# Rational Prioritization of Visual Information

Computer Vision Group[1*]

[1]Department of Computer Science and Artificial Intelligence, University of Granada, CITIC-UGR, 18071 Granada, Spain

[*]To whom correspondence should be addressed; E-mail: jags@decsai.ugr.es

**Way back when we began to develop this approach, our objectives were clear. We aspired to develop a sound, practical theory for organizing the quantizer-dependent quality of encoding in progressive image transmission. The basic elements of this theory would be a novel scheme for information prioritization, the mechanism of bit rate allocation among competing quantizers, and the strategy for coder performance evaluation.**

## Introduction

In the coding community the standard approach is often to use the peak signal to noise ratio (PSNR) for coder evaluation, the largest reduction in some square error for information priorization, and a theory of (minimal) distortion as a function of rate for bit allocation. Thus, any coding scheme that does not attempt to minimize some square-error cannot be expected to prove its worth with a curve of PSNR versus bit rate, which may be a constraint on the formulation of new coding schemes capable of making an intelligent use of visual information. This may be justified assuming the correctness of the PSNR, but what are the actual properties of the PSNR? For example, does it take into account the effectiveness of the information, so discriminating relevant structures from unwanted detail and noise? Does it examine whether the properties of the original image at significant points are equal to the properties of the decoded output at corresponding locations? The point is that whereas we have no evident affirmative answer to these and other questions, the PSNR does not appear capable of predicting visual distinctness from digital imagery as perceived by human observers.

Regarding the issue of information prioritization, standard schemes prioritize the code bits often according to their reduction in distortion, and a major objective in this context is to select the most important information—which yields the largest distortion reduction—to be transmitted first, where the distortion is usually a squared-error metric. Since the quality of the reconstructions at different bit rates strongly depends on the visual distinctness of the perceived data, the information selected to be transmitted first by any prioritization scheme at each truncation time should achieve the largest visual distinctness over still-to-be-transmitted data. The natural question is whether a squared-error metric is capable to rank order visual information with respect to the visual distinctness as measured by humans, and thus, the largest squared-error reduction can be used to prioritize, with reliability, the most important information according to their distinctness.

With respect to the issue of distribution of bits, in the standard approach of Rate-Distortion Theory (the idea of rate distortion was introduced in 1948 by Shannon) a bit rate allocation problem among competing quantizers (e.g., spatial regions) is optimally solved for a given bit budget if the marginal change in distortion is the same for all regions. Again, the squared-error metric is the most popular distortion measure used for continuous alphabets. Its advantages are its simplicity and its relationship to least squares prediction. To our understanding, the problem with the standard approach of rate (square-error) distortion is that there exist some questions yet to be answered concerning the properties that obey its solution for bit allocation among competing regions. For example, if we view the regions as citizens of a society, does this solution respect the views of the regions (citizens)? Is this solution blind to the kind of objects that the regions contain? Could it be interpreted as a fair aggregation of individual interests? Does the solution of the distribution problem change by virtue of a change in the scale of the benefits regions receive from their respective allocations?

## Developing an Axiomatic Approach for Rationality in Progressive Transmission

Here we propose that a different approach to solve the problems of evaluation, prioritization, and distribution can be to first state some general principles that the solution of the problem in each case must obey, and then derive the solution that satisfies exactly the principles. The axioms may, of course, be incompatible. It is not rare that one would like to impose more axioms that are jointly compatible. It may also happen that the axiomatic solution resulting from a list of axioms that all seem appealing is found to behave unsatisfactorily in some significant example. To overcome this problem, one must formalize the example and state an additional axiom that specifies how the solution should behave in this situation, and finally determine the greatest subset of axioms from the original list that are compatible with the new axiom. Of course, compatibility may hold for several distinct such subsets. In any case, the critical difference with respect to the approaches discussed above is that we will be able to predict exactly the behavior of the axiomatic solution according to its principles. For example, the principles of rationality avoid certain forms of behavioral inconsistency in situations in which choices are to be made among available quantizers for their prioritization; the principles of cooperation among quantizers may be needed to increase their risk tolerance in variable-resolution compression, and the principle of justice provides conditions for fair quantizer formation.

In a rational system for transmission, a discrete wavelet transform provides a representation of the original image. A tree structure, called a spatial orientation tree, naturally defines the spatial relationship in the pyramid that results from the transformation. Each node of the tree corresponds to a pixel, and its direct descendants (offspring) correspond to the pixels of the same spatial orientation in the next finer level of the pyramid. Transform coefficients in a spatial orientation tree correspond to a particular region of the original image, and thus, each spatial orientation tree is associated with one spatial region. Individual trees may be grouped together to form a reduced number of quantizers that convey structural information about the picture to the rational transmission. A "just" quantizer formation will give no tree a cause for "reasonable regret" in rational progressive transmission (Reference (*1*)).

That is, they are all able to achieve the same overall success. The basic assumption is that justice requires compensating individual spatial orientation trees for aspects of their prioritization for which they are not responsible and which hamper their achievement of whatever is valuable in their own transmission. Differences for which they are responsible may be ruled by rationality (Reference (*2*)).

A simple condition to perform a just quantizer formation can be the equality of the a priori importance of the spatial orientation trees that are grouped together in one quantizer, from which we understand the key role of the a priori importance of a tree in the development of a theory for just quantizer formation.

A prioritization protocol whereby the ordering of importance is determined within a rational approach, chooses at each truncation time among alternative quantizers for further transmission in such a way as to avoid certain forms of behavioral inconsistency (Reference (*2*)). The system may exhibit either a risk-seeking posture with respect to "gambles" on quantizer-dependent quality of encoding or risk-averse behavior.

By changing its risk attitude within a rational approach that avoids certain forms of behavioral inconsistency, a quantizer may modify the gain in benefit that results from a particular bit stream candidate to be transmitted at a truncation time. At medium and high bit rates, quantizers exhibit only low risk tolerance since they are aware that the next truncation time might be the last one (Reference (*4*)).

Anyway, since at extremely low bit rates the target bit rate may be far away, quantizers are able to exhibit higher risk tolerance, and as a consequence, they will have a greater possibility of accelerating their benefit gain. The cooperation among subsets of the quantizers may be needed to increase the risk tolerance at very low bit rates within a rational approach and still prioritize first the more relevant pieces of information at each truncation time (References (*3*)): The members of any coalition of quantizers can then negotiate a feasible change in the risk attitudes of the quantizers of the coalition that would benefit them all. The final risk tolerance of different quantizers comes from the balance of power among the coalitions of quantizers; and the prioritization protocol chooses to transmit, at each truncation time, a bit stream for the quantizer that receives the highest payment (per coding bit) in a coalitional game that minimizes the dissatisfaction of coalitions.

Experimental results should illustrate the comparative performance of the rational system against the state of the art in progressive transmission. Reference (*6*) shows the principles of a visual distinctness measure that can be used to evaluate image compression methods. The book ends with an epilogue that summarizes the key results and conclusions plus four appendixes containing basic background material.

All software (with documentation) developed in the book may be accessed on the Internet site `http://decsai.ugr.es/cvg/REWIC` or by anonymous ftp to `decsai.ugr.es` with the path `pub/cvg/software`. All material is made available to other researchers for academic use only.

This is intended to be a simple and accessible study on the role of rationality, cooperation, and justice in progressive image transmission. From the above discussion, it is clear that we were drawn to the problem of progressive transmission from backgrounds in theories of distributive justice and game theory, because of the difficulty of capturing the concept of relative information for predicting visual distinctness from 2D digital images. We hope that you find a few key ideas and techniques that provide intuition toward new questions, and also that our answers to problems of evaluation, prioritization, and distribution allow extensive interpretation. For example, the information theoretic measure for predicting visual distinctness and the expected increase in utility for information prioritization are related (Reference (*7*)).

# Experimental Results

We have developed an algorithm, named after REWIC, following the rational approach to progressive transmission as described in (*2*). Also we have developed an algorithm called "Rational Embedded coder with CONstraints (RECON)" (see (*5*)), which implements the rational system for transmission at very low bit rates.

Here we firstly provide a set of psychophysical experiments to test the comparative subjective quality of images reconstructed using the state of the art in progressive transmission, SPIHT (*8*), REWIC, and RECON. Second, we perform a thorough comparison of RECON and SPIHT on a data set of 100 test images using an objective coder selection procedure.

## 0.1 Subjective coder evaluation

### 0.1.1 Experiment 1

Here we perform a comparison of subjective performance of the state-of-the-art coder in progressive transmission SPIHT with RECON and REWIC, using a psychophysical experiment. To this aim, test image $\#24$ (Fig. 1) was first compressed to the same very low bit rates using the three compression methods. Figure 2 shows the respective reconstructions at 0.0156, 0.0312, 0.0625, and 0.08 bpp.

Fifteen volunteers subjectively evaluated the reconstructed images following an ITU-R Rec-


Figure 1: Data set of standard $512 \times 512$ grayscale test images.

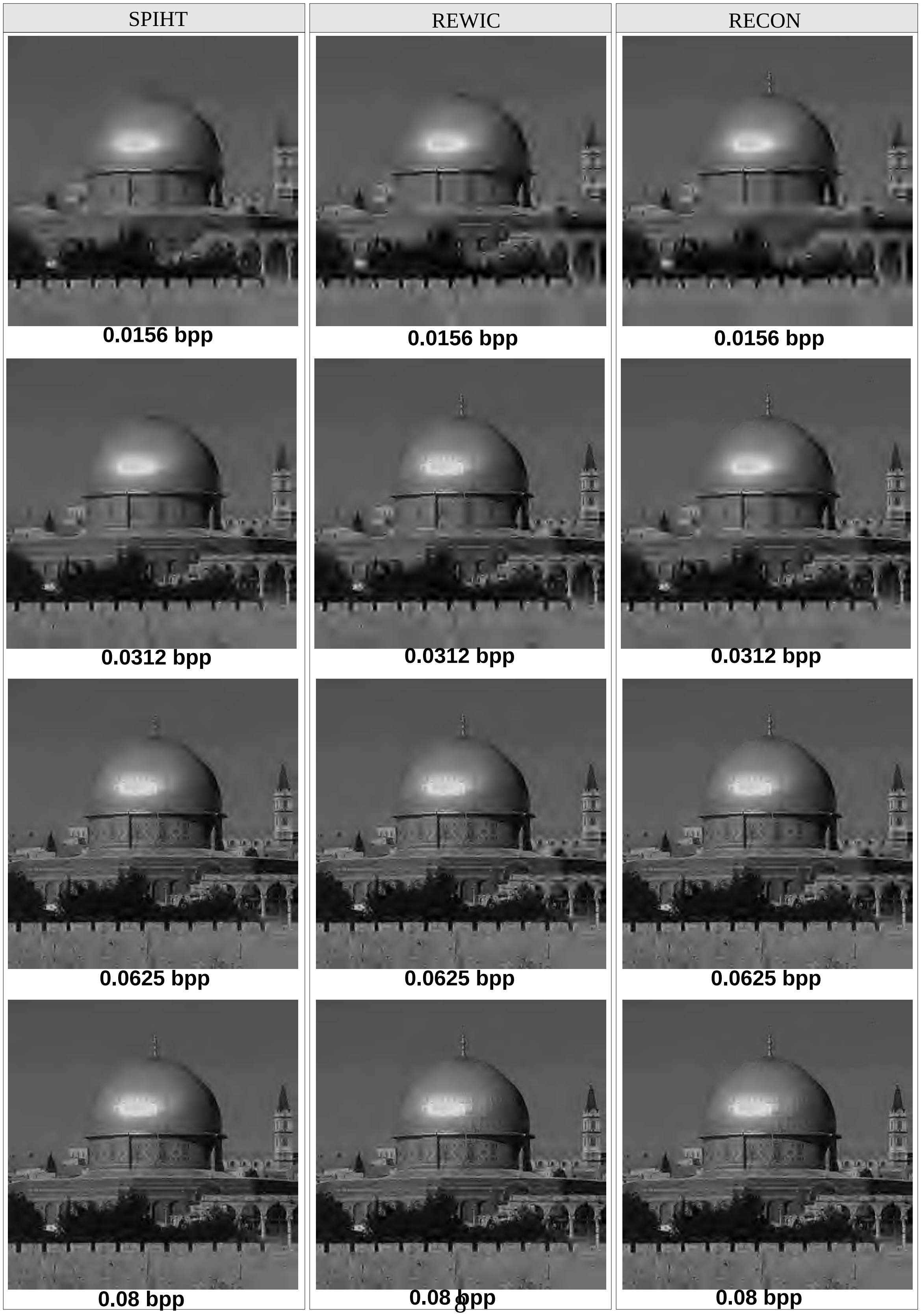


Figure 2: Reconstructions of the test image # 24 using SPIHT, REWIC, and RECON at 0.08, 0.0625, 0.03125, and 0.015625 bpp.

Table 1: Quality factors given by human observers.

| bpp | MEAN QUALITY FACTOR | | |
|---|---|---|---|
| | *SPIHT* | *REWIC* | *RECON* |
| **0.015625** | 1.00 | 1.07 | 1.33 |
| **0.03125** | 1.27 | 2.07 | 2.07 |
| **0.0625** | 1.80 | 2.27 | 2.47 |
| **0.08** | 2.33 | 2.93 | 2.73 |
| $MEAN$ | 1.60 | 2.09 | 2.15 |

ommendation 500-10. Table 1 summarizes mean quality factors for reconstructions illustrated in Fig. 2. As can be seen from this table, quality factors predict a better visual fidelity using RECON than with the SPIHT reconstructions. We know that the visual quality of SPIHT decoded outputs is bad at 0.0625, 0.03125, and 0.015625 bpp. However, the visual quality of RECON reconstructions is bad only at 0.015625 bpp.

Computational times (in seconds) for REWIC and RECON on several test images of the database are shown in Table 2. The times here are for an Intel Pentium IV at 2.4 GHz.

### 0.1.2 Experiment 2

The test image $\#3$ (Fig. 1) was first compressed at 0.0156, 0.0312, 0.0625, and 0.08 bpp using SPIHT and RECON. Figure 3 shows the respective reconstructions.

Fifteen volunteers nonexpert in image compression subjectively evaluated the reconstructed images using ITU-R Recommendation 500-10. Table 3 summarizes the mean quality factors for different decoded outputs using the compression methods.

Figure 4 shows plots of rate vs. PSNR and rate vs. CG for SPIHT and RECON at 0.0156, 0.0312, 0.0625, and 0.08 bpp. As can be seen from these plots, the PSNR predicts that the SPIHT results in a higher image fidelity than RECON, which does not appear to correlate with the subjective quality estimated by human observers (see Table 3). On the contrary, the overall

Table 2: Computational time for REWIC and RECON.

| *Computational time (in seconds) for REWIC and RECON* | | | |
|---|---|---|---|
| $IMAGE$ | Rate (bpp) | REWIC | RECON |
| # 3 | $0.015625$ | $1.3$ | $1.5$ |
| | $0.03125$ | $1.3$ | $1.7$ |
| | $0.0625$ | $1.3$ | $1.9$ |
| | $0.125$ | $1.3$ | $1.9$ |
| | $0.25$ | $1.3$ | $1.9$ |
| | $0.5$ | $1.4$ | $2.0$ |
| # 24 | $0.015625$ | $1.2$ | $1.5$ |
| | $0.03125$ | $1.2$ | $1.6$ |
| | $0.0625$ | $1.2$ | $1.8$ |
| | $0.125$ | $1.3$ | $1.8$ |
| | $0.25$ | $1.3$ | $1.8$ |
| | $0.5$ | $1.3$ | $1.9$ |
| # 65 | $0.015625$ | $1.1$ | $1.3$ |
| | $0.03125$ | $1.1$ | $1.5$ |
| | $0.0625$ | $1.1$ | $1.7$ |
| | $0.125$ | $1.2$ | $1.7$ |
| | $0.25$ | $1.2$ | $1.7$ |
| | $0.5$ | $1.2$ | $1.8$ |

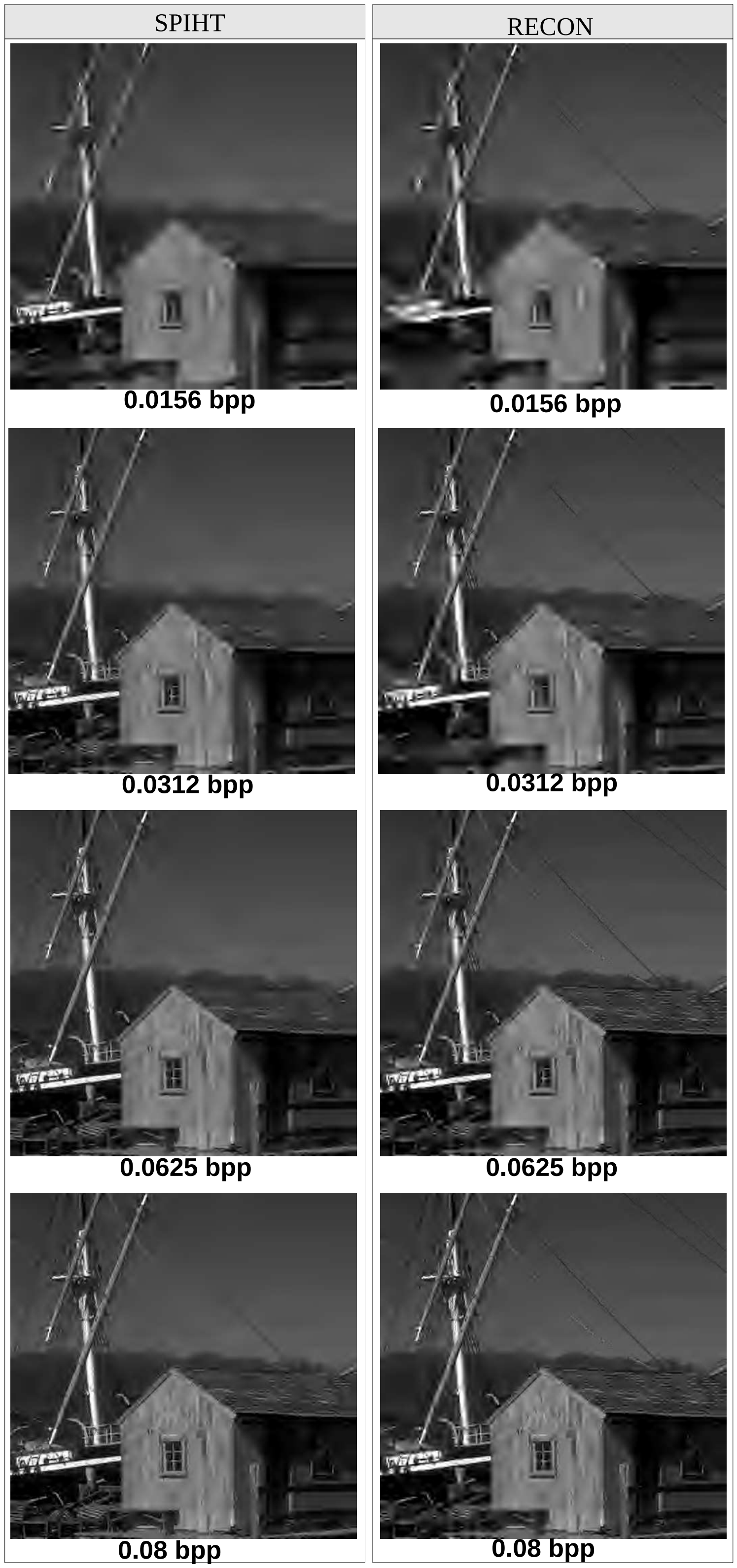


Figure 3: Reconstructions of the test image #3 using SPIHT and RECON at 0.0156, 0.0312, 0.0625, and 0.08 bpp.

Table 3: Quality factors given by human observers.

| bpp | MEAN QUALITY FACTOR | |
|---|---|---|
| | *SPIHT* | *RECON* |
| **0.015625** | 1.00 | 1.07 |
| **0.03125** | 1.47 | 1.87 |
| **0.0625** | 1.80 | 2.47 |
| **0.08** | 2.47 | 2.93 |
| $MEAN$ | 1.69 | 2.09 |

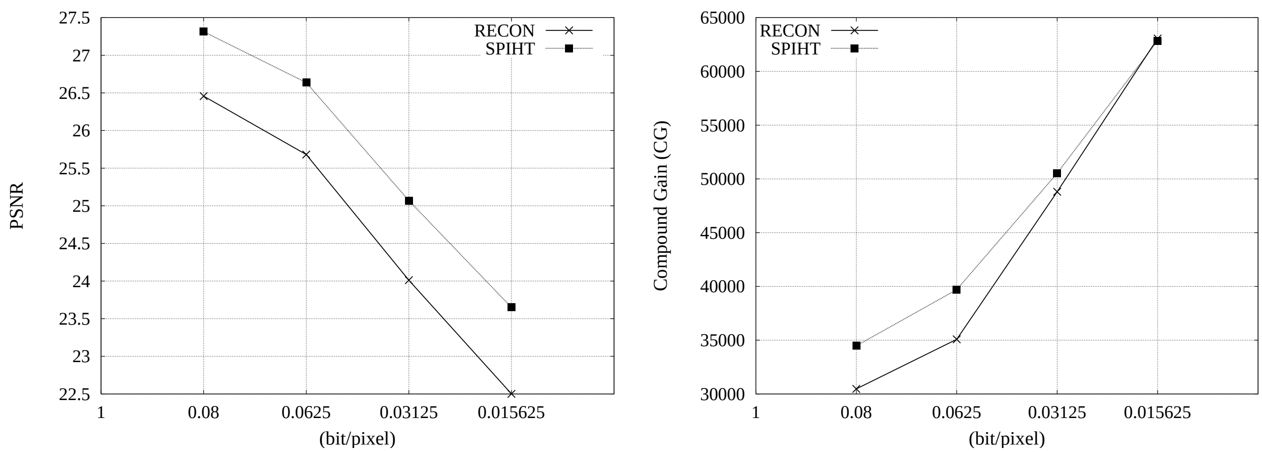


Figure 4: Plots of rate vs. PSNR and rate vs. CG for SPIHT and RECON at 0.0156, 0.0312, 0.0625, and 0.08 bpp.

impression is that, as predicted by the CG, RECON results in a higher image fidelity than SPIHT, which correlates with subjective fidelity by human observers in Table 3. Recall that an optimal coder in the CG sense tends to produce the lowest value of the CG.

### 0.1.3 Experiment 3

A new test image was compressed using the SPIHT and RECON coders. Figure 5 shows the respective reconstructions at 0.0156, 0.0312, 0.0625, and 0.08 bpp. A psychophysical experiment was also performed, and again, 15 volunteers subjectively evaluated the reconstructed images using the ITU-R Recommendation 500-10. Table 4 summarizes the mean quality factors that

Table 4: Quality factors given by human observers.

| bpp | MEAN QUALITY FACTOR | |
|---|---|---|
| | *SPIHT* | *RECON* |
| **0.015625** | 1.00 | 1.07 |
| **0.03125** | 1.13 | 1.87 |
| **0.0625** | 2.20 | 2.87 |
| **0.08** | 3.13 | 3.13 |
| $MEAN$ | 1.87 | 2.24 |

were provided by this subjective evaluation.

Figure 6 shows plots of rate vs. PSNR and rate vs. CG for RECON and SPIHT at 0.08, 0.0625, 0.03125, and 0.015625 bpp. The PSNR predicts that SPIHT results in a higher image fidelity than RECON, which does not appear to correlate with the subjective quality estimated by human observers (Table 4). On the contrary, as can be seen from Fig. 6, the CG predicts that RECON results in a higher image fidelity than SPIHT, which correlates with the subjective fidelity by human observers given in Table 4.

## 0.2 Objective coder evaluation

Here we perform a more thorough comparison of RECON and SPIHT, based on the objective coder selection procedure presented in (*6*). Tests reported here were performed on the data set composed of 100 standard $512 \times 512$ grayscale test images shown in Fig. 1.

Tables 5 and 6 summarize the results of this experiment on the test images of the data set in Fig. 1: 25 out of 100 test images (25%) have passed conditions (1) and (2) in the coder selection procedure. Hence, RECON is significantly better than SPIHT with a high confidence level for 25% of the data set of test images, whereas SPIHT is better than RECON for 1% of the images.

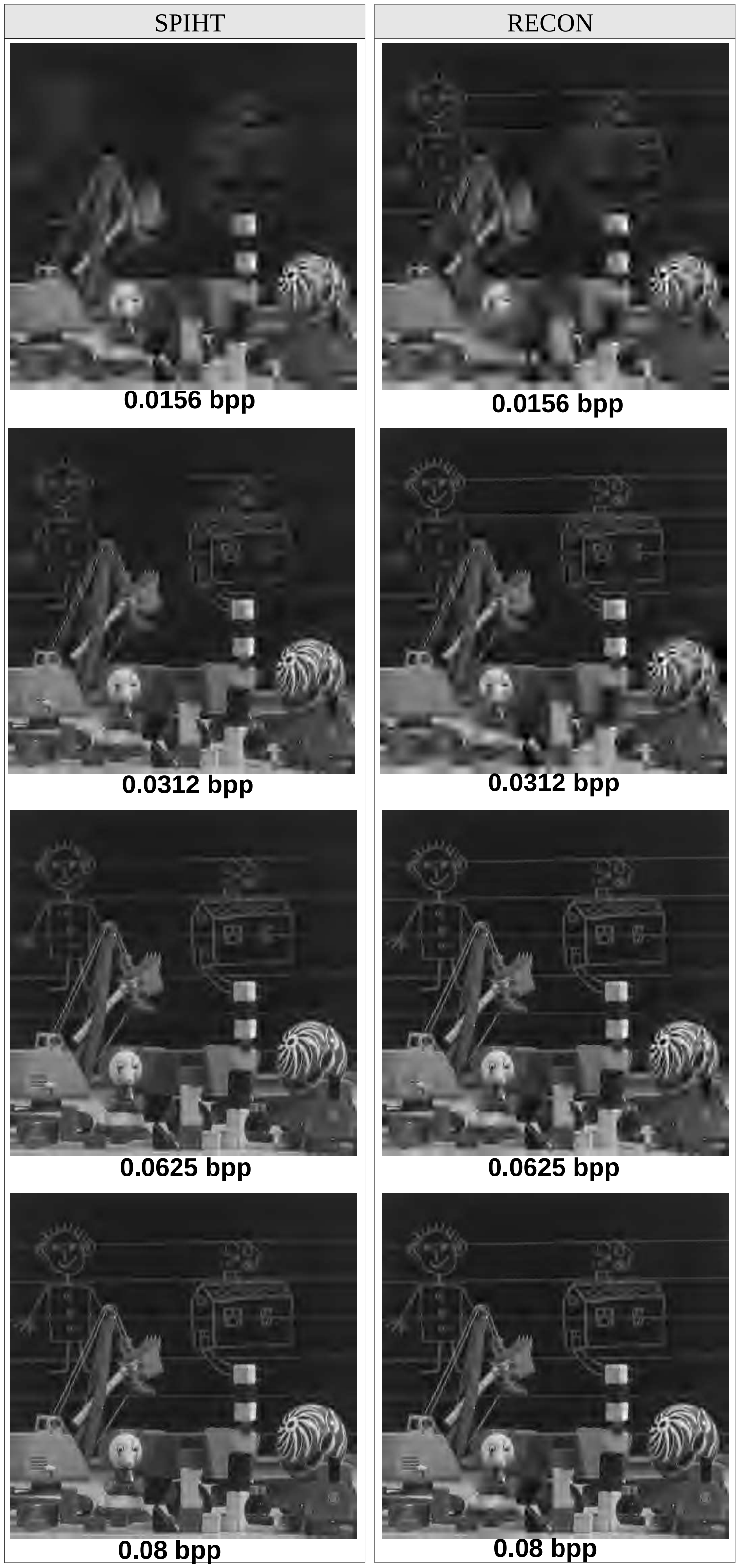


Figure 5: Reconstructions of test image #65 using SPIHT and RECON at 0.0156, 0.0312, 0.0625, and 0.08 bpp.

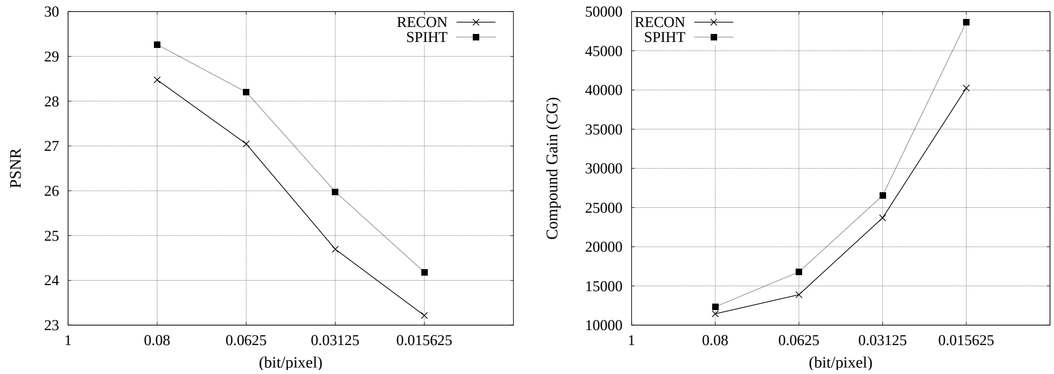


Figure 6: For image $\#65$, plots on rate vs. PSNR and rate vs. CG for RECON and SPIHT at 0.08, 0.0625, 0.03125, and 0.015625 bpp.

Table 5: Coder selection procedure.

| *CODER SELECTION PROCEDURE WITH % CONFIDENCE (SPIHT/RECON)* | | | |
|---|---|---|---|
| Image Number | Condition 1 (y/n) | Condition 2 (y/n) | Confidence |
| $16, 25, 26, 27, 35, 39, 41, 42, 49, 55, 63, 65, 66, 71, 77, 81, 88, 89, 93, 95$ | $y$ | $y$ | 99% |
| 2 | $y$ | $y$ | 95% |
| 36, 57, 61, 67 | $y$ | $y$ | 90% |

Table 6: Comparative performance of RECON and SPIHT.

| Total Percentage of Images at which RECON/SPIHT is Significantly Better than SPIHT/RECON with at Least 90% Confidence | |
|---|---|
| **RECON** *better than* **SPIHT** | 25% |
| **SPIHT** *better than* **RECON** | 1% |

# References and Notes


1. J.A. García, R. Rodriguez-Sánchez, and J. Fdez-Valdivia, "Justice in quantizer formation for rational progressive transmission," *Opt. Eng.*, Vol. 43(9), pp. 2105-2119, (2004).

2. J.A. García, R. Rodriguez-Sánchez, J. Fdez-Valdivia, and X. R. Fdez-Vidal, "Rational systems exhibit moderate risk aversion with respect to 'gambles' on variable-resolution compression," *Opt. Eng.*, Vol. 41(9), pp. 2216-2237 (2002).

3. J.A. García, R. Rodriguez-Sánchez, J. Fdez-Valdivia, and X. R. Fdez-Vidal, "CORAL: Collective rationality for the allocation of bits," *Opt. Eng.*, Vol. 42(4), pp. 1000-1012 (2003).

4. J.A. García, R. Rodriguez-Sánchez, J. Fdez-Valdivia, and J. Martinez-Baena, "Self-control of quantizers' risk attitude in rational embedded wavelet image coding" *Opt. Eng.*, Vol. 42(11), pp. 3215-3234

5. J.A. García, R. Rodriguez-Sánchez, and J. Fdez-Valdivia, "An embedded coder providing better image quality at very low bit rates," *Opt. Eng.*, Vol. 43(3), pp. 615-627 (2004).

6. J.A. García, J. Fdez-Valdivia, X. R. Fdez-Vidal, and R. Rodriguez-Sánchez, "Information theoretic measure for visual target distinctness," *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol. 23(4), pp. 362-383 (2001).

7. R. Rodriguez-Sánchez, J. Fdez-Valdivia, A. Toet, and J.A. García, "The relationship between information priorization and visual distinctness in two progressive image transmission schemes," *Pattern Recognit.*, Vol. 37(2), pp. 281-297 (2004).

8. A. Said, and W.A. Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. on Circuit and System for Video Technology*, **6**(3), 243-250, (1996).

**Acknowledgments.** This paper was sponsored by the Spanish Board for Science and Technology (MICINN) under grant TIN2010-15157.