal in the same subband to make the normalization invertible<sup>1</sup>. In this way, most of the masking interactions are neglected in this invertible but oversimplified gain control model.

The coding scheme proposed here is also based in the scalar quantization of the non-linear normalized representation. However, here we propose to use the full gaussian neighborhood over space, scales and orientations, instead of the simplified neighborhood recommended in the standard. The practical use of this full neighborhood masking model is made possible because we apply the series expansion inversion method<sup>2</sup> to come back to the wavelet domain after the quantization.

### 3. RESULTS AND CONCLUSIONS

Figures 1 and 2 compare the results (images and MSE rate-distortion) of the proposed (full masking) model with the three different simplified perception models considered in JPEG2000.

On one hand, the JPEG2000 models that do not consider adaptive masking (linear and point-wise non-linear models) achieve poor results: they over saturate the colors and introduce false colors in the opponent chromatic direction. Note the blue edges in the red-green image (fig.1) and the false red details in the blue-yellow image (fig.2). The simplified divisive normalization model removes these color artifacts but it reduces the contrast of high-frequency details and the sharpness of the edges. On the other hand, the proposed model appropriately reproduces color while preserving the spatial sharpness of the image. Moreover, rate-distortion results show that the proposed model reduces about 35% the volume of the signal for constant MSE at 0.2-0.3 bits/pix.

These results suggest that the consideration of *all* the masking interactions in the gain control model (which is technically possible thanks to the series expansion inversion method<sup>2</sup>) is a key issue in order to improve the visual quality and the rate-distortion performance of JPEG2000.

#### Acknowledgement

TIC2003-01504 (CICYT-FEDER), Grupos04-08 (GV), and the program DN8644.

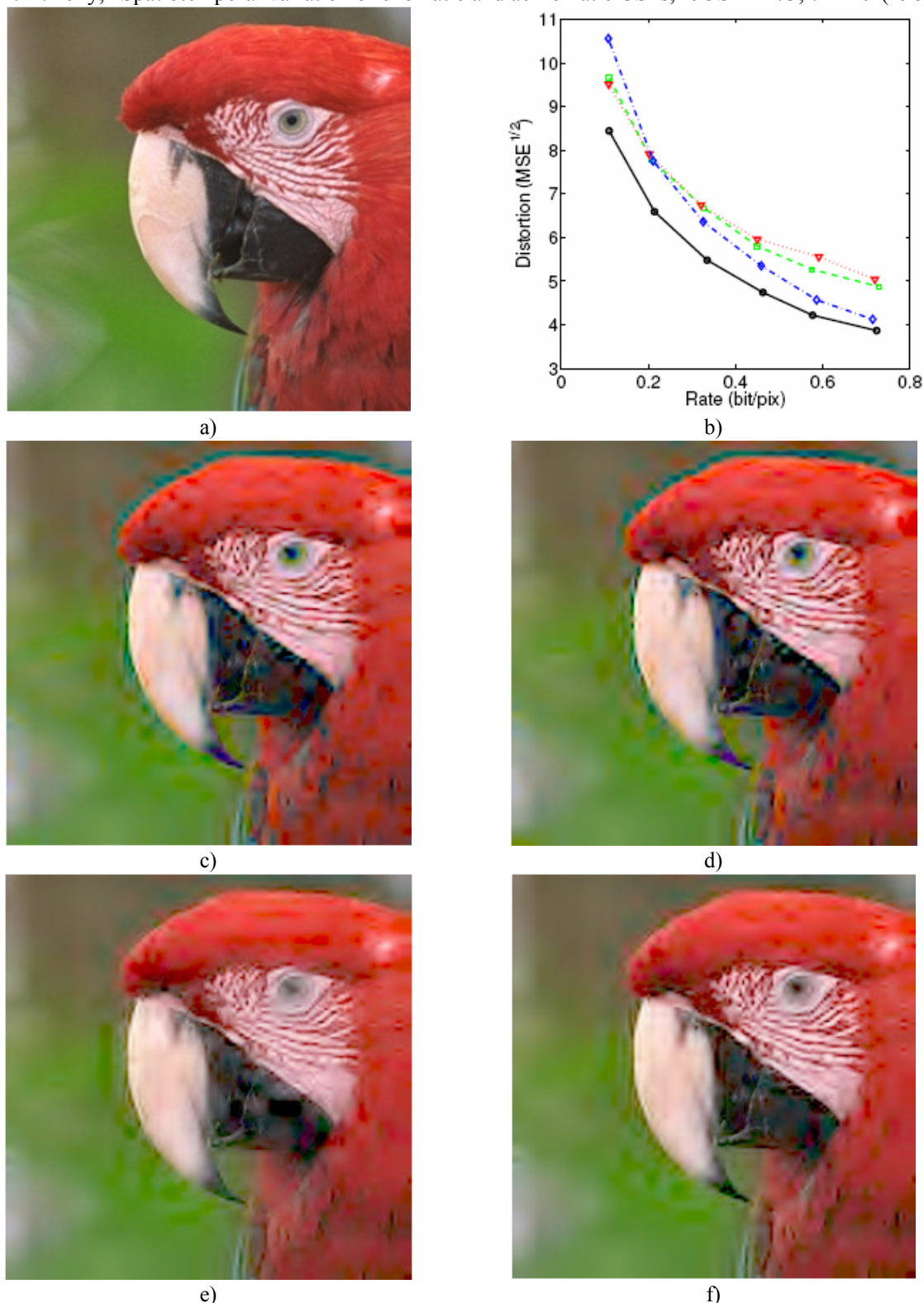
#### References

1. W. Zeng, S. Daly and S. Lei, "An overview of the visual optimization tools in JPEG2000," *Sign. Proc. Im. Comm.* 17(1), 85-104 (2002).
2. J. Malo, I. Epifanio, R. Navarro, and E. Simoncelli, "Non-linear image representation for efficient perceptual coding," to appear in *IEEE Trans. Im. Proc.* (2005).
3. A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression* (Kluwer, Boston, 1992).
4. A. B. Watson, "Perceptual-components architecture for digital video," *JOSA A* 7, 1943-54 (1990).
5. J. Malo, J. Gutierrez, I. Epifanio, F. Ferri and J. M. Artigas, "Perceptual feed-back in multigrid motion estimation using an improved DCT quantization," *IEEE Trans. Im. Proc.* 10 (10), 1411-1427 (2001).
6. O. Schwartz and E. P. Simoncelli, "Natural signal statistics and sensory gain control," *Nat.Neur.* 4, 819-25 (2001).
7. A. B. Watson and J. Solomon, "A model of visual contrast gain control," *JOSA A* 14, 2379-91, 1997.
8. G. K. Wallace, "The JPEG still picture compression standard," *Com. of the ACM* 34(4), 31-43 (1991).
9. S. Daly, "Application of a noise-adaptive CSF to image data compression," *Opt.Eng.* 29, 977-87 (1990).
10. A. B. Watson, "DCT quantization matrices visually optimized for individual images," In B.E. Rogowitz, editor, *Human Vision, Visual Processing and Digital Display IV*, volume 1913 (1993).

11 J. Malo, F. Ferri, J. Albert, and J. Soret, "Comparison of perceptually uniform quantization with average error minimization in image transform coding," *Electronics Letters* 35(13), 1067-1068 (1999).

12 E. Martinez-Uriegas. Color detection and color contrast discrimination thresholds. In *Proceedings of the OSA Annual Meeting ILS-XIII*, pp 81, Los Angeles, 1997.

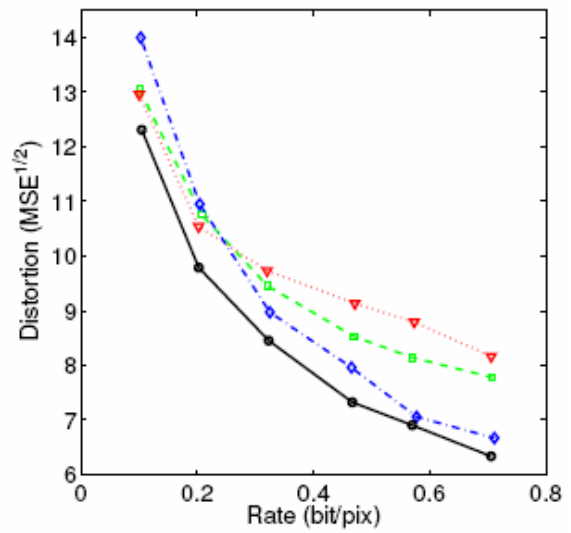
13 D. H.Kelly, "Spatiotemporal variation of chromatic and achromatic CSFs," *JOSA A* 73, 742-49 (1983).



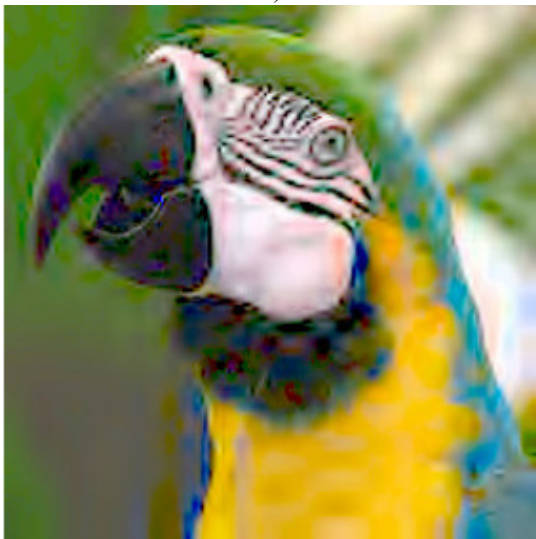
**Figure 1:** Coding results (0.2 bits/pix). **a)** Original. **b)** Rate-Distortion: linear JPEG2000 (dashed), Point-wise non-linear JPEG2000 (dotted), Simplified gain control JPEG2000 (dash-dot), and the proposed method (solid). **c)** linear JPEG2000. **d)** Point-wise non-linear JPEG2000. **e)** Simplified gain control JPEG2000. **f)** The proposed method.



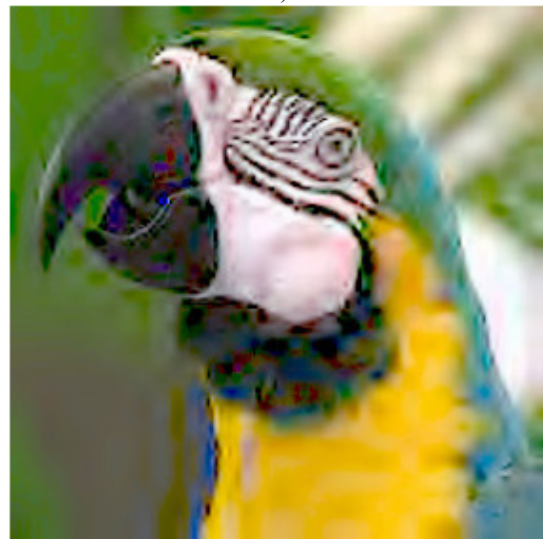
a)



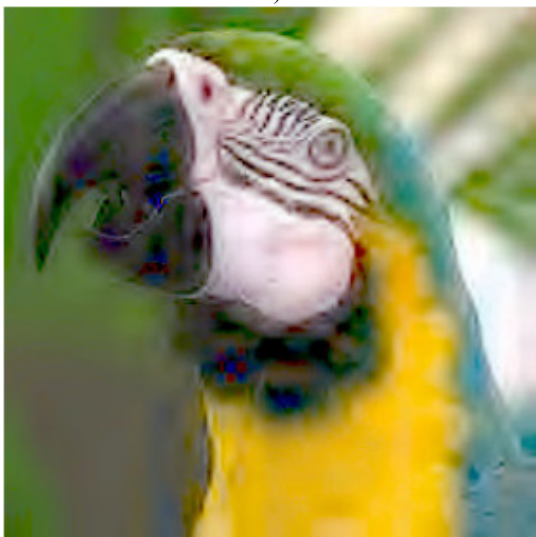
b)



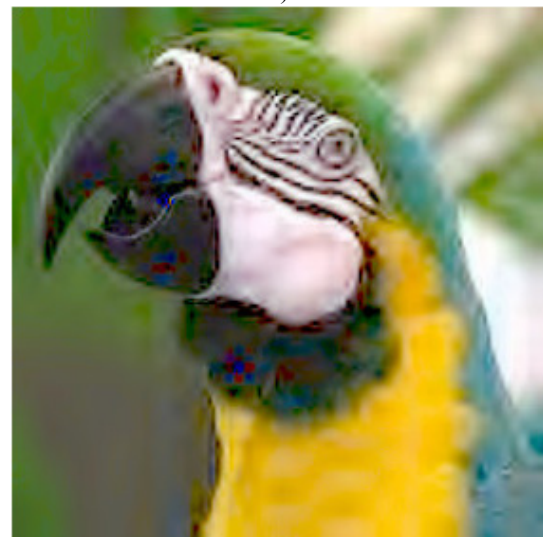
c)



d)



e)



f)

**Figure 2:** Coding results (0.2 bits/pix). **a)** Original. **b)** Rate-Distortion: linear JPEG2000 (dashed), Point-wise non-linear JPEG2000 (dotted), Simplified gain control JPEG2000 (dash-dot), and the proposed method (solid). **c)** linear JPEG2000. **d)** Point-wise non-linear JPEG2000. **e)** Simplified gain control JPEG2000. **f)** The proposed method.