IMAGE STRUCTURE PRESERVING OF LOSSY COMPRESSION IN THE SENSE OF PERCEPTUAL DISTORTION WHEN USING ANISOTROPIC DIFFUSION PREPROCESSING

Iván Kopilovic[1] Tamás Szirányi[2]

Abstract

In this paper we show how the image structure can be better preserved at lossy compression, when compression artifact is reduced with perceptually adaptive filtering by means of anisotropic diffusion, performed on the image before coding. Adaptive filtering extracts and preserves the important visual information, and it will be discussed in the context of the non-linear scale-space theory. Among the proposed non-linear diffusion equation we shall identify the one that is adequate for preprocessing. The effect of linear and non-linear diffusions prior to Joint Photographic Experts Group's (JPEG) standard baseline compression is analyzed and compared. It is shown that the selection of the appropriate preprocessing parameters at a specified bit-rate greatly reduces the compression artifacts. The selection criteria were determined using a perceptual error metric based on a model of the human visual system (HVS). The proposed method preserves well the main structure of the decompressed image and it does not change the perceptible image quality of the decompressed image considering all the details, while removing most of the compression artifacts.

1.Introduction

Application of the non-linear scale-space theory to non-recoverable, low bit-rate image compression will be presented. In particular, the method discussed is a preprocessing that improves the JPEG baseline coding standard [15], and it may be considered as a blocking effect reduction algorithm.

[1] University of Veszprém, Department of Image Processing and Neurocomputing, Egyetem u. 10., H-8200 Veszprém, Hungary
[2] Analogical and Neural Computing Laboratory, Comp. & Automation Inst., Hungarian Academy of Sciences, Kende u. 13-17., H-1111 Budapest, Hungary

There exist several blocking effect reduction methods ranging from simple smoothing of the image before compression to sophisticated quantizer design [4], lapped transform techniques [6] or postprocessing [11] of the decompressed image. Smoothing entails the loss of edge information, other methods involve changes in the standard compression and/or decompression algorithm. Postprocessing is done on images, where the original information is already seriously damaged.

In contrast to these methods, we have introduced a preprocessing step [21] performed before coding that will reduce the artifacts 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 extracting the perceptually important information, achieved by non-linear diffusion processing.

We used a perceptual error metric [10] as a quality criterion to establish a selection rule [9] for the parameters of the preprocessing to achieve the best compression quality for a given bit-rate. The effectiveness of the algorithm will be evaluated and demonstrated.

2. Non-linear scale-space and adaptive filtering

By perceptually adaptive filtering we shall mean operations on images that suppress noise and enhance edge information. 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. The result of this procedure is the extraction of the perceptually important information consisting of uniform region and enhanced boundaries.

The search for such operations ended in a unified mathematical framework called non-linear scale-space theory established basically in [7,8,13]. Non-linear scale space theory represents the extraction of information as a series of operations. It is modeled 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 the scale-space representation of the particular information extraction operation. Some natural assumptions can be made on these operators. Among the most important are translation and rotation invariance, non-appearance of new details, recursive structure, and local (differential) nature. The fundamental result of the theory is that certain class of scale-space representations verifying the above properties is generated by partial differential equations (PDEs), and conversely, a wide class of PDEs generates a scale-space representation verifying the above mentioned properties.

According to these results, a wide class of diffusion-like non-linear PDEs generates scale-space representation, where the scale parameter of the scale-space representation corresponds to the "time" parameter of the PDE, and it will be termed the "diffusion parameter" or the "scale" interchangeably. The most important PDEs in this respect are the linear diffusion [7,13], the Perona-Malik-Catte anisotropic diffusion [2,16], the non-linear isotropic [18], and the pure anisotropic diffusion [2].

Diffusion parameter is expressed in degrees of the 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 obtained this way will be called the relative diffusion.

3. A HVS-Based Evaluation Scheme

There have been several attempts to develop image distortion metrics based on the sensation of the human visual system [10,14], motivated by the well known fact that pixel-by-pixel based distortion measures fail to give correct assessment.

A perceptual error metric found in [10] was used in the experiments to compare and evaluate the compression results. The error metric accounts for several mechanisms of the human visual system. The inputs of the algorithm are the viewing parameters (see Table 1) 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, since cells of the visual cortex have similar responses [5]. For each band, the effect of masking and contrast sensitivity is computed to weigh the error coefficients. This perceptually masked error will be denoted by *E*.

A logarithmic measure can be obtained from the masked error, it is called the masked peak signal to noise ratio and is computed [10] as $MPSNR(f,g) = 20\log_{10}(1/E(f,g))$, where f and g are two images. We shall also use a subjective image quality assessment scale defined in [3], which assigns quality grades from 1 (the worst) to 5 (the best). The corresponding distortion measure will be denoted by $Q(f,g)$, where f and g are images. In [10] a relation between the masked error E and Q was established with the formula $Q(f,g) = 5(1 + 159.5E(f,g))^{-1}$.

Parameter settings contained in Table 1 were used.

Table 1

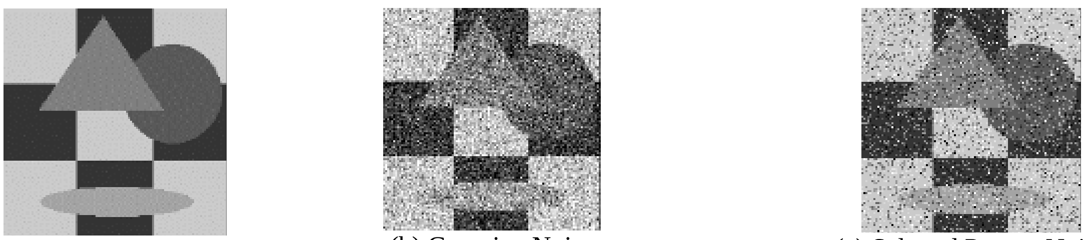
Viewing configurations used for image distortion measurements.

	Computer Display	HDTV
Viewing distance (m)	0.5	1.8
Pixels/inch	72	48
Pixels/degree	24.7	60.3
Pixel size (deg)	0.040	0.017

4. Testing the performance of the PDEs

We have tested the performance of well-known PDEs as perceptually adaptive filters for compression artifact reduction. We set up a model of information extraction process. A noise free artificial test image shown in Fig. (1a) was used to represent the essential visual information that was to be extracted. The image was added two kinds of noises: Gaussian noise and “salt and pepper” noise (Fig. 1b and 1c).

Both of the noisy images were preprocessed with diffusion and then compressed with JPEG. The results are shown in Table 2 together with the perceptual error versus the original image (Fig. 1a). The best perceptually adaptive filter falls out to be the PMC-AD, since it suppresses noise while performing shape enhancement and gives the best results in terms of perceptual error metric, i.e. it results in an image that is the closest to the original.



(a) Original

(b) Gaussian Noise
$MPSNR_{C.D.}$=31.32 dB $MPSNR_{HDTV}$=34.77 dB

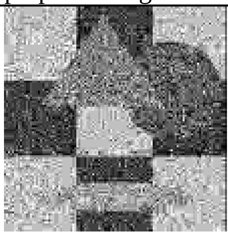
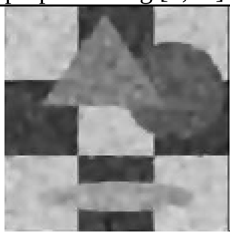
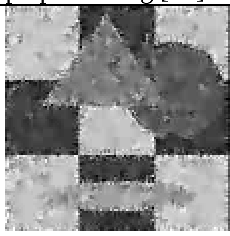
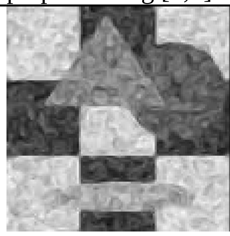
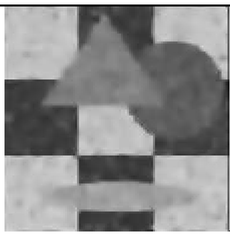
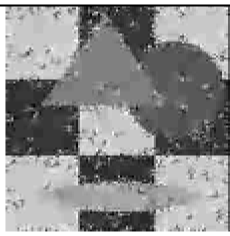
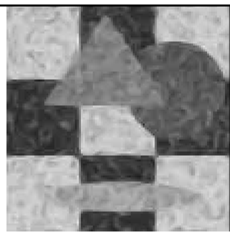
(c) Salt and Pepper Noise
$MPSNR_{C.D.}$=32.53 dB, $MPSNR_{HDTV}$=35.85 dB

Figure 1

Artificial test image used for testing different types of anisotropic diffusions. Gaussian noise was of 10dB, salt and pepper noise affected 15% of pixels.

Table 2

Compression results for the test image with and without preprocessing. The first row contains the results for the noisy image from Fig. 1(b), the compression bit-rate is 0.8 bits/pixel. The second row contains the results for the noisy image from Fig. 1(c), the compression bit-rate is 0.64 bits/pixel.

	JPEG with no preprocessing	JPEG with PMCAD preprocessing [2,16]	JPEG with NLID preprocessing [18]	JPEG with P-AD preprocessing [2,8]
Gaussian noise	 $MPSNR_{C.D.}$= 34.52 dB $MPSNR_{HDTV}$ = 37.46 dB $PSNR$=14.34dB	 $MPSNR_{C.D.}$= 49.37 dB $MPSNR_{HDTV}$ = 50.75 dB $PSNR$=25.41dB	 $MPSNR_{C.D.}$= 43.22 dB $MPSNR_{HDTV}$ = 44.29 dB $PSNR$=20.13dB	 $MPSNR_{C.D.}$= 44.15 dB $MPSNR_{HDTV}$ = 45.03 dB $PSNR$=22.33dB
Salt & pepper noise	 $MPSNR_{C.D.}$= 37.02 dB $MPSNR_{HDTV}$ = 39.91 dB $PSNR$=15.83dB	 $MPSNR_{C.D.}$= 49.12 dB $MPSNR_{HDTV}$ = 51.37 dB $PSNR$=25.03dB	 $MPSNR_{C.D.}$= 42.38 dB $MPSNR_{HDTV}$ = 43.22 dB $PSNR$=19.64dB	 $MPSNR_{C.D.}$= 46.2 dB $MPSNR_{HDTV}$ = 47.33 dB $PSNR$=23.83dB

5. Parameter Settings for the Preprocessing Algorithm

Most of the compression methods result in unrealistic, sometimes annoying artifacts at low bit rates, as shown if Figures 3. JPEG compression produces a characteristic blocking-artifact, and in case of wavelet compression, there is a characteristic compression distortion as well; parts of the image become blurred, while in other parts of the image false textured regions appear (Figures 3). We have seen that preprocessing by PMC-AD reduces the compression error. The question is how to select the appropriate scale (diffusion parameter) to achieve best result.

We set two criteria for a good performing image quality enhancement. It should not affect the original information content of the image or it should cause as little change as possible, i.e. it should preserve the main image information as much as possible. On the other hand, it should eliminate the formation of compression artifacts. Motivated by these two constraints, we have introduced [9] two characteristic features of a compression algorithm relative to a fixed adaptive filter that is given by the scale-space representation T:

1. *The perceptual scale* $t_p(c)$ of the compression at a given bit rate is defined as the greatest scale

at which the perceptual error of the preprocessed compression is not greater than that of the plain compression. This refers to the maximal scale, i.e. the maximal loss of details, which remains imperceptible at the given bit rate.

2. *The scale capacity* $t_k(c)$ of the compression at a given bit rate is defined as the smallest scale at which the compression algorithm operates below a specified error level, i.e. for images processed with smaller scale than $t_k(c)$, compression will give rise to undesirable error.

Note that for perceptual scale, the error is measured relative to the original, while for scale capacity it is measured relative to the enhanced image. It is desired [9] that the selected scale *t(c)* for the bit rate *c* obey the inequality $t_k(c) \le t(c) \le t_p(c)$, i.e. it must fall between scale-capacity and perceptual scale. We have established such a general scale-selection function in [9]. It is shown in Figure 2. It is independent of the viewing condition, i.e. it is valid through a wider range of pixels/degree values, and it specifies the scale to be used with AD for each bit rate.

The scale selection function *t* is used to determine the number of AD iterations *j* in the preprocessing step according to the rule $j(c) = [t(c)/\ l]$, where 0<*c*≤8 is the bit-rate, and l is the diffusion step size.

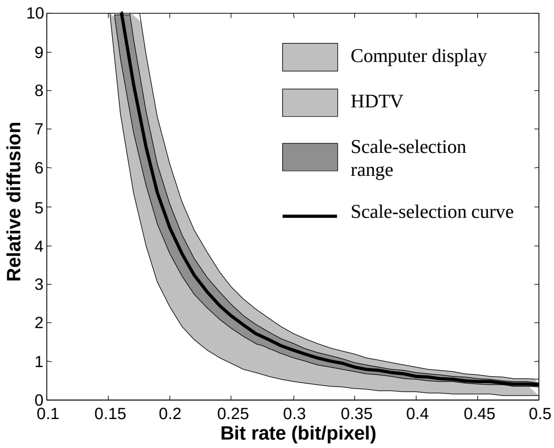


Figure 2

General scale-selection range and the scale selection curve $t(c) = ac^{-1}\exp(bc^{-1})$, *where a = 0.0816 and b = 0.496.*

6. Experimental Demonstrations

In this section we demonstrate the image enhancement property of the AD through several experimental examples of different test-images. In Figures 3 the middle row show the results on the test image “Barbara”.

Figure 3

Compression results for Barbara. JPEG, AD preprocessed JPEG and wavelet compression are compared. Wavelet compression was accomplished with [22]. Diffusion parameter was chosen according to the scale selection curve. If h is the size of the image on 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 case. MPSNR error values are indicated as well (CD = Computer display).

Figure 4 shows the compression diagrams (bit-rate vs. HVS quality-rate) for this test image. Figure 5 shows the result for image Lena. Similar MPSNR charts were obtained for Lena as in Figure 4 for Barbara. Snapshots of compression results together with the edge-maps for further test images are shown through Figure 6 and 7. Note that the edge-maps of AD preprocessed images are “cleaner” and contain less noise. In these examples scales at a given bit-rate were selected according to the established *general* scale selection curve (Figure 2). Tuning scales to particular images is possible as well, though it may be quite time consuming. Despite the generality of the scale selection the resulted charts in Figures 4 demonstrates well the weaker performance of the linear diffusion.

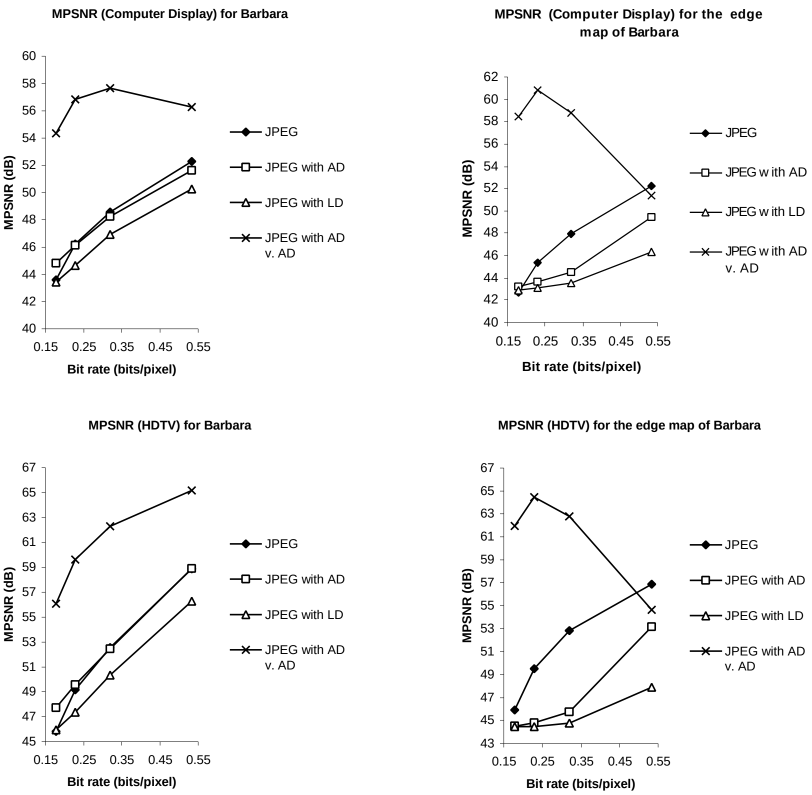


Figure 4

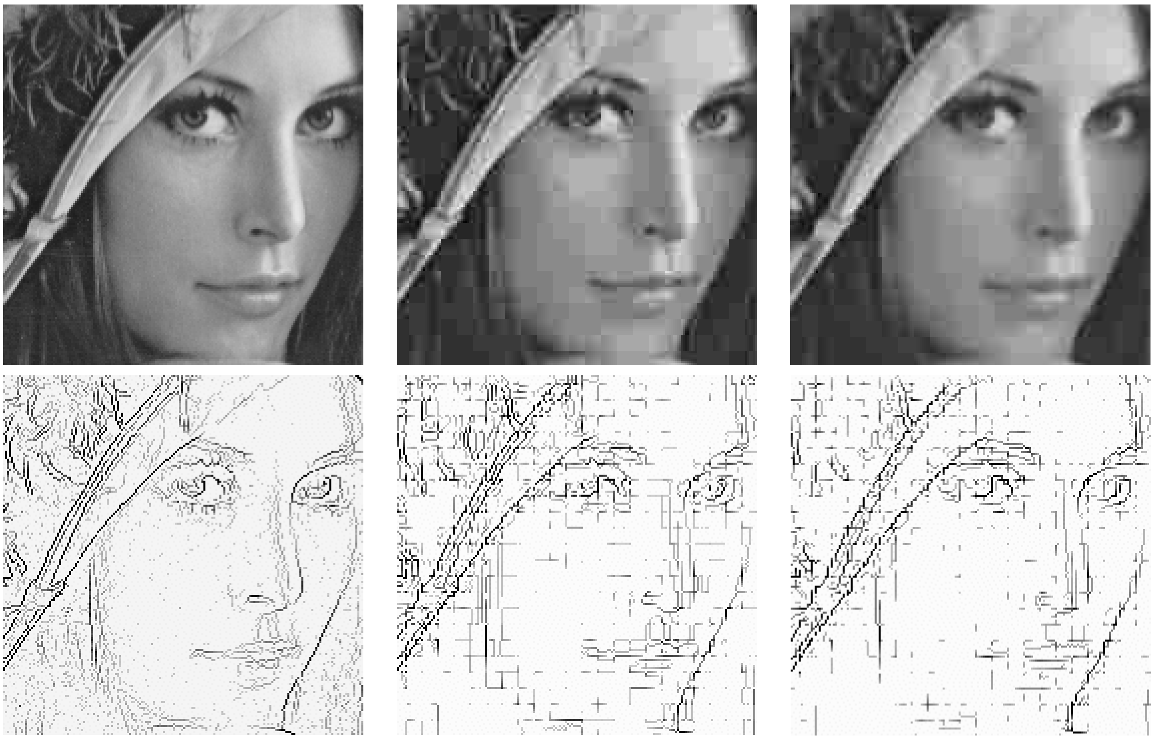
MPSNR values for compression results compared to the original image with different viewing configurations and at different bit-rates. Results are shown for test image Barbara. Scales at a given bit-rate were selected according to the established general scale selection curve. Tuning scales to particular images is possible as well, though it may be quite time consuming. Despite the generality of the scale selection the above charts demonstrate well the weaker performance of the linear diffusion. As concerning the edge map errors, JPEG error values are alleviated by added local frequency content contributed by edge artifacts, especially at low bit-rates, where there is a significant amount of quantization error.

As concerning the edge map errors, JPEG error values are alleviated by added local frequency content contributed by edge artifacts, especially at low bit-rates, where there is a significant amount of quantization error.

From these experiments and diagrams we can conclude that

- The AD preprocessed image is transmitted through the compression process with much lower distortion than compressing only the original image.

- The compressed AD preprocessed image results in nearly the same MPSNR value than the transformed original one, comparing with the original image in all details.
- AD gives significantly better results than using a simple smoothing in the preprocessing step.



(a) original

(b) JPEG
$MPSNR_{CD}$ = 53.51 dB
$MPSNR_{HDTV}$ = 56.94 dB

(c) JPEG with AD preprocessing
$MPSNR_{CD}$ = 53.95 dB
$MPSNR_{HDTV}$ = 57.42 dB

Figure 5

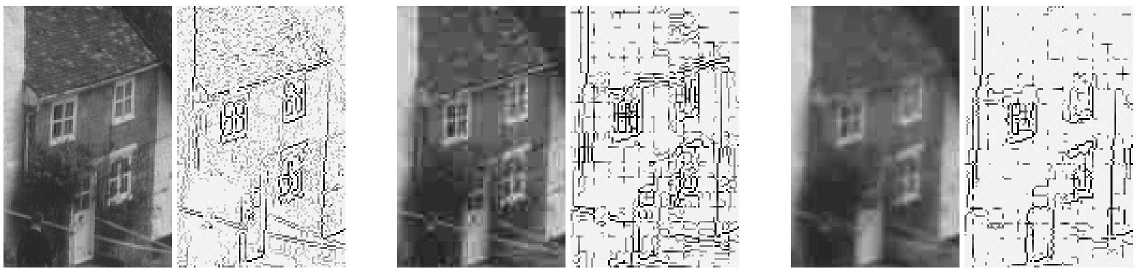
Detail of Lena, its JPEG compressed versions and the result of AD preprocessing. Bit-rate is 0.25 bits/pixel. Diffusion parameter was chosen according to the scale selection curve. Edge maps were obtained with Sobel's method. If h is the size of the image in the paper then images should be observed from a distance of d=11.14h for the HDTV case and from d=4.56h for the computer display case. Error in terms of MPSNR is indicated (CD = computer display).

7. Computational complexity of the implementation

The number of PMC AD iterations is bit-rate dependent. It was shown [9] that with decreasing bit-rate the number of AD iterations grows according to the scale selection function (Figure 2).

A good choice of diffusion step-sizes can reduce the number of operations needed. Error measurements have shown that the step-size for AD can fall in the interval [0.1, 0.25]. Fast implementation suggests the choice of λ_{LD}=0.125=1/8 and λ_{AD}=0.25=1/4, since in this case, multiplication by λ_{LD} and λ_{AD} is a simple shift operation. Adaptive implementation should include features such as keeping record of, and executing the non-linear step only for the points of the edgy locations, or not performing the pre-diffusion at each AD iteration, etc. Multigrid implementation can also speed up the processing, as described in [1], where the execution time is decreased by an order of magnitude.

Partial differential equations allow parallel computation of the numerical scheme for several or all the pixels, which can be exploited using partially or fully parallel computing architectures (MMX, CNN VLSI chip [12]). With parallel imaging circuits (like CNN [12]), the whole compression process together with AD preprocessing and other image enhancement steps [19] could be very fast, as well as dynamic coding of moving images [20]. Running time estimates have shown that at 0.25 bits/pixel, the preprocessing on a Pentium II 400MHz MMX takes 0.4 seconds, on a CNN array [12] it takes 0.08 seconds, while the non-linear isotropic approximation [18] takes 0.01 seconds.



(a) original

(b) JPEG
$MPSNR_{CD} = 51.16$ $MPSNR_{HDTV} = 56.36$

(c) AD preprocessed JPEG
$MPSNR_{CD} = 50.38$ $MPSNR_{HDTV} = 54.65$

Figure 6

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. Error in terms of MPSNR is indicated (CD = computer display).

8. Conclusion

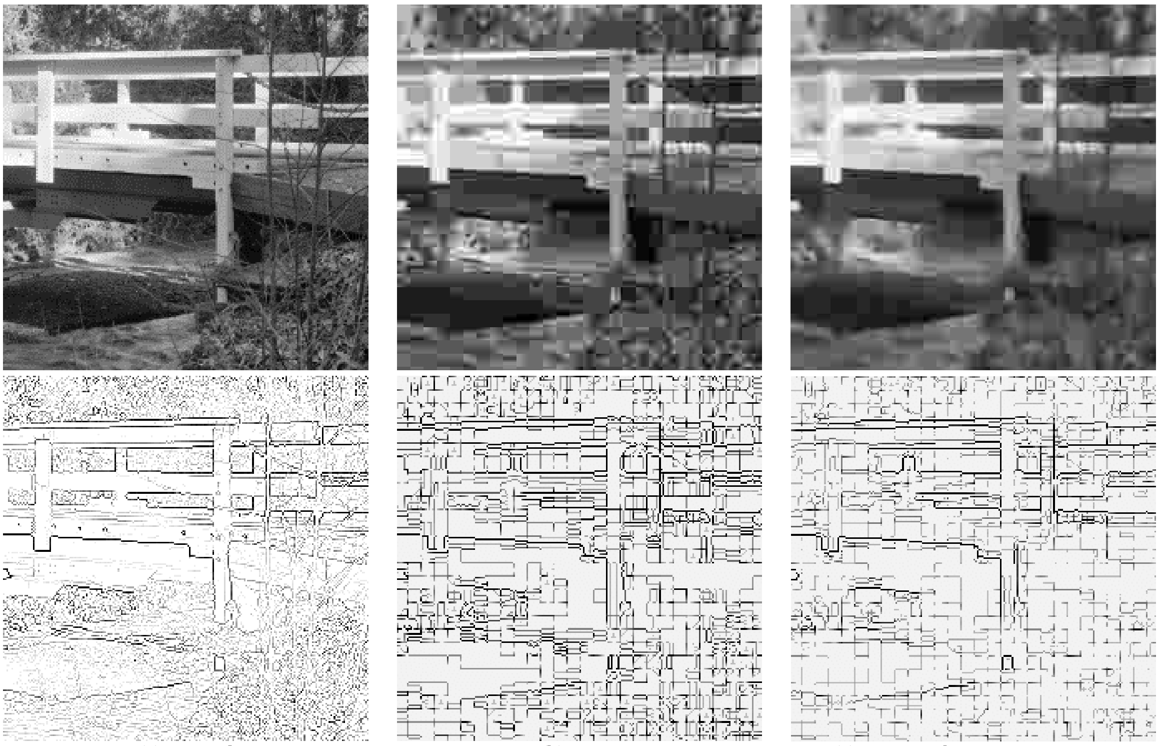
We have shown that a scale-space representation generated by a non-linear anisotropic diffusion can be effectively used to enhance JPEG compression and prevent the formation of artifacts. Using perceptual distortion measurement, it was also shown that AD is a better means of preprocessing than ordinary smoothing (LD) both with respect to perceptual error metric and edge-adaptivity.

Now, we summarize the advantages of using AD preprocessing previous to compression:

- AD enhances and preserves the main image information.
- The AD preprocessed image is transmitted through the compression process at much lower distortion than compressing only the original image (MPSNR and PSNR values are much better for AD than for the other case).
- The compressed AD preprocessed image results in nearly the same MPSNR value than the

transformed original one, comparing with the original image in all details.

- The compression-rate and the scale of AD can be coupled considering the selection curve.
- The AD preprocessing method requires no change in the standard decoding algorithm while it provides image quality enhancement for low bit-rate lossy image compression.



(a) original

(b) JPEG
$MPSNR_{CD}$ = 44.31 dB
$MPSNR_{HDTV}$ = 47.29 dB

(c) JPEG with AD preprocessing
$MPSNR_{CD}$ = 44.56 dB
$MPSNR_{HDTV}$ = 47.55 dB

Figure 7

Detail of Bridge, 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=6.43h for the HDTV case and from d=15.7h for the computer display case. Error in terms of MPSNR is indicated (CD = computer display).

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 [10] 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] ACTON, S.T., Multigrid Anisotropic Diffusion, IEEE Trans. on Image Proc., Vol.7., No.3, 1998, pp. 280-291.

[2] CATTÉ, F., COLL, T., LIONS, P.L., 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] DEBRUNNER, V.E., CHEN, L., LI, H.J., Lapped Multiple Bases Algorithms for Still Image Compression Without Blocking Effect, IEEE Trans. on Image Proc.,Vol.6., No.9., 1997, pp. 1316-1321

[7] FLORACK, L., Image Structure, Kluwer Academic Publishers, 1997.

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

[9] KOPILOVIC, I., SZIRÁNYI, T., Non-linear Scale Selection for Image Compression Improvement Obtained by HVS-based Distortion Criteria, ICIAP, Venice, IAPR, Italy, 1999.

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

[11] 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

[12] LINAN, G., ESPEJO, S., DOMINGUEZ-CASTRO, R., ROCA, E., RODRÍGUEZ-VÁZQUEZ, A., A Mixed Signal 64x64 CNN Universal Machine Chip, MicroNeuro'99, IEEE, Granada, Spain, 1999, pp.61-68

[13] LINDEBERG, T., HAAR ROMENY, B. M., Linear scale-space I-II, Geometry - Driven Diff. In Computer Vision, Kluwer Academic Publishers , 1992, pp.1-72

[14] MALO, J., PONS, A.M., FELIPE, A., ARTIGAS, J.M., Characterisation of the human visual system threshold performance by a weighting function in the Gabor domain, Journal of Modern Optics,Vol.44, No.1,1997, pp. 127-148.

[15] PENNEBAKER, W.B., MITCHELL, J.L., JPEG Still Image Data Compression Standard, Van Nostrand Reinhold, 1993

[16] PERONA P., SHIOTA T., MALIK J., Anisotropic Diffusion, Geometry - Driven Diff. In Computer Vision, *Kluwer* Academic Publishers , 1992, pp.73-92

[17] PROST, R., BASKURT, A., JPEG Dequatisation Array for Regulatized Decompression, IEEE Trans. on Image Proc.,Vol.6., No.6., 1997, pp. 883-888

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

[19] SZIRÁNYI, T., CSAPODI, M., Texture Classification and Segmentation by Cellular Neural Network using Genetic Learning, Computer Vision and Image Understanding, Vol. 71, No. 3, pp. 255-270, 1998

[20] SZIRÁNYI, T., CZÚNI, L., Image Compression by Orthogonal Decomposition Using Cellular Neural Network Chips, Int. J. Circuit Theory and Applications, Vol. 27, No. 1, pp.117-134, 1999

[21] SZIRÁNYI, T., KOPILOVIC, I., TÓTH, B.P., Anisotropic Diffusion as a Preprocessing Step for Efficient Image Compression, Proc. of the 14th ICPR, Brisbane, IAPR, Australia, August 16-20, 1998, pp.1565-1567

[22] Wavelet coder by G. Davis, http://pascal.dartmouth.edu/~gdavis/wavelet/wavelet.html.

[23] Independent JPEG Group's CJPEG, version 6a, 7-Feb-96, Independent JPEG Group's DJPEG, version 6a, 7-Feb-96, Copyright (C) 1996, Thomas G. Lane.