

Non-linear Scale-Selection for Image Compression Improvement Obtained by Perceptual Distortion Criteria

Iván Kopilovic

University of Veszprém, Department of Image Processing and Neurocomputing,
H-8200 Veszprém, Egyetem u. 10., Hungary

Tamás Szirányi

University of Veszprém, Department of Image Processing and Neurocomputing and
Analogical and Neural Computing Laboratory, Comp. & Automation Inst., Hungarian Academy of
Sciences, H-1111 Budapest, Kende u. 13-17., Hungary

Abstract

The effect of linear and non-linear diffusions prior to Joint Photographic Experts Group's (JPEG) standard baseline compression is analysed and compared. We shall show that anisotropic diffusion acts as an adaptive filtering that extracts and preserves the important visual information, and it will be shown how it can be used as a preprocessing method for compression artefact reduction. The selection of the appropriate preprocessing parameters at a specified bit rate greatly reduces the compression artefacts. The selection criteria were determined using a perceptual error metric based on a model of the human visual system.

1. Introduction

Application of the non-linear scale-space theory to non-recoverable, low bit rate image compression will be presented. The method discussed was first proposed in [18] as an improvement on the JPEG baseline coding standard [13].

Compression enhancement methods range from simple preliminary smoothing to more sophisticated methods involving quantizer design [4] or postprocessing [10,15]. Smoothing entails the loss of edge information, at postprocessing the data is already corrupted. Other methods call for change in the compression standard.

In contrast to these methods, we have introduced a preprocessing step [18] performed before coding that will reduce the artefacts while avoiding the loss of information on "important" edges, and which involves no changes in the standard coding and decoding scheme. It is based on

adaptive filtering achieved by non-linear diffusion processing. A unified framework for a number of non-linear diffusion equations is the non-linear scale-space theory. Within this context we shall consider the most important diffusion equations found in literature.

Throughout this work, a perceptual error metric [9] was used as a quality criterion. Using this metric, we shall establish a selection rule for the parameters of the preprocessing to achieve the best compression quality for a given bit rate.

2. Scale-space and adaptive filtering

Perceptually adaptive filtering will mean suppression of noise and enhancement of edge information. In many applications, small details and textured regions do not influence the overall content of the image, and therefore they are considered to be noise at a given resolution. This assumption is reasonable, since at low bit rates textured regions facilitate the appearance of compression artefacts. An adaptive filter should remove the noise completely in uniform regions and perform directional filtering along the edges (shape enhancement), but it should exclude filtering across the edges.

A unified mathematical framework for these operators is the non-linear scale-space theory [7,11,14]. It represents the extraction of information as a series of operations. We can model this as a family $(T(t))_{t \geq 0}$ of operators acting on images. Each operator corresponds to a level of abstraction or the compactness of information representation. The greater the parameter t is, the more details will be omitted. The level of abstraction is called the scale of the representation, and the family $(T(t))_{t \geq 0}$ is called a scale-

space representation. Some natural assumptions can be made on these operators. These are simple properties that an information extraction operation is expected to verify: *translation invariance, non-creation of details, recursive structure, local (differential) nature.*

The fundamental result of the theory is that the certain class of scale-space representations verifying the above assumptions is generated by (non-linear) diffusion-like partial differential equations (PDEs), and conversely, a wide class of PDEs generates a scale-space representation [7,6,11]. Accordingly, the scale parameter of a scale-space representation corresponds to the “time” parameter of its generating PDE, and so the parameter t will be termed the “diffusion parameter” or the “scale” interchangeably. The most important PDEs in this respect that are found in the literature are listed in Table 1.

Parameter t is expressed in degrees of visual angle, but we shall express it in a dimension free form $t(t_0) = p^2 t_0$, where t_0 is the diffusion parameter and p is the pixels/degree value for a specific display setting. Parameter t obtained this way will be called the *relative diffusion*. For PMC-AD and NLI-AD the term $G_s * f$ (Table 1) was computed with linear diffusion, and therefore we shall call it the *pre-diffusion* [2] term with parameter s .

3. The perceptual error metric

The perceptual error metric found in [9] was used in the experiments to compare and evaluate the compression results, which is motivated by the well known fact that pixel-by-pixel based distortion measures fail to give correct assessment. The error metric accounts for several mechanisms of the human visual system. These are

basically the contrast sensitivity and the frequency/orientation pattern masking. For further reference on these phenomena and their application see [9] or [12] where some basic methods are reviewed.

The input of the algorithm is the viewing parameters (Table 3) and the two images to be compared. The reference image serves as a background for the error image. Original and the error image are decomposed into orientation and spatial frequency bands with Gabor-like filters, a model for the response of the cell in the visual cortex [5]. For each band the effect of masking and contrast sensitivity is computed to weigh the error coefficients. The error obtained will be called perceptually masked error, and we shall denote it by E . It can be expressed in a logarithmic form called the masked peak signal to noise ratio and is computed as $MPSNR(f, g) = 20 \log_{10}(1/E(f, g))$, where f and g are two images [9].

We shall also use a subjective image quality assessment scale shown in Table 2. It can be obtained with the formula $Q(f, g) = 5(1 + 159.5E(f, g))^{-1}$ [9].

We use the distortion measurement algorithm with parameter settings contained in Table 3.

5. The preprocessing algorithm

Most of the compression methods result in unrealistic, sometimes annoying artefacts at low bit rates. JPEG has a characteristic blocking-artefact (Figures 1,2,3 and 6).

PDEs of Table 1 were tested as perceptually adaptive filters for compression artefact reduction in [8]. The adequate perceptually adaptive filter fell out to be the PMC-AD, since it suppresses noise while performing shape enhancement. It preserves edge information and gives the best results in terms of perceptual error metric.

Table 1 Diffusion equations used in image processing. The general form of the equations is $\partial_t u(t, x) = F(u(t, \cdot))(x)$, where $t > 0$, $x \in \mathbf{R}^2$. The Gaussian kernel with parameter s is denoted by $G_s(x) = (4\pi s)^{-1} \exp(-(4s)^{-1} \langle x, x \rangle)$, x in \mathbf{R}^2 .

<i>Linear Diffusion (LD)</i>	$F(f) = \text{div}(\text{grad } f)$	Isotropic, linear, non-adaptive, equivalent to low-pass Gaussian filtering, $T_f(t)f = u_f(t, \cdot) = G_t * f$. See [11].
<i>Perona – Malik – Catté Anisotropic Diffusion (PMC-AD)</i>	$F(f) = \text{div}(g \circ (\ \text{grad}(G_s * f)\) \text{grad } f)$ $g = \exp\left(-\ K^{-1} \text{id}_{\mathbf{R}^2}\ ^4\right)$, $K > 0$.	Anisotropic, non-linear, edge-adaptive. Diffusion is anisotropically inhibited at edges by function g , depending on their strength (magnitude of the gradient). The role of G_s is to reduce noise sensitivity. Enhances the shape of objects, maintains the edge-information and suppresses noise. See [14] and [2].
<i>Non-linear Isotropic Diffusion (NLID)</i>	$F(f) = g \circ (\ \text{grad}(G_s * f)\) \text{div}(\text{grad } f)$, g is the same as above.	Non-linear, isotropic, edge-adaptive. Diffusion is inhibited isotropically by g and therefore the shape enhancement is poor. The role of G_s is the same as for PMC-AD. Simple implementation. Proposed in [17].
<i>Pure Anisotropic Diffusion (P-AD)</i>	$F(f) = \langle D^2 f \frac{\text{grad}^+ f}{\ \text{grad } f\ }, \frac{\text{grad}^+ f}{\ \text{grad } f\ } \rangle$, $\text{grad}^+ f = (-\partial_2 f, \partial_1 f)$	Non-linear, anisotropic, edge-adaptive. Diffusion is purely anisotropic, diffusion takes place only in the direction of edges and not across them. Useful for shape enhancement. Low noise suppression. See [7]. ($D^2 f$ denotes the second derivative of f).

Table 2 The subjective image quality assessment scale [3]. The corresponding masked error values are given.

Q	Impairment	Grade	MPSNR (dB)	Masked Error(E)
5	Imperceptible	Excellent	∞	0
4	Perceptible,not annoying	Good	56	$1.5 \cdot 10^{-3}$
3	Slightly Annoying	Fair	47.5	$4.3 \cdot 10^{-3}$
2	Annoying	Poor	40.5	$9.4 \cdot 10^{-3}$
1	Very Annoying	Bad	32	$2.5 \cdot 10^{-2}$

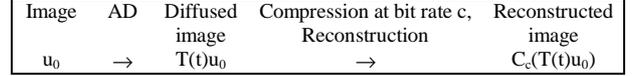
Table 3 Viewing configurations used for image distortion measurements.

	Computer Display (CD)	HDTV
Viewing distance (m)	0.5	1.8
Pixels/inch	72	48
Pixels/degree	24.74	60.32
Pixel size (deg)	0.04042	0.01658

Accordingly, we shall use AD instead of PMC-AD and $(T(t))_{t \geq 0}$ will denote the scale-space representation corresponding to it. Now, we shall determine the adequate parameterisation of the AD preprocessing.

We set two criteria for a good performing image quality enhancement. As first, it should not affect the original information content of the image. As second, it

should eliminate the formation of compression artefacts. These two constraints are contradictory and a compromise must be obtained. Let us keep in mind the following process:



Motivated by the above ideas, we introduce two characteristic features of a compression algorithm relative to a fixed adaptive filter $(T(t))_{t \geq 0}$, the *perceptual scale* and the *scale capacity*.

Perceptual scale of the compression at bit rate c is $t_p(c) = \sup\{t \geq 0 \mid E[C_c(T(t)u_0), u_0] < E[C_c(u_0), u_0] + \epsilon_0(c)\}$ where c is the bit rate, C_c is the compression/reconstruction operator at bit rate c , u_0 is the image, and $\epsilon_0(c)$ is a specified error tolerance. This refers to the maximal scale, i.e. the maximal loss of details, which remains imperceptible at the given bit rate. If the noise is removed adaptively with AD, the formation of artefacts is avoided, but certain information is lost as well. Perceptual scale is the scale at which AD begins to corrupt those details that belong to the essential visual content of the image, and the difference becomes “visible” for the perceptual distortion measuring algorithm. The error tolerance was 1%.

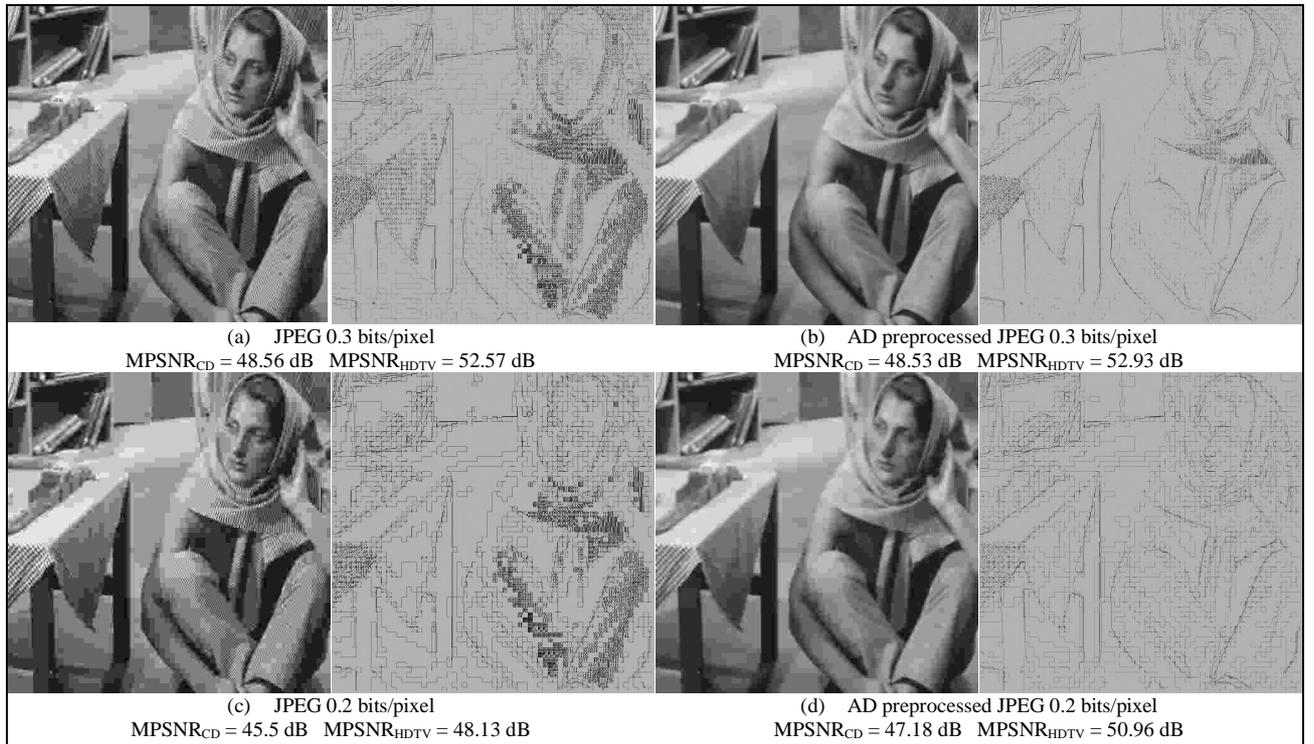
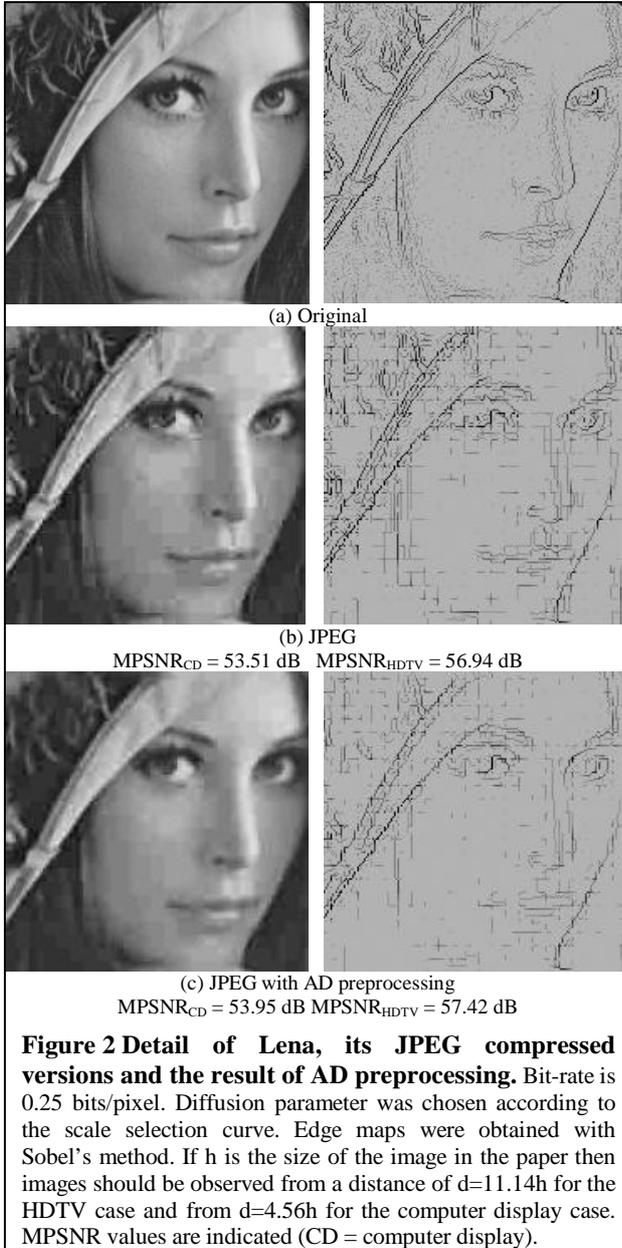


Figure 1 Compression results for test image Barbara and the corresponding edge maps. Diffusion parameter was chosen according to the scale selection curve. Edge maps were generated with Sobel’s method. If h is the size of the image in the paper then images should be observed from a distance of $d=6.75h$ for the HDTV case and from $d=2.768h$ for the computer display.



Scale capacity of the compression at bit rate c is defined as $t_x(c) = \inf\{t \geq 0 \mid E[C_c(T(t)u_o), T(t)u_o] < \varepsilon_1\}$, where C_c is the compression/reconstruction operator at bit rate c , u_o is the original image and ε_1 is a specified error level. Scale capacity is thus the smallest scale at which the compression algorithm operates below a specified error level on the preprocessed image. For images processed with smaller scale than $t_x(c)$, compression will give rise to undesirable error. At a fixed compression rate, there is a scale above which the

formation of artefacts is not significant any more. This scale is the scale capacity of the compression algorithm. For error level we have chosen the masked error equivalent for grade 4 of the CCIR scale ($E=1.5 \cdot 10^{-3}$) of Table 3.

According to these observations, we conclude that a reasonable scale-selection range for a bit rate c is $t_x(c) \leq t(c) \leq t_p(c)$. It remains to determine the values $t_x(c)$ and $t_p(c)$ for all bit rates c .

To determine this functional dependency, numerical experiments were performed. Three standard test images were considered: Barbara (Figure 1), Lena (Figure 2) and Golden hill (Figure 3). Each of the images were compressed to three different bit rates after diffusion with variable scale parameter, and the compression error was measured in terms of perceptually masked error (E). AD was run with fixed values of K and pre-diffusion parameter ($K=0.3, s=1.5$).

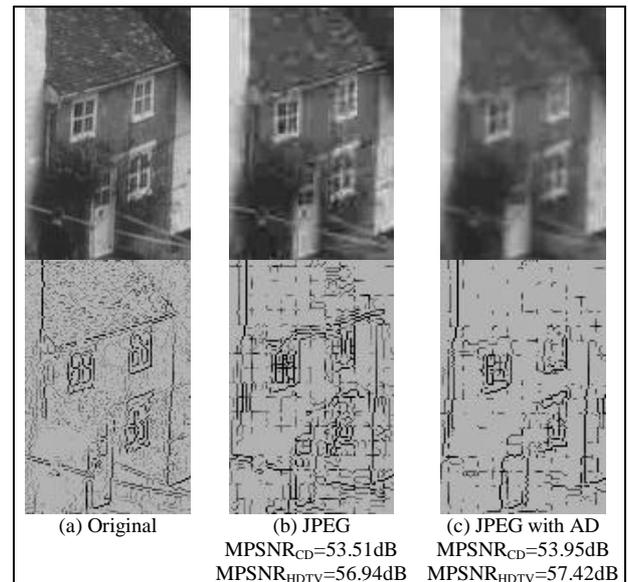


Figure 3 Detail of Goldhill, its JPEG compressed versions and the result of AD preprocessing. Bit-rate is 0.25 bits/pixel. Diffusion parameter was chosen according to the scale selection curve. Edge maps were obtained with Sobel's method. If h is the size of the image on the paper then images should be observed from a distance of d=10h for the HDTV case and from d=24h for the computer display case. MPSNR values are indicated (CD = computer display).

Error measurements were performed for two viewing conditions: HDTV and computer display (CD) (Table 3). Results for image Lena for the HDTV case is shown in Figure 4, column 1. Edge-maps of images were compared too (Figure 4, column 2).

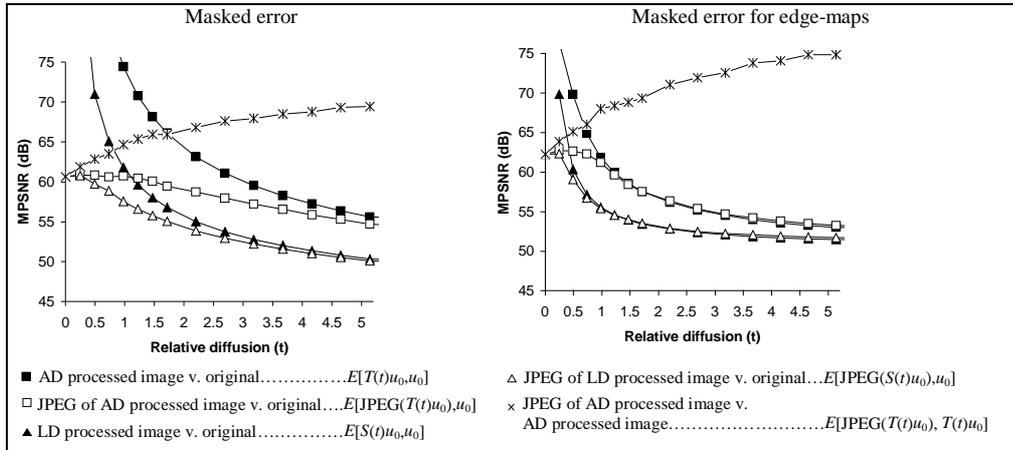


Figure 4 Perceptually masked errors (in terms of MPSNR) for test image Lena, and its edge-maps as a function of scale with HDTV configuration, at different bit rates. Evidently, linear diffusion (denoted by triangles) is outperformed by AD (denoted by squares). Results for edge-maps show that LD corrupts, while AD tends to preserve edges. Strong edges are very important constituents of the visual content of an image. The parameters of PMC-AD were the following: $K=0.1$, $n=4$, $s=0.35$. The compression bit-rate is 0.32 bits/pixel.

These results show clearly, that AD outperforms LD (smoothing) both in terms of perceptual error and edge-adaptivity.

By averaging results for the three images, overall tendency of the error was obtained, which was used to determine functions t_p and t_κ for computer display and HDTV configuration, and the scale-selection ranges lying between these two curves. Putting together the curves obtained for HDTV and computer display configuration, we were able to select a scale-selection function with values falling into both scale-selection ranges, in this way we obtained a scale-selection rule valid through a wider range of viewing conditions. In Figure 5 scale-selection ranges are shown, with their intersection (general scale-selection range) and the scale-selection function.

Snapshots of compression results for test images together with the edge-maps are shown through Figures 1, 2, 3 and 6. Note that the edge-maps of AD pre-processed images are “cleaner” and contain less noise.

The difference between LD and AD is, that in the latter case image information is filtered inadaptively, while in the former case it is performed in an adaptive way.

We have thus established a scale-selection rule for AD preprocessing, and in one hand, we have seen that AD gives significantly better results than a simple smoothing in the preprocessing step.

6. Conclusion

We have shown that a scale-space representation generated by an anisotropic diffusion can be used to enhance JPEG compression and prevent the formation of artefacts. Using perceptual distortion measurement

algorithm that simulates a human observers assessment, a scale-selection strategy was presented for different bit rates of compression. It was shown that AD is a better means of pre-processing than ordinary smoothing (LD) both with respect to perceptual error metric and edge-adaptivity.

The preprocessing method requires no change in the standard decoding algorithm and provides image quality enhancement for low bit rate lossy image compression, which is sometimes an essential issue, if there are transmission or storage constraints.

Similar result were obtained for wavelet coding [8].

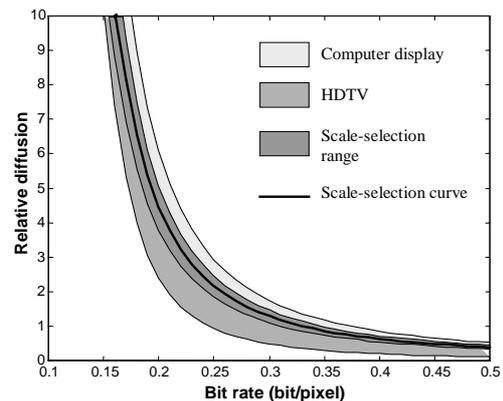
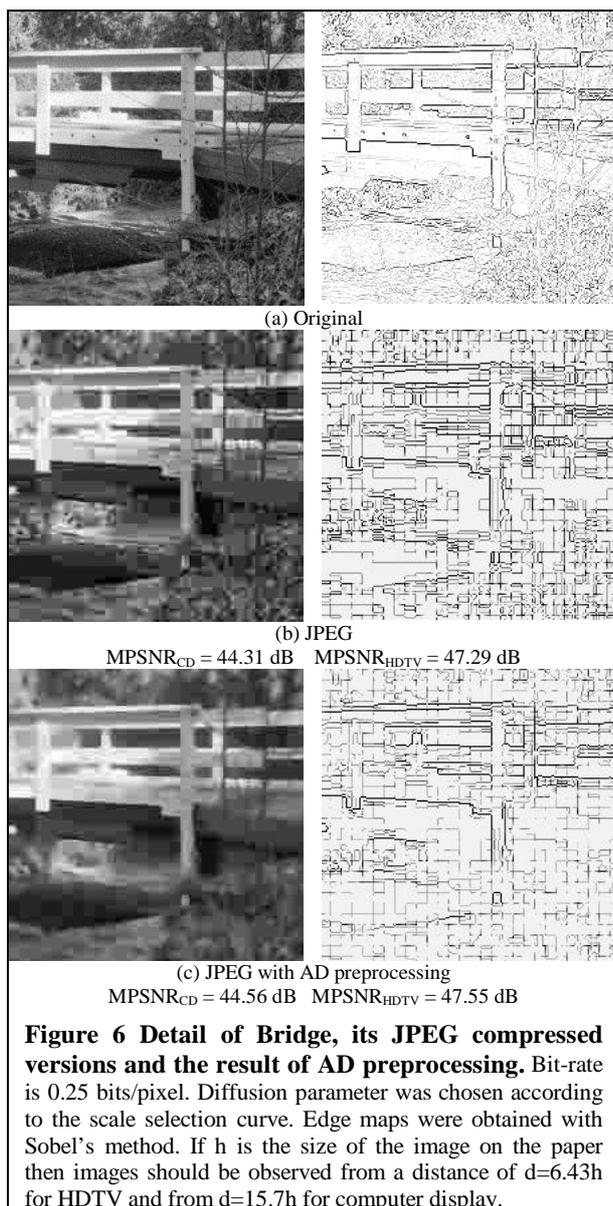


Figure 5 Scale-selection ranges for computer display and HDTV configuration, and the general scale-selection range with the scale-selection function. The scale-selection function is of the form $t(c) = ac \exp(bc)$, where c is the bit-rate, $a = 0.0102$ and $b = 0.062$.

Acknowledgements

Special thanks are due to Stefan Winkler and Christian Lambrecht and the Signal Processing Laboratory (LTS) of EPFL (Lausanne) for the HVS-based distortion measuring software [9] supported for us for testing. This work was supported by the Hungarian Research Fund (OTKA) and by the Swiss Research Fund (No. 7UNPJ048236).



References

[1] Al-Shaykh, O. K., Mersereau, R. M.: *Lossy Compression of Noisy Images*, IEEE Trans. on Image Proc., Vol.7., No.12., 1998, pp. 1641-1652

[2] Catté, F., Coll, T., Lions, P. L. and Morel, J. M.: *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM J. Numerical Anal., 1992, Vol. 29, pp.182-193

[3] CCIR, "Method for the Subjective Assessment of the Quality of Television Pictures", 13th Plenary Assembly, Recommendation 500, Vol.11, pp.65-68, 1974

[4] Crouse, M., Ramchandran, K.: *Joint Thresholding and Quantizer Selection for Transform Image Coding: Entropy-Constrained Analysis and Application to Baseline JPEG*, IEEE Trans. on Image Proc., Vol.6., No.2., 1997, pp. 285-297

[5] Daugman, J. G.: *Two-dimensional Spectral Analysis of the Cortical Receptive Field Profiles*, Vision Research, Vol. 24, 1980, pp. 891-910

[6] Florack, L.: *Image Structure*, Kluwer Academic Publishers, 1997

[7] Guichard, F., Morel, J., *Partial Differential Equations and Image Iterative Filtering*, State of the Art in Numerical Analysis, Oxford University Press, 1997.

[8] Kopilovic, I., Szirányi, T.: *Image structure preserving of lossy compression in the sense of perceptual distortion when using anisotropic diffusion preprocessing*, FSPIPA'99, IAPR, Budapest, Hungary, 6-7 Sept. 1999.

[9] Lambrecht, C. J. B., Verscheure, O.: *Perceptual Quality Measure Using a Spatio - Temporal Model of the Human Visual System*, Proceedings of SPIE, V.2668, pp.450-461, San Jose, CA, 1996.

[10] Lee, Y., L., Kim, H., C., Park, H., W.: *Blocking Effect Reduction of JPEG Images by Signal Adaptive Filtering*, IEEE Transactions on Image Processing, Vol.7, No.2, 1998, pp.229-234

[11] Lindeberg, T., Haar Romeny, B. M.: *Linear scale-space I-II, Geometry - Driven Diff. In Computer Vision*, Kluwer Academic Publishers, 1992, pp.1-72

[12] Malo, J., Pons, A. M., Artigas, J. M.: *Subjective image fidelity metric based on bit allocation of the human visual system in the DCT domain*, Image and Vision Computing, 1997, pp. 535-548

[13] Pennebaker, W. B., Mitchell, J. L. : *JPEG Still Image Data Compression Standard*, Van Nostrand Reinhold, 1993

[14] Perona, P., Shiota, T., Malik, J.: *Anisotropic Diffusion, Geometry - Driven Diff. In Computer Vision*, Kluwer Academic Publishers, 1992, pp.73-92

[15] Prost, R., Baskurt, A.: *JPEG Dequantisation Array for Regularized Decompression*, IEEE Trans. on Image Proc., Vol.6., No.6., 1997, pp. 883-888

[16] Rosenholtz, R., Watson, A. B.: *Perceptual adaptive JPEG coding*, Proc. IEEE Int.Conf. on Image Processing, Vol. 1, 1996, pp. 901-904

[17] Roska, T., Szirányi, T.: *Classes of Analogic CNN Algorithms and Their Practical Use in Complex Processing*, Proc. IEEE Non-linear Signal and Image Proc. June, 1995, pp.767-770

[18] Szirányi, T., Kopilovic, I., Tóth, B. P.: *Anisotropic Diffusion as a Preprocessing Step for Efficient Image Compression*, Proceedings of the 14th ICPR, Brisbane, IAPR, Australia, August 16-20, 1998, pp. 1565-1567

[19] JPEG Group's CJPEG, v 6a, JPEG Group's DJPEG, v 6a, Copyright (C) 1996, Thomas G. Lane