ELSEVIER

Image and Vision Computing 15 (1997) 535–548

image AND vision COMPUTING

Subjective image fidelity metric based on bit allocation of the human visual system in the DCT domain

J. Malo*, A.M. Pons, J.M. Artigas

Departament d'Óptica, Facultat de Física, Universitat de València, C/Dr. Moliner 50, 461 00 Burjassot, València, Spain

Received 15 July 1996; accepted 4 December 1996

Abstract

Until now, subjective image distortion measures have partially used diverse empirical facts concerning human perception: non-linear perception of luminance, masking of the impairments by a highly textured surround, linear filtering by the threshold contrast frequency response of the visual system, and non-linear post-filtering amplitude corrections in the frequency domain. In this work, we develop a frequency and contrast dependent metric in the DCT domain using a fully non-linear and suprathreshold contrast perception model: the *Information Allocation Function* (IAF) of the visual system. It is derived from experimental data about frequency and contrast incremental thresholds and it is consistent with the reported noise adaptation of the visual system frequency response. Exhaustive psychophysical comparison with the results of other subjective metrics confirms that our model deals with a wider range of distortions more accurately than previously reported metrics. The developed metric can, therefore, be incorporated in the design of compression algorithms as a closer approximation of human assessment of image quality. © 1997 Elsevier Science B.V.

Keywords: Subjective image fidelity; Non-linear perception model; DCT domain

1. Introduction

The importance of measuring image differences is central to many image and video processing algorithms.

The most straightforward application is image encoding design and optimisation. The final codebook depends on the selected metric controlling the iterative process which reduces the distance between the original and the encoded image [1,2]. It is known that in the applications judged by a human observer, Euclidean metrics such as mean square error lead to unsatisfactory results [3,4]. In those cases, the encoding procedure has to be designed to minimise the *subjective distortions* of the reconstructed signal. Characteristics of the human viewer (in fact a numerical model of its requirements and limitations) must be included to obtain a perceptually weighted metric that reproduces opinion of the observers [5]. Up to now, most successful image and video codec standards (JPEG [6], H.261 and MPEG-X [7,8]) can be qualified as perceptual oriented coders as they include certain human visual system (HVS) characteristics to match the encoding process to the observer requirements. All such techniques carry out a transform of the signal to a frequency domain (in particular DCT), not only due to the energy compactation and decorrelation properties of these transforms, but also due to the fact that HVS sensitivity is highly uneven in these domains [5,9]. An optimum frequency dependent bit allocation can be done in such a way that many transform coefficients can be safely discarded. These standards [6–8] and other compression schemes [10–12] have used an ad hoc knowledge of HVS threshold frequency response, but the amplitude stepsize non-linearities, when taken into account [10–12], are still based on statistical LBG design of the codebook using Euclidean metrics. If a reliable subjective metric were employed, the standard LBG algorithm for codebook design could be used without other ad hoc previous corrections.

The applications of subjective distortion metrics are not restricted to the design of a perceptually matched quantizer. Wherever image differences are computed and perceptual criteria have to be satisfied, this kind of measures can be applied. Two less exploited examples of this are motion estimation and adaptive noise cancellation. In region matching methods for motion estimation, the estimated displacement at a particular point is the vector that minimises *some difference measure* between a neighbourhood of the point in that frame and a neighbourhood of the displaced point in the previous frame [13]. In adaptive image enhancement through linear or non-linear filtering, the filter coefficients

* Corresponding author. E-mail: jesus.malo@uv.es.

0262-8856/97/$17.00 © 1997 Elsevier Science B.V. All rights reserved
PII S0262-8856(96)00004-2

locally change to minimise *some difference measure* between the output of the filter and the target reference [14,15].

All these techniques could be perceptually matched if a reliable subjective distortion measure were used to guide them.

The basic idea of any perceptually weighted metric is that subjective differences between two images cannot be directly derived from the *given* images, but from their *perceived version*. According to this, the way we model the information reduction process that the visual system applies to the input signals, is the key to obtain a good subjective fidelity metric.

Until now, the reported subjective distortion measurements have been based on selective weighting of image differences in the spatial or/and the frequency domain [3,4,16–21]. This selectivity has been founded on a perception model implementing at least one of these processes:

1. Application of a point-wise cube-root non-linearity on intensity to take into account the non-linear response of the photoreceptors to luminance [17,21,22].
2. Masking effects, weighting the intensity differences according to the values of the gradient, the local activity, the local contrast or the local luminance [3,4,21], to take into account the major sensitivity of HVS to deviations in sharp edges, and the masking of noise in highly textured areas.
3. Linear band-pass filtering by the visual system's threshold frequency response [16–22]. The perception models based on a linear filtering scheme assume that frequency sensitivity of the HVS is inversely proportional to the contrast detection threshold of sinusoidal patterns. The visual system frequency response, named *Contrast Sensitivity Function* (CSF) in the psychophysical literature [23–25], is empirically defined as the inverse of the contrast detection threshold of sinusoidal gratings of different frequencies. These linear and threshold models do not include suprathreshold non-linear performance of HVS [25,26].
4. Application of non-linear functions (logarithmic or square root functions) to the amplitude spectrum after the linear stage [18–20,22]. The response to the amplitude of the input, the sensitivity and the resolution step size of any detector are closely related. The Weber law states that the HVS has a logarithmic varying resolution step size for luminances (zero spatial frequencies). This implies high sensitivity for low luminances and lower sensitivity for higher luminances. This uneven amplitude sensitivity is also qualitatively true for the non-zero frequencies [25], but the actual resolution step size for amplitudes is not logarithmic [26]. The aim of the post-filtering non-linear functions applied to the spectrum is to emphasise the differences between the original and the distorted spectra in the perceptually significant areas (favouring low amplitudes). However, these logarithmic or square-root corrections are just heuristic extensions of the Weber law to the non-zero frequencies.

In this work we develop a frequency and contrast dependent metric in the DCT domain using a fully non-linear and suprathreshold perception model: the *Information Allocation Function* (IAF) of the visual system [26,27]. This approach qualitatively improves the previous perception models used for subjective distortion measurement because, due to the nature of its experimental foundations, it is not a stage-after-stage sequential model made of disconnected characteristics of the HVS, but it includes the effects from photoreceptors to post-transform suprathreshold non-linearities. The bit allocation model has been derived from experimental data about frequency and contrast incremental thresholds of HVS, and it is consistent with the reported noise adaptation of the CSF [23]. The IAF has been recently employed to guide the design of a perceptually adapted vector quantizer giving better subjective performance that JPEG standard at the same bit rate [27]. In this work, experimental results are presented that show how the distortions predicted by the IAF metric are linearly related to the opinion of the observers. We also compare the performance of our metric with other distortion measures partially based on processes 1 to 4. The experimental comparison confirms that our model deals with a wider range of distortions more accurately than these previously reported metrics.

2. Modelling bit allocation properties of the human visual system: The IAF

The existence of tolerances to changes in contrast of gratings or DCT base functions implies that the visual system maps the continuous contrast range into a finite set of discrete perceptions. Such a discrete encoding implies that perception can be considered as a redundancy removal process analogous to vector quantization. These considerations gave rise to a novel model of human visual performance focused in its bit allocation properties [26,27]. The Information Allocation Function (IAF) gives the amount of information assigned by the system to encode each area of the frequency-contrast domain. A inversely proportional dependence with the size of the tolerances of HVS to changes in that domain was proposed for this function [26,27]:

$$\mathrm{IAF}(f,C)=\frac{\mathrm{d}^2 I}{\mathrm{d}f\,\mathrm{d}C}=\frac{K}{\Delta f(f,c)\,\Delta C(f,C)}, \tag{1}$$

where $\Delta f(f,C)$ and $\Delta C(f,C)$ are the experimental frequency and contrast incremental thresholds of gratings in each point of the frequency-contrast domain. The experimental procedure to determine these incremental thresholds consists of measuring the minimum amount of frequency or amplitude deviation needed to make just discriminable two successively presented sinusoidal patterns of a given frequency and contrast [26]. The results obtained by such a technique

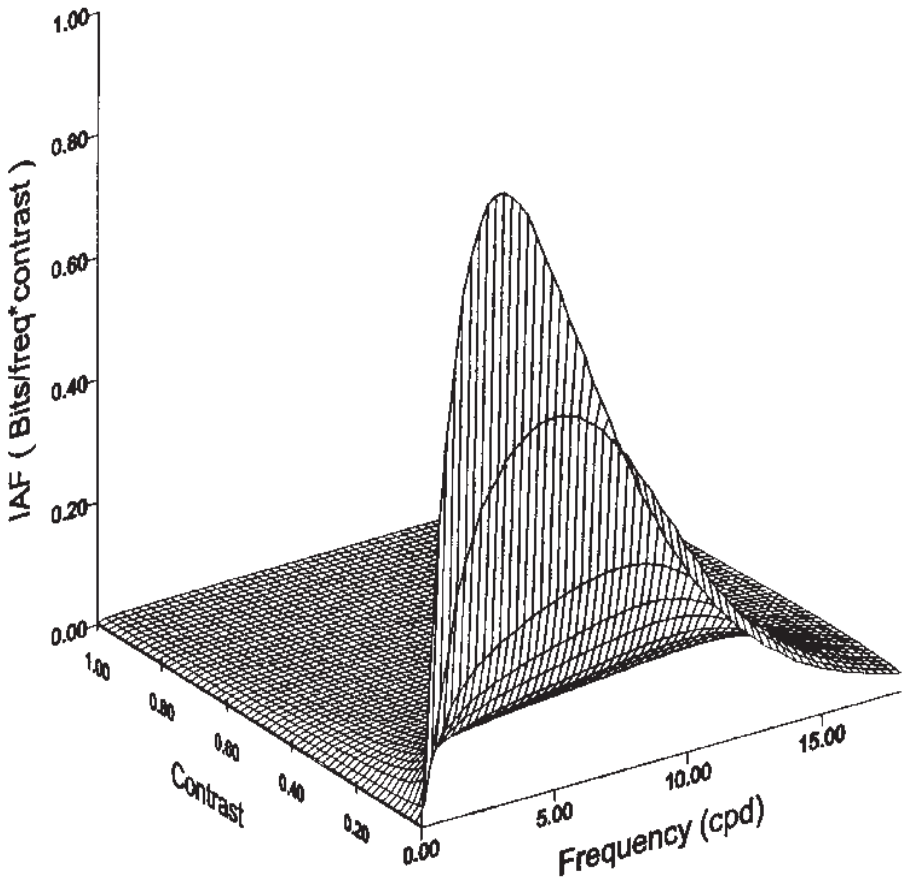


Fig. 1. Relative bit allocation of the visual system in the frequency-contrast DCT domain.

do not include spatial domain masking effects, however, it must be stressed that they do include all the possible non-linearities in the spatial and the frequency domains because they are not obtained from some intermediate stage but from the whole response of the HVS. The explicit expression for the IAF is:

$$\mathrm{IAF}(f,C) = \frac{K}{C_T(f) + C\left(\left(\frac{C_T(f)}{k(f)}\right)^{1/n(f)} + C\right)^{-1}\left(k(f)C^{n(f)} - C_T(f)\right)}, \tag{2}$$

where $C_T(f)$ is the absolute contrast detection threshold for a grating of frequency f (the inverse of the CSF(f)) and $k(f)$ and $n(f)$ are the empirical variation of the parameters of $\Delta C(f,C)$ [26]:

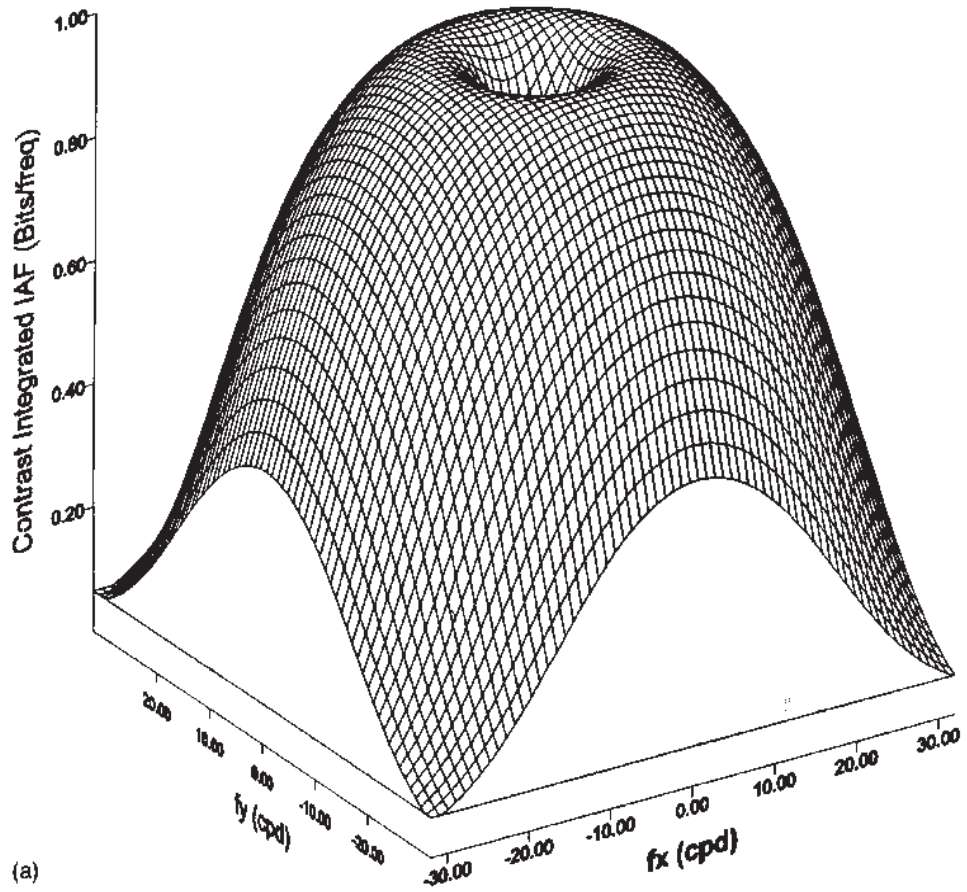
$$k(f) = -0.079389 \log_{10}(f) + 0.322725, \tag{3a}$$

$$n(f) = 0.84 \frac{f^{1.7}}{0.54534 + f^{1.7}}. \tag{3b}$$

This bit allocation model obtained from the experimental data implies higher resolution of the visual system for middle frequencies and the low contrast range (see Fig. 1), in agreement with classical perception models [24,25]. Nevertheless, quantitative comparison reveals that the IAF has a wider effective band-pass than the threshold CSF filter [24] (see Fig. 2), and that suprathreshold non-linearities are no longer a simple logarithmic or square-root transduction for each non-zero frequency. Note that considering the suprathreshold characteristics of HVS (contrast dependent term in the denominator of eqn (2)) substantially modifies the relative importance of each frequency coefficient. In other words, if one neglects suprathreshold characteristics of HVS, one gets the threshold CSF, and substantial disagreement with the opinion of the observers will be obtained. The shape of the IAF is consistent with the reported noise level dependence of the CSF [23].

3. Measuring subjective differences with the IAF

The subjective distortion measure presented in this paper is based on the fact that the amount of information assigned by the system to encode each area of the frequency-contrast DCT domain can be a good measure of the relative importance given by the system to the differences between the encoded signals in that area. The basic idea is: as more information is allocated by the HVS in one area, more visual importance is given to impairments in that area. An IAF-like frequency-contrast sensitivity surface has been successfully employed in the design of perceptually adapted lossy compression algorithms [23,27]. This fact supports the idea that the IAF could be an accurate weight to measure the contributions of differences in the DCT domain to the global distance between images.



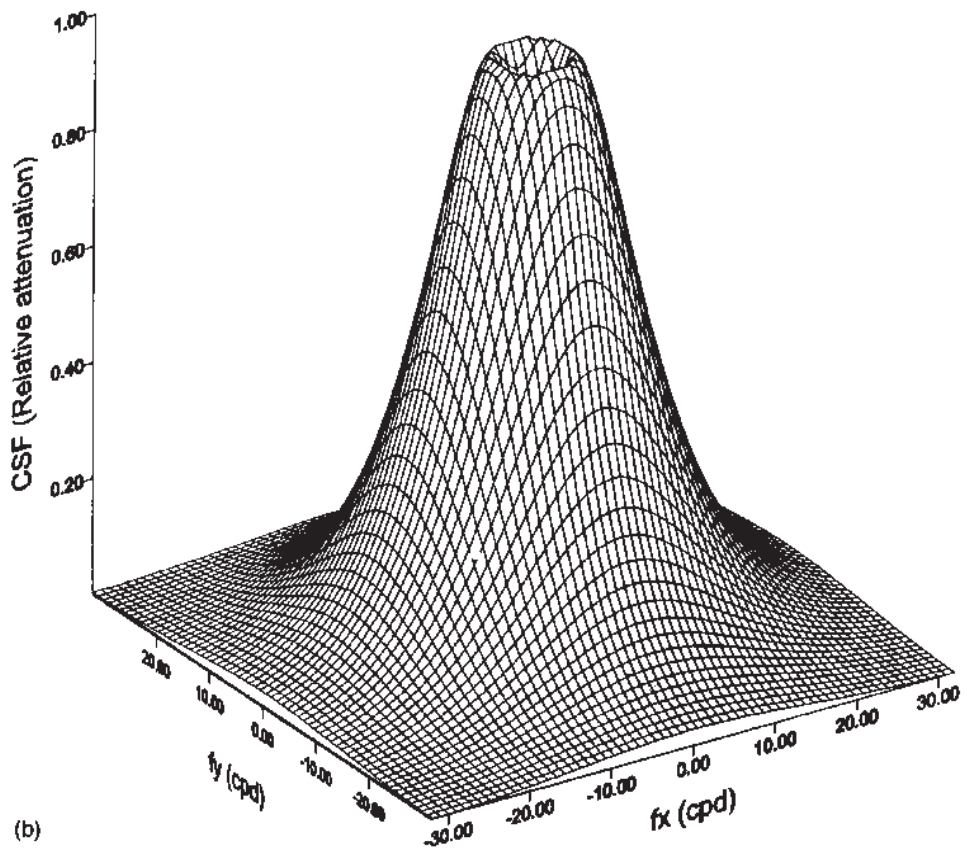


Fig. 2. A comparison between the contrast integrated IAF: (a) relative bit allocation of the visual system in the frequency domain, and (b) the CSF: the threshold sensitivity of the HVS in the frequency domain.

According to this, the measurement of the subjective difference between an original image i_1 and its distorted version i_2 can be done weighting the differences between the perceived images in the DCT domain with the IAF in that domain:

$$d(i_1,i_2)^2=\int \mathrm{IAF}(f,C)(\mathrm{QDCT}[i_1]-\mathrm{QDCT}[i_2])^2\,df, \quad (4)$$

where QDCT stands for the quantized DCT, f is the spatial frequency of each coefficient and C is the contrast for amplitude $(\mathrm{QDCT}[i_1]-\mathrm{QDCT}[i_2](f))/2$. Notice that in this model, perception is considered as a variable-precision quantization process given by the IAF, according to the method exposed in [13]. In this way, the integral (4) can be interpreted as the sum of the differences between the perceived images assuming the IAF model of redundancy removal.

4. Other error measurements

In this section we briefly outline some of the distortion measures reported in the literature [3,4,16–21] which are representative of the metrics that use the visual perception characteristics described in the introduction. We will use these algorithms to emphasise the quantitative improvement obtained from the IAF metric.

(1) Standard objective energy based metric: The mean square error (MSE):

$$d(i_1,i_2)^2=\int (i_1(x)-i_2(x))^2\,dx. \quad (5)$$

In spite of its well known disagreement with subjective assessment, standard MSE has been included here in order to analyse the nature of the distortions of the test images, and to see how the different metrics improve the straightforward result given by the MSE.

(2) Modified MSE including photoreceptors cube-root non-linear response [9]:

$$d(i_1,i_2)^2=\int \left(i_1(x)^{1/3}-i_2(x)^{1/3}\right)^2 dx. \quad (6)$$

(3) Masking based metric I. This metric assumes that perception can be characterised by a simple blurring process (local averaging) with two additional features: there is a facilitation of error detection in sharp edge zones, and there is masking of errors in highly textured areas [3,4]. According to this, the considered metric computes the differences between the locally averaged images, weighting each difference by a factor directly proportional to the gradient of the image and inversely proportional to the local variance:

$$d(i_1,i_2)^2=\int \left((M[i_1(x)]-M[i_2(x)])\frac{|\nabla i_1(x)|}{1+2\,\mathrm{Var}[i_1(x)]}\right)^2 dx, \quad (7a)$$

where $M[i(x)]$ is the mean luminance in a neighbourhood, N, of point x:

$$M[i(x)]=\int_N i(x)\,dx \quad (7b)$$

and $\mathrm{Var}[i(x)]$ is the local variance evaluated in the neighbourhood of x:

$$\mathrm{Var}[i(x)]=\int_N (i(x)-M[i(x)])^2\,dx. \quad (7c)$$

(4) Masking based metric II. This error measure takes into account the reduction of sensitivity in high luminance areas weighting the image differences by the inverse of the local luminance [4]:

$$d(i_1,i_2)^2=\int \left(\frac{(i_1(x)-i_2(x))}{M[i_1(x)]}\right)^2 dx. \quad (8)$$

(5) Linear filtering based metric [16,18]. Linear filtering based models of perception assume that the perceived image is a band pass filtered version of the original image. The characteristic HVS filter is the threshold frequency response, the CSF. Assuming this, the perceptual difference should be computed from the filtered images, weighting the error spectra by the frequency response of the system:

$$d(i_1,i_2)^2=\int (\mathrm{FT}[i_1(x)](f)-\mathrm{FT}[i_2(x)](f))^2\mathrm{CSF}(f)^2\,df. \quad (9)$$

(6) Non-linear photoreceptors response plus linear threshold filtering [17,21]:

$$d(i_1,i_2)^2=\int \left(\mathrm{FT}\left[i_1(x)^{1/3}\right](f)-\mathrm{FT}\left[i_2(x)^{1/3}\right](f)\right)^2\mathrm{CSF}(f)^2\,df. \quad (10)$$

(7) Non-uniform weighting of spectrum differences: Subjective Quality Factor (SQF):

$$d(i_1,i_2)=\int_{f_1}^{f_2}\frac{|\mathrm{FT}[i_1(x)](f)-\mathrm{FT}[i_2(x)](f)|}{f}\,df. \quad (11)$$

(8) Non-uniform weighting of spectrum differences plus linear threshold filtering and nonlinear contrast response: Square Root Integral (SQRI) [18,19]

$$d(i_1,i_2)=\int_{f_1}^{f_2}\frac{((\mathrm{FT}[i_1(x)](f)-\mathrm{FT}[i_2(x)](f))\mathrm{CSF}(f))^{1/2}}{f}\,df. \quad (12)$$

5. Experiments

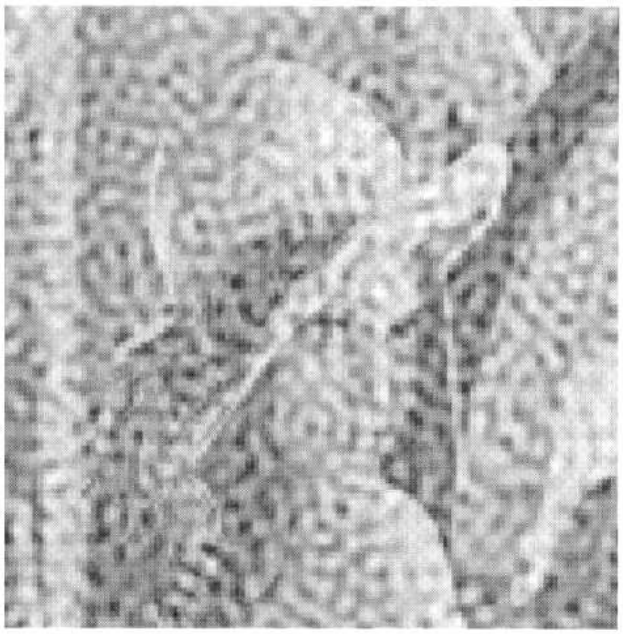
To test the reliability of the proposed metric and compare its performance with the results of the other described metrics under a variety of distortions, two different standard psychophysical experiments [21,28] have

(a)

(b)

(c)

(d)



(e)

(f)

(g)

(h)

(i)

(j)

(k)

(l)

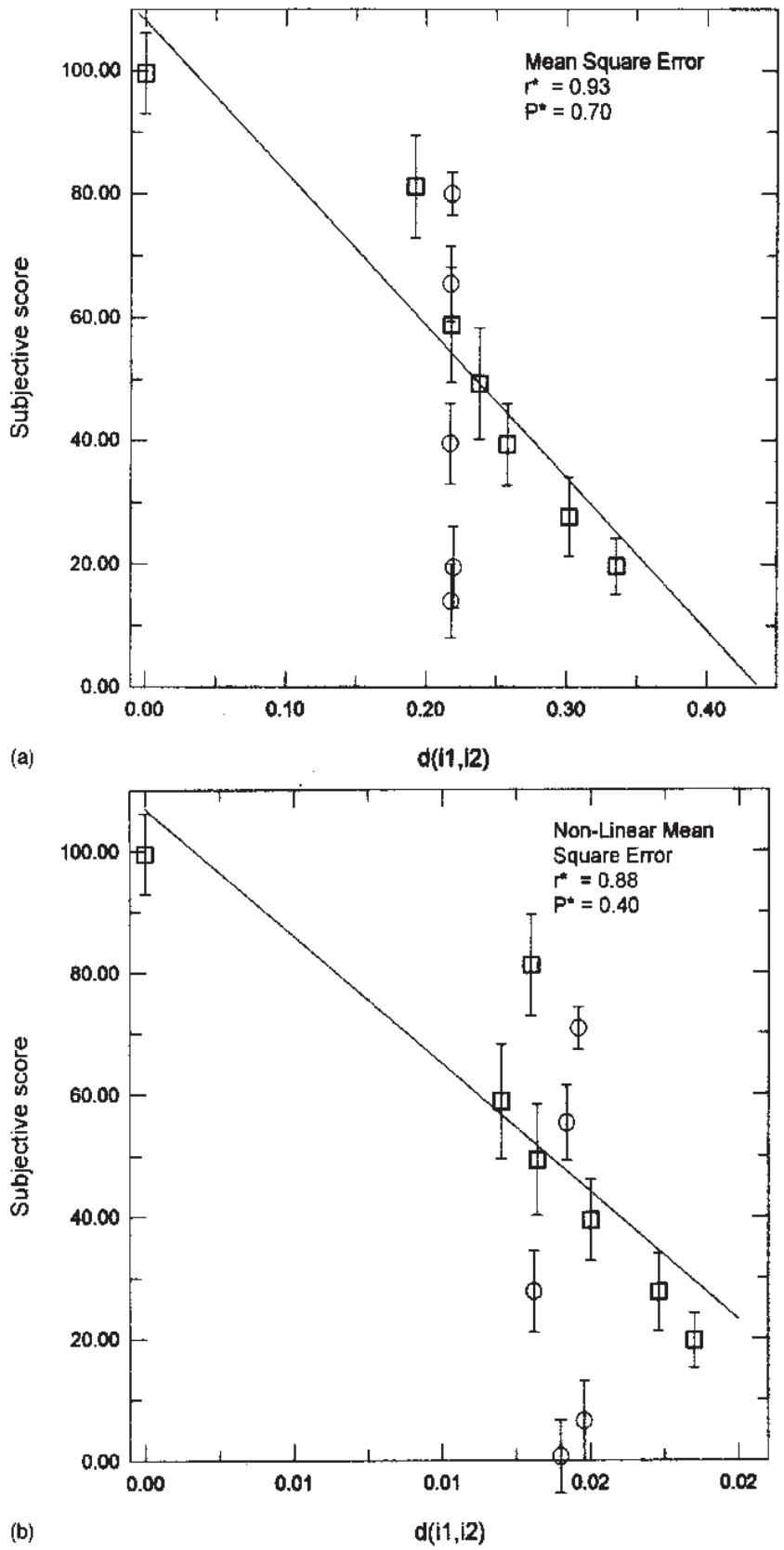
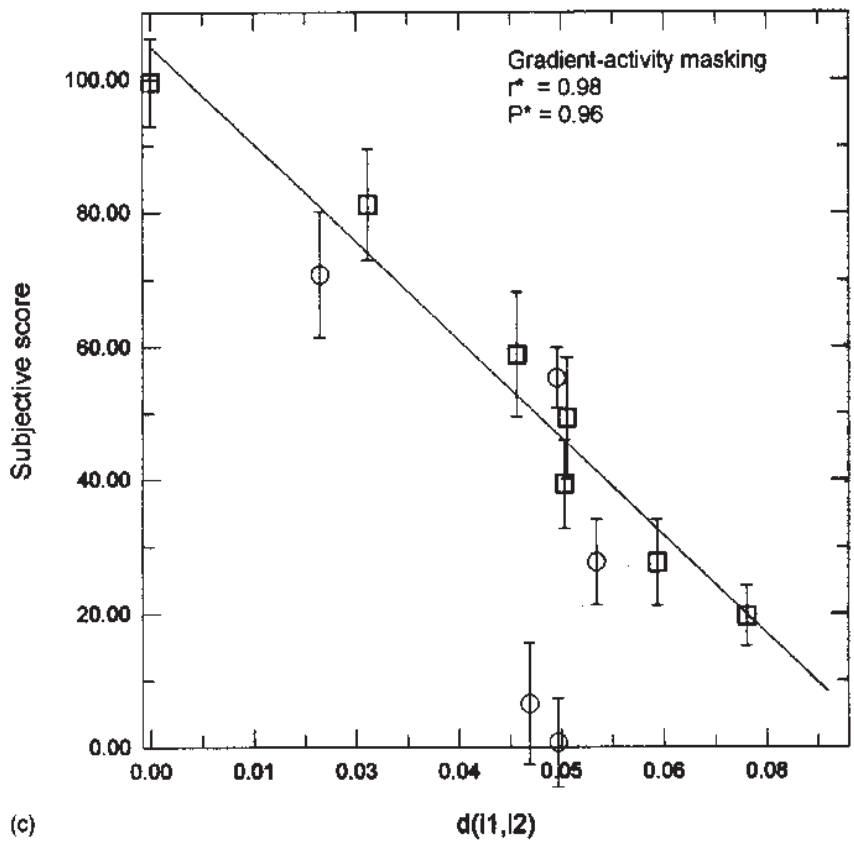
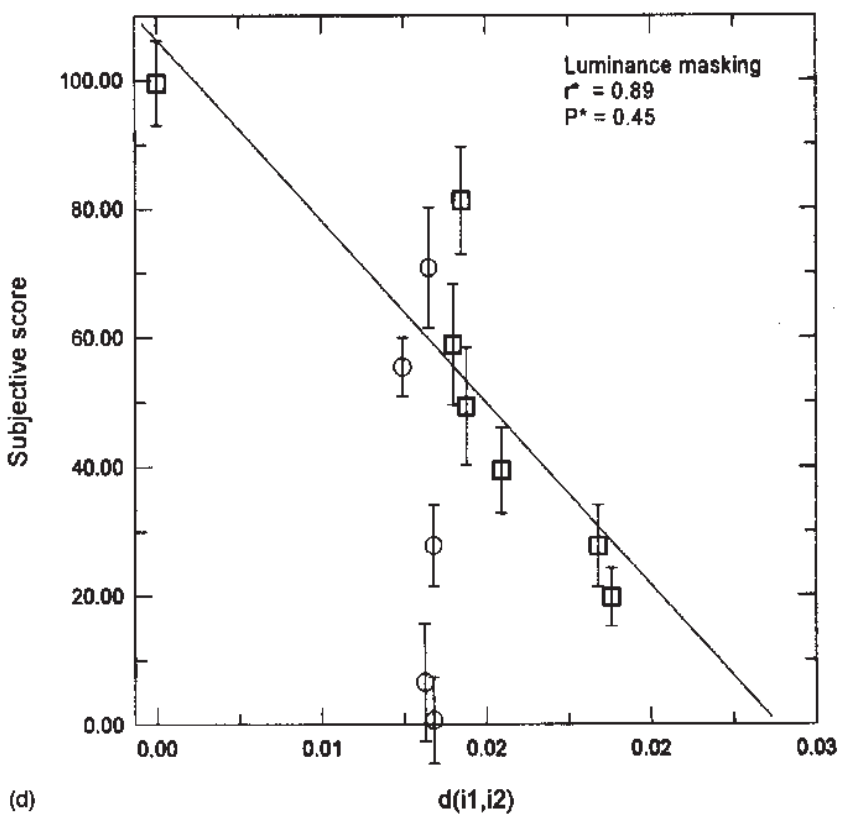


Fig. 4. Observers score vs. distortion computed by the different algorithms in experiment I (circles) and experiment II (squares). The correlation coefficient (r) and the probability of the χ^2 test (P) of the fit are given inside each figure. When they are marked with an asterisk (r^*,P^*), it means that the straight line only fits the data of experiment II. In those cases, the metric can not hold the coloured noise data (see the corresponding r and P-values for the experiment I and both experiments in Table 1).

Fig. 3. Examples of the test images used in the experiments. Experiment I, images (a–f): Original image corrupted by coloured noise of the bands [0,1] cl/deg (a); [1,2] cl/deg (b); [2,4] cl/deg (c); [4,8] cl/deg (d); [8,16] cl/deg (e); and [16,32] cl/deg (f). The signal to noise energy ratio is 100:5. Experiment II, images (g–l): JPEG-compressed images at different bit rates: 0.80 bits/pix (g); 0.38 bits/pix (h); 0.29 bits/pix (i); 0.18 bits/pix (g); 0.13 bits/pix (k); 0.08 bits/pix (l).



(c)

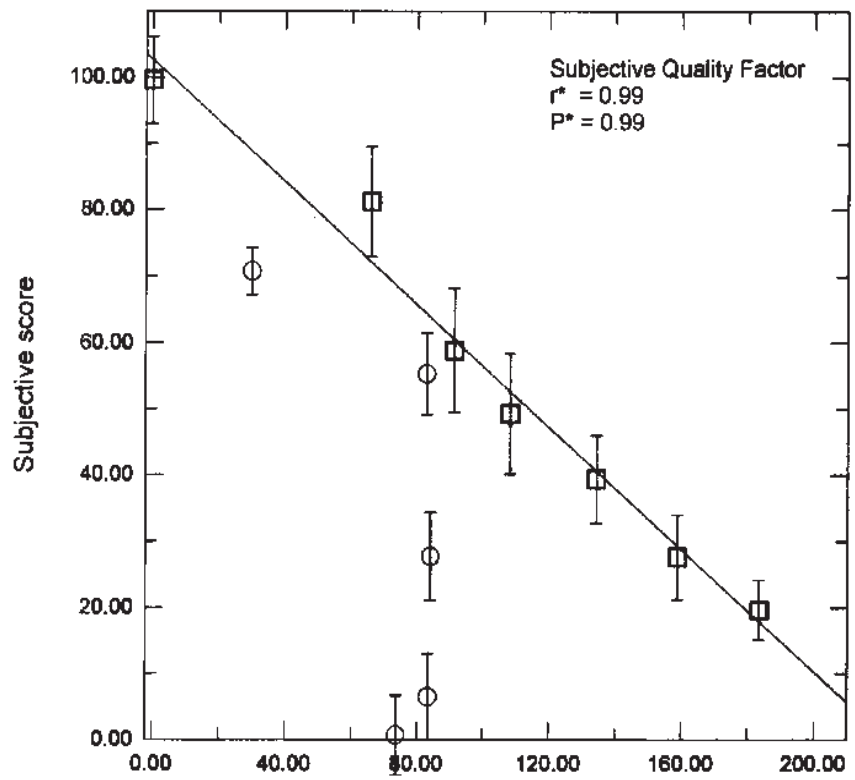


(d)

Fig. 4. Continued.

been carried out:

- Experiment I: The task of the observers was ordering a set of degraded images with constant energy noise but different spectrum, (coloured noise), so that an equal MSE would be obtained. The coloured noise was obtained filtering a white-noise random signal with a square filter of the proper band-pass. In experiment I the tests were distorted versions of a natural image



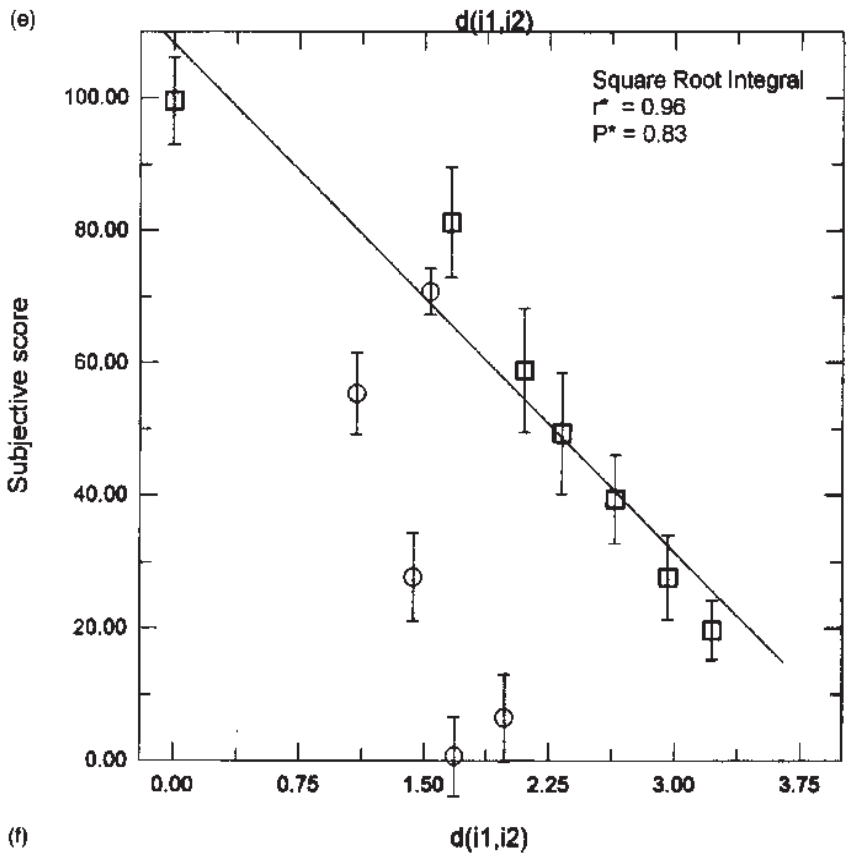
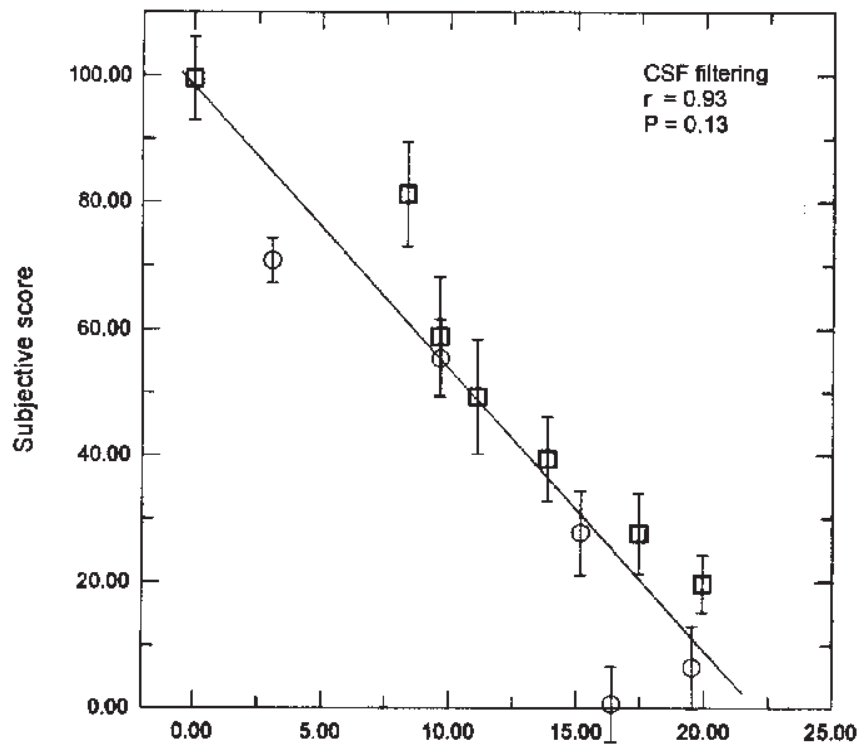


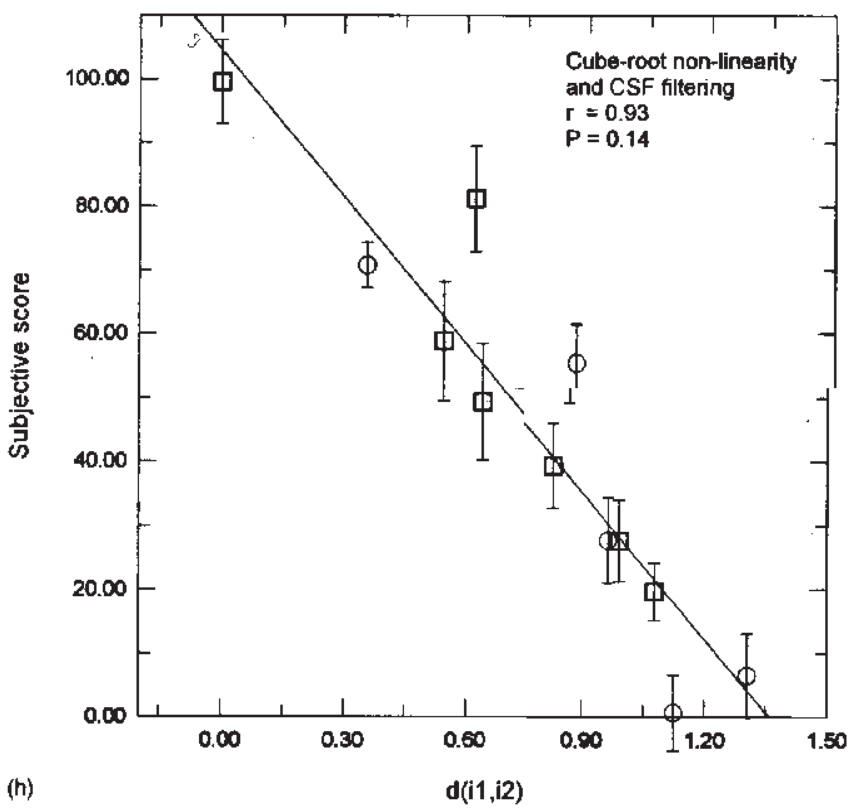
Fig. 4. Continued.

with coloured noise of the bands: [0,1], [1,2], [2,4], [4,8] and [8,16] cycles per degree.

- Experiment II: The task of the observers was assigning a numerical score to the distortion of JPEG compressed images at different bit rates and with non-linear mean square error. To avoid blocking effect distortions, no subblock division was considered: the quantization matrices were applied to the full-image transform. The objective of using this kind of test images is not to evaluate a particular image compression standard, but



(g)



(h)

Fig. 4. Continued.

to generate degraded images with smooth frequency shaped noise of different energies. In experiment II the tests were JPEG-coded versions of natural images to the bit rates: 0.08, 0.13, 0.18, 0.29, 0.38, 0.80 and 5.62 bits/pix.

Examples of the employed test images used in the experiments are shown in Fig. 3. Double stimulus continuous quality scales with adjectival descriptors were used to record the quality judgements [28]. In both cases, the opinions of the observers were normalised to fall in the range [0,100]. Experiments I and II were carried out by 33 and 28 different observers, respectively. 8-bit gray-scale

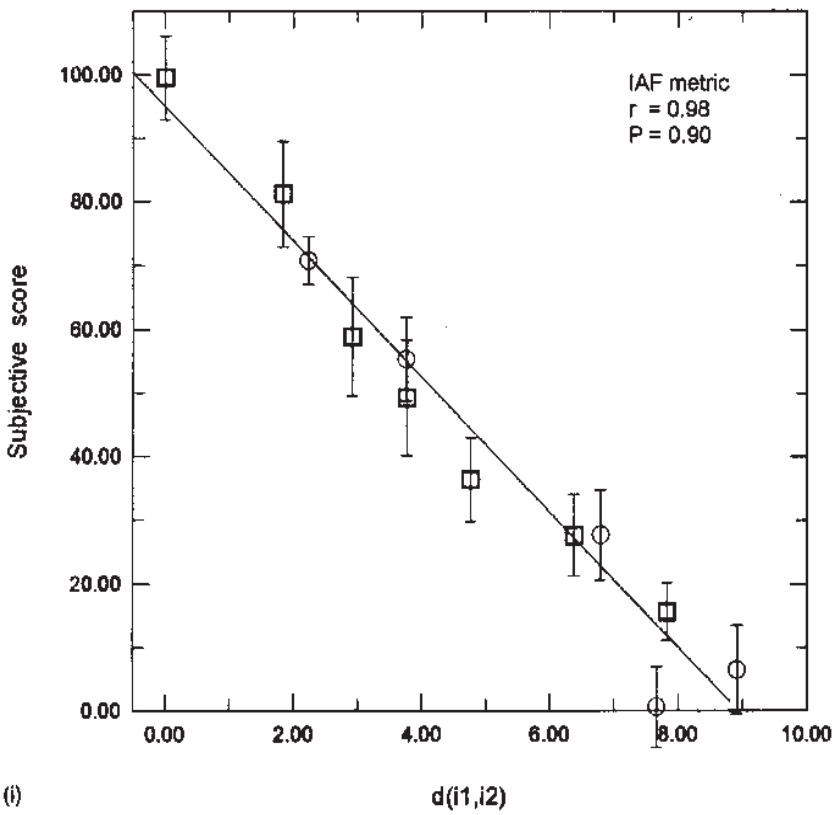


Fig. 4. Continued.

images were presented in a 21″ gamma corrected colour monitor (Mitsubishi Diamond Pro 20) subtending 3.8° of visual angle at a viewing distance of 1.5 m. Mean luminance was fixed to be 50 cd m^{-2}. The display was viewed binocularly under fluorescent room lights.

6. Results and discussion

In this section we analyse the performance of our metric and other proposed metrics representing the subjective quality score versus the distance computed with the different algorithms. As experiment I and II are different examples of the same general assessment process, they should follow the same curve when plotted versus algorithms results. In other words, a good metric should give comparable distance values – distance values expressed in the same scale – in spite of the different nature of the evaluated distortions. The free parameters of each model have been fixed as indicated by their respective authors.

Fig. 4 displays the quality assessment of the observers plotted versus the numerical results given by eqns (4)–(12) in experiment I (circles) and experiment II (squares). In Table 1, the goodness of the metrics to match the observers opinion in both experiments is analysed. The correlation coefficient (r), and the value of the χ^2 test for each experiment and metric are given. The χ^2 parameter is the probability that χ^2 is larger or equal to the sum of squared errors ($P(\chi^2 \geq \varepsilon^2)$) In Table 1 the results are given for each isolated experiment and for the whole jointly considered

Table 1
Correlation coefficient, r, and χ^2 parameter, $P(\chi^2 \geq \varepsilon^2)$, for each experiment and metric; each isolated experiment and the whole data are considered; a good metric should be well-behaved in any case

	Experiment I		Experiment II		Both experiments	
	r	P	r	P	r	P
Mean square error (MSE)	–	–	0.93	0.70	0.67	1e − 19
Non-linear MSE	0.13	1e − 18	0.88	0.40	0.68	2e − 17
Gradient-activity masking	0.64	8e − 8	0.98	0.96	0.80	1e − 6
Luminance masking	0.41	3e − 15	0.89	0.45	0.59	2e − 14
Subjective quality factor (SQF)	0.34	1e − 16	0.99	0.99	0.57	6e − 13
Square root integral (SQRI)	0.65	3e − 10	0.96	0.83	0.60	8e − 8
CSF filtering	0.94	0.07	0.97	0.95	0.93	0.13
Non-linearity and CSF filtering	0.91	0.01	0.93	0.72	0.93	0.14
IAF metric	0.97	0.28	0.98	0.96	0.98	0.90

data. A good metric should be well-behaved in any case. P-values below the range [0.01,0.1] indicate that the model does not handle the data [29]. In each figure, the values of r, and P of the fit are given. When the correlation coefficient and the probability of the χ^2 test are marked with an asterisk, (r^*,P^*), it means that the straight line only fits the data of experiment II. In such a case, the metric cannot reproduce the coloured noise data.

Fig. 4 and the confidence data of Table 1 show that some metrics completely fail to predict the results of experiment I. As can be seen in Fig. 3, in both experiments the quality of the distorted test images ranges from very bad to good quality, so if a particular metric cannot deal with the data of a particular experiment, it is due to a lack of generality of its underlying model. The values obtained in the χ^2 confidence measurement may seem too high in some cases, but, as it is well known, these values not only depend on the goodness of the model to match the data, but on the measurement errors too [29]. High variance measurements decreases the sum of squared errors, ε^2, increasing the probability of χ^2 to be larger than ε^2. In this case, high variance is due to the psychophysical nature of the experiments. In this kind of experiments the variance cannot be decreased below certain limits because the observers responses are intrinsically disperse and outliers always occur. This fact may increase the reliability lower limit of P, but the relative ranking remains unchanged. In this case, the experimental data (and their variances) are the same for all the considered models, so they are comparable using the χ^2 parameter.

As we can see in Fig. 4(a), in spite of the clear non-linearity of the results, the MSE fits to a reasonably accuracy the data of the experiment II (squares), but, obviously, it fails to predict the subjective score for the images corrupted by equal energy coloured noise (circles).

Fig. 4(b) shows that if we only include the photoreceptors non-linearities, differences between test images of experiment I are introduced, but the results are not yet valid to handle the subjective opinion in this case.

Fig. 4(c) and 4(d) show the still poor results of the spatial-domain masking based metrics. Although the results of the gradient-activity metric are really good for experiment II, the goodness of the fit for all the data is below the acceptability limit ($P(\chi^2 \geq \varepsilon^2) = 1{\cdot}10^{-6} \ll 0.1$) in spite of the relative good alignment of the data $r = 0.80$. In Fig. 4(e) and 4(f) appear the results of $1/f$ weighted differences in the frequency domain: SQF and SQRI. These metrics achieve very good results when fitting the experiment II data, but they fail again to adjust the results of the experiment I. In the case of Square Root Integral metric, good agreement with subjective quality assessment has been reported in noisy imagery [19], but those results always assume white noise.

The best results are obtained by the CSF-based frequency methods and our IAF metric (Fig. 4(g), 4(h) and 4(i)). These metrics can deal with the results of both experiments giving acceptable fits ($P(\chi^2 \geq \varepsilon^2) > 0.1$) for the whole set of data.

The results show that the equal-energy coloured noise assessment can be reproduced only if the uneven spatial frequency sensitivity of HVS is taken into account (by means of the CSF or the IAF). The SQRI metric do employ the CSF but its effect is perturbed by the $1/f$ factor, overestimating the low frequency contributions to the perceptual error. An accurate prediction of the assessment of frequency dependent or band localised noise is not an academic problem attending to the wide use of the transform domain encoding techniques for still images and video signals. In such techniques the quantization noise is unevenly distributed in the spatial frequency domain according to *statistical* and *perceptual* criteria [6,7,12], so it is very important to know the perceptual effect of selective error allocation.

Fig. 4(g), 4(h) and 4(i) and the quantitative results of the χ^2 test show that there is little benefit from introducing first-stage photoreceptors non-linearities, and that there is a significant improvement in using the fully non-linear model based on bit allocation of HVS.

7. Concluding remarks

In this paper, a new subjective image fidelity metric in the DCT domain has been presented. It is based on an unified fully non-linear and suprathreshold contrast perception model: the Information Allocation Function (IAF) of the visual system. Due to its global empirical nature, this model includes the effects form photoreceptors point-wise non-linearities to post-transform suprathreshold non-linearities.

Exhaustive psychophysical experiments show that the distortion obtained by the IAF-based algorithm is linearly related to the score given by the observers under a variety of noise conditions. The quality of the fit is better in the IAF model than in previously reported metrics. The developed metric can, therefore, be incorporated in the design of compression algorithms as a closer approximation of human assessment of image quality.

Acknowledgements

This work was supported by Generalitat Valenciana/ Diputació de València IVEI scientific research project No. 96/003-035. The authors wish to acknowledge useful discussions with J. García, F. Ferri, J. Albert and P. Capilla.

References

[1] A. Gersho, R.M. Gray, Vector Quantization and Signal Compression, Kluwer, Dordrecht, The Netherlands, 1992.
[2] Y. Linde, A. Buzo, R.M. Gray, An algorithm for vector quantizer design, IEEE Trans. Comm. 28 (1980) 84–95.
[3] J.O. Limb, Distortion criteria of the human viewer, IEEE Trans. on Sys. Man and Cybern. 9 (12) (1979).
[4] H. Marmolin, Subjective MSE measurements, IEEE Trans. on Sys. Man, Cybern. 16 (3) (1986) 486–489.

[5] N. Jayant, J. Johnston, R. Safranek, Signal compression based on models of human perception, in: Proc. IEEE 81 (10) (1993) 1385–1422.
[6] G.K. Wallace, The JPEG still picture compression standard, Comm. ACM. 34 (4) (1991) 31–43.
[7] D. LeGall, MPEG: A video compression standard for multimedia applications, Comm. ACM 34 (4) (1991) 47–58.
[8] ISO/IEC, 13818 draft international standard: Generic coding of moving pictures and associated audio, part 2: Video, 1993.
[9] A.N. Akansu, R.A. Haddad, Multiresolution Signal Decomposition: Transforms, Subbands and Wavelets, Academic Press, San Diego, 1992, pp. 1–7.
[10] B. Macq, Weighted optimum bit allocations to ortogonal transforms for picture coding, IEEE Trans. 10 (1992) 875–883.
[11] K.N. Ngan, H.C. Koh, W.C. Wong, Hybrid image coding scheme incorporating human visual system characteristics. Opt. Eng. 30 (7) (1991) 940–947.
[12] B. Macq, H.Q. Shi, Perceptually weighted vector quantization in the DCT domain, Electr. Lett. 29 (15) (1993) 1382–1384.
[13] J.L. Barron, D.J. Fleet, S.S. Bauchemin, Performance of optical flow techniques. Int. J. Comp. Vis. 12 (1) (1994) 43–77.
[14] S. Haykin, Adaptive Filter Theory, 3rd ed., Prentice Hall, New Jersey, 1996.
[15] I. Pitas, A.N. Venetsanopoulos, Non-linear Digital Filters: Principles and Applications, Ch. 9: Adaptive non-linear filters, Kluwer, Dordrecht, The Netherlands, 1990, pp. 267–311.
[16] N.B. Nill, BR. Bouzas, Objective image quality measure derived from digital image power spectra, Opt. Eng. 32 (4) (1992) 813–825.
[17] LA. Saghri, P.S. Cheatham, A. Habibi, Image quality measure based on a human visual system model, Opt. Eng. 28 (7) (1989) 813–819.
[18] P.J.G. Barthen, Evaluation of subjective image quality with the square-root integral method, JOSA A. 7 (10) (1990) 2024–2031.
[19] P.J.G. Barthen, Evaluation of the effect of noise on subjective image quality, in: Proc. of the SPIE 1453, Human Vision, Visual Processing and Digital Display II, 1991.
[20] D.J. Granrath, The role of human visual models in image processing, Proc. IEEE 69 (5) (1981) 552–561.
[21] D.R. Fuhrmann, J.A. Baro, J.R. Cox, Experimental evaluation of psychopysical distortion metrics for JPEG-encoded images, Proc. of the SPIE 1913 (1993) 179–190.
[22] W.K. Pratt, Digital Image Processing, John Wiley, New York, 1992.
[23] S. Daly, Application of a noise-adaptive Contrast Sensitivity Function to image data compression, Opt. Eng. 29 (8) (1990) 977–987.
[24] D.H. Kelly, Receptive-field-like functions inferred from large-area psychophysical measurements, Vision Res. 25 (12) (1985) 1895–1900.
[25] G.E. Legge, A power law for contrast discrimination, Vision Res. 21 (1981) 457–467.
[26] J. Malo, A.M. Pons, J.M. Artigas, Bit allocation of the human visual system inferred from contrast incremental thresholds of sinusoidal gratings, Perception 24 (Supl.) (1995) p. 86.
[27] J. Malo, A.M. Pons, J.M. Artigas, Bit allocation algorithm for codebook design in vector quantization fully based on human visual system non-linearities for suprathreshold contrasts, Electr. Lett. 24 (1995) 1229–1231.
[28] R. Aldridge, J. Davidoff, M. Ghanbari, D. Hands, D. Pearson, Measurement of scene-dependent quality variations in digitally coded television pictures, IEE Proc. Vis. Image and Signal Proc. 142 (3) (1995) 149–154.
[29] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes in C, 2nd ed., Ch. 15: Modelling of data, Cambridge University Press, Cambridge, 1992, pp. 659–666.