

***salienShrink*: SALIENCY-BASED WAVELET SHRINKAGE**

Konstantinos Rapantzikos, Yannis Avrithis, Stefanos Kollias

School of Electrical and Computer Engineering,
National Technical University of Athens
e-mail: {rap,iavr}@image.ntua.gr, stefanos@cs.ntua.gr

ABSTRACT

This paper describes *salienShrink*, a method to denoise images based on computing a map of salient coefficients in the wavelet domain and use it to improve common denoising algorithms. By salient, we refer to those coefficients that correspond mostly to pure signal and should therefore be preserved throughout the denoising procedure. We use a computationally efficient model to detect salient regions in the bands of the multiresolution wavelet transform. These regions are used to obtain a more accurate estimate of the noise level, improving the performance of existing well known shrinkage methods. Extensive experimental results on the *BiShrink* method show that the proposed method effectively enhances PSNR and improves the visual quality of the denoised images.

Index Terms— saliency detection, wavelet shrinkage, image denoising

1. INTRODUCTION

Image denoising remains one of the fundamental problems in the field of image processing and still poses challenges to the researchers. Noise in images is usually present due to capturing instruments, transmission, quantization etc. and despite the huge number of published techniques each approach has its own assumptions, advantages and limitations. Hence, usually successful denoising depends on the task at hand.

Discrete Wavelet Transform (DWT) has been widely used to suppress noise in images, since properties like sparsity and multiresolution structure fit well with the denoising procedure. Sparsity refers to the property of the wavelet transform to concentrate signal energy into a small number of large coefficients enabling therefore the reduction of noise with appropriate thresholding reduction of the absolute coefficient values [1]. *Wavelet shrinkage*, a well known non-linear coefficient thresholding method, is based on this fact. Preliminary shrinkage methods used either a universal threshold, like *VisuShrink* [1], or a subband adaptive one like the *SureShrink* [9] and *BayesShrink* [10]. Extensions of these methods have been appeared and prove to be more successful since they exploit dependency between coefficients either at the same or at different scales. Literature is rich in methods predicting the contribution of wavelet coefficients based only on intra-

scale or both on intra- and inter- scale dependencies [2][4][3][5][6][7][8].

Such approaches are either based on a local neighborhood or assume a specific coefficient model extracted through statistical analysis. Usually they exploit the multiresolution properties of the wavelet transform by identifying across scale correlations between wavelet coefficients and local correlation between neighborhood coefficients. These approaches, no matter if they employ a model or not, require the computation of appropriate statistical properties at the coefficient or subband level. Chen *et al.* proposed *NeighShrink*, a wavelet thresholding scheme, by incorporating neighboring coefficients [4]. Sendur & Selesnick assumed non-Gaussian bivariate distributions and proposed non-linear threshold functions (shrinkage functions) derived from this model using Bayesian estimation theory [6][7].

In the proposed work, we use the notion of saliency in the wavelet domain and propose a simple way to produce denoised images both of high PSNR and high visual quality. We assume that the high salient regions detected by our method correspond to regions of high signal structure and low noise (e.g. areas with many edges). The adopted method for saliency detection is based on a simplified version of the selective tuning (ST) model proposed by Tsotsos *et al.* [11]. According to the proposed model, the visual input is decomposed into several bands using the DWT or the Dual-Tree Complex Wavelet Transform (DTCWT) and each subband is searched for salient coefficients. This search is accomplished using an across-scale Winner-Take-All (WTA) network that detects the most salient coefficients in terms of magnitude and is inspired by the ST model. The method results in a final saliency map (S) for each of the subbands that represents the high and low saliency regions in terms of gray values.

For comparison purposes, we use publicly available implementations of well established algorithms and provide results on a variety of images both in terms of PSNR and visual quality at different levels of noise. The computed saliency map is used to enhance *BiShrink*, the denoising method proposed by Sendur & Selesnick [6], both for the real DWT and the DTCWT.

2. SALIENCY-BASED DENOISING

2.1. Computing saliency in the wavelet domain

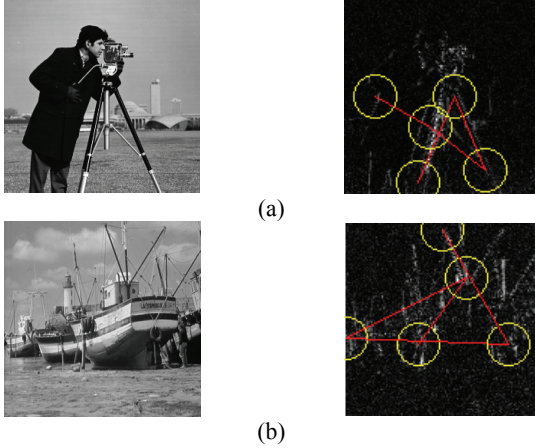


Fig. 1 original images and 5 attended regions of a random wavelet band (better viewed in color); (a) “cameraman”; (b) “boat”

In order for DWT to be performed, a pair of low-pass $h_\phi(\cdot)$ and high-pass $h_\psi(\cdot)$ filters is applied to the image in both the horizontal and vertical directions. The filter outputs are then subsampled by a factor of two, generating a set $D_{DWT} = \{D_i\}$, with $i = \{1,2,3\}$ of three high pass bands that correspond to horizontal, vertical and diagonal details and a low-pass subband A that corresponds to the approximation coefficients. Although DWT has been quite successfully applied to many applications it suffers from orientation selectivity and shift variance. Recently, the DTCWT was proposed by Kingsbury, which has good directional selectivity and its subband responses are approximately shift-invariant [13]. DTCWT gives rise to wavelets in six distinct directions and therefore produces a set of six detail subbands $D_{DTCWT} = \{D_i\}$, with $i = \{1, \dots, 6\}$. There are two wavelets in each direction, one for the real and one for the complex part. Sendur & Selesnick have used DTCWT successfully in image denoising [6][7]. More details on their technique will be given in section 2.2.

We approach the task of selecting salient wavelet coefficients by using a saliency detection method inspired by existing models of visual attention mechanisms. Most of these computational models decompose the visual input into a multiresolution structure, where each level is a simplified version of the previous one (usually Gaussian filtered) and apply simple operations across scales to detect the most prominent pixel neighborhoods [12]. In the proposed method, we exploit the multiresolution property of the DWT and apply a simplified version of the selective tuning model to the detail subbands. We first select the strongest item at the top level of the processing hierarchy (coarser level of the WT) using a WTA process and then propagate the winning location to each of the higher scales of representation. We use this scheme, because important (high valued) wavelet coefficients remain intact through scales. This process includes an Inhibition-Of-Return (IOR) mechanism, which ensures that the saliency map is scanned in order of decreasing saliency by the focus of attention, and

generates the model's output in the form of spatio-temporal attentional scanpaths. The stopping criterion of the WTA approach or in other words the number of regions to be attended will be discussed in the next subsection. The final saliency map is produced at the desired level for each of the subbands. We compute one such map S_i^l for each subband i , at each level l of the decomposition.

2.2. Denoising

The extended BiShrink method that we use for our experiments is proposed by Sendur *et al.* [7] and is based on non-Gaussian distributions to model intra- and inter-scale dependencies. A short description of their method and the proposed framework are given in this section.

If coefficient w_{2k} is at the same position as the k^{th} coefficient w_{1k} , but at a coarser scale then the noisy observations of w_{1k} and its parent w_{2k} are $y_{1k} = w_{1k} + n_{1k}$ and $y_{2k} = w_{2k} + n_{2k}$ respectively. A non-linear bivariate shrinkage function using the Maximum-A-Posteriori (MAP) estimator is then derived and the MAP estimator of w_{1k} becomes [6]

$$\hat{w}_{1k} = \frac{\left(\sqrt{y_{1k}^2 + y_{2k}^2} - \frac{\sqrt{3}\sigma_n^2}{\sigma} \right)_+}{\sqrt{y_{1k}^2 + y_{2k}^2}} \cdot y_{1k} \quad (1)$$

where

$$(g)_+ = \begin{cases} 0, & \text{if } g < 0 \\ g, & \text{otherwise} \end{cases} \quad (2)$$

The estimator requires the knowledge of the noise variance σ_n^2 and the marginal variance σ^2 for each wavelet coefficient. Noise variance is usually estimated using a robust median estimator from the finest scale wavelet coefficient of the D_3 subband (diagonal) [7][4].

$$\hat{\sigma}_n^2 = \frac{\text{median}(|D_3|)}{0.6745} \quad (3)$$

Marginal variance is calculated by

$$\hat{\sigma}_y^2 = \frac{1}{M} \sum_{y_i \in N(k)} y_i^2 \quad (4)$$

where M is the number of coefficients in a neighborhood $N(k)$ around y_i . Finally, σ is estimated as

$$\hat{\sigma} = \sqrt{(\hat{\sigma}_y^2 - \hat{\sigma}_n^2)_+} \quad (5)$$

We exploit Eq. 4 for determining the stopping criterion of the WTA process (section 2.1) by allowing the system to further attend regions while the ratio between the $\hat{\sigma}$ of the currently and the previously attended region remains in a pre-defined range $[\epsilon_1, \epsilon_2]$. By rephrasing the main assumption of the proposed work (regions of high saliency correspond to regions of high signal presence), we may say

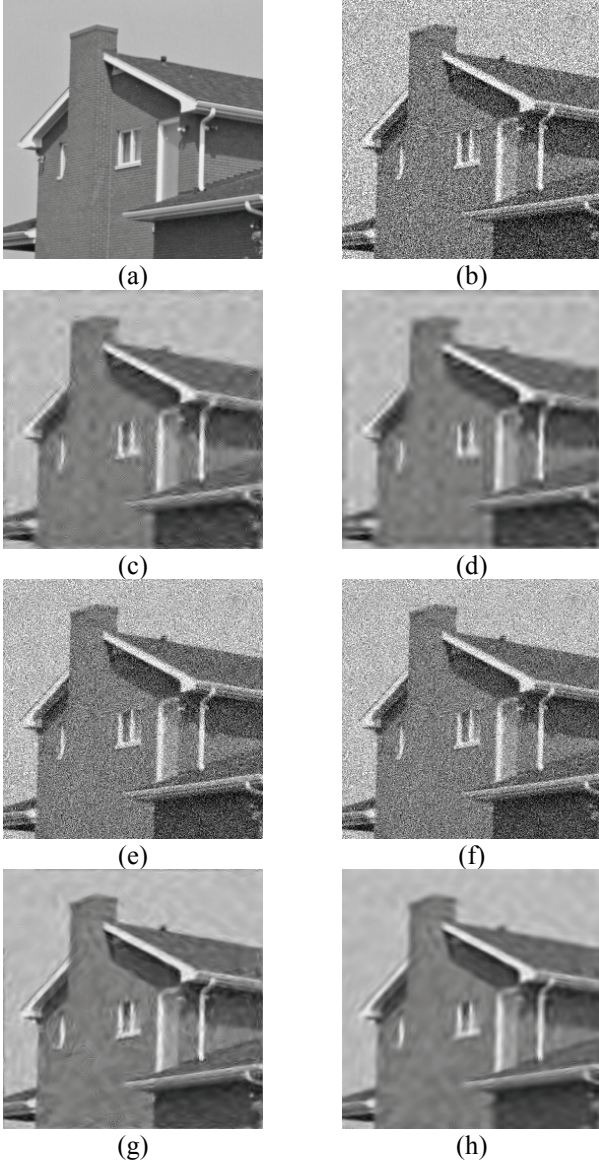


Fig. 2 (a)-(b) original and noisy image with $\sigma=35$; (c) DWT-salienShrink; (d) DWT-BiShrink; (e) BayeShrink; (f) Donoho; (g) DTCWT-salienShrink; (h) DTCWT-BiShrink

that salient regions contain more signal structure than noise. Keeping this in mind, we could lower the global estimate of noise variance for these regions in order to avoid removing coefficients related to pure signal. The S_i^l , as computed by the IOR mechanism, have low values inside the inhibited regions (pure signal) and high values outside them (noise). Hence, we multiply each saliency map S_i with the estimated noise variance after Gaussian smoothing and normalizing it in the range $[0, 1]$ as

$$\hat{\sigma} = \sqrt{(\hat{\sigma}_y^2 - S_i^l \cdot \hat{\sigma}_n^2)_+} \quad (6)$$

Most common shrinkage algorithms are based on predicting the signal's standard deviation and use it to obtain a

threshold $[1][9][10]$. As a consequence, the proposed way of obtaining more accurate standard deviation estimates may be similarly used for these techniques. In section 3, we present results only on the BiShrink method due to space limits.

3. EXPERIMENTAL RESULTS

Images found often in the literature are used for gauging the proposed denoising method. All images are of 256×256 grayscale and corrupted by additive white Gaussian noise. For comparison purposes, we show results for the Sendur *et al.*'s methods with and without saliency bias, for the BayeShrink and for the denoising method based on Donoho *et al.*'s work included in the Matlab image processing toolbox 5.1. We use the public available implementation of BiShrink [7] for implementing the proposed method and our own implementation of BayesShrink. For the presented experiments, we set the range $[\epsilon_1, \epsilon_2] = [0.8, 1.2]$ and the size of IOR to be a ratio of the min input image dimension $(\frac{1}{6} \cdot \min\{\text{rows}, \text{cols}\})$. The performance of the denoising algorithms is measured in terms of peak signal-to-noise-ratio (PSNR), which can be defined as follows

$$PSNR = 20 \cdot \log_{10} \left(255 / \sqrt{\frac{1}{N} \sum_k (v_k - r_k)^2} \right) \quad (7)$$

where v is the original and d the denoised image respectively. In tables 1, 2 we present PSNR results for three images. Each number is the average of 5 runs of the corresponding algorithm. All implementations along with the PSNR results and the full resolution original/denoised images can be found in

www.image.ntua.gr/~rap/icip07/salienShrink/

Table 1 "house" - PSNR(dB)

σ	15	20	25	30	35
DWT-salienShrink	31.29	29.75	28.45	27.38	26.48
DWT-BiShrink	30.43	28.58	27.15	26.13	25.31
BayeShrink	29.2	26.58	26.04	25.59	24.80
Donoho	28.40	25.32	22.93	20.98	19.42
DTCWT-salienShrink	32.71	31.32	30.21	29.32	28.61
DTCWT-BiShrink	32.74	31.27	30.12	29.14	28.32

Table 2 "boat" - PSNR(dB)

σ	15	20	25	30	35
DWT-salienShrink	27.49	25.95	24.79	23.99	23.30
DWT-BiShrink	26.27	24.73	23.65	22.97	22.42
BayeShrink	24.54	23.08	22.46	21.97	21.50
Donoho	25.65	25.16	24.17	22.93	21.61
DTCWT-salienShrink	28.62	27.23	26.33	25.44	24.76
DTCWT-BiShrink	27.81	26.37	25.40	24.51	23.82

Overall, DWT-salienShrink results in an average gain of 1.2dB over DWT-BiShrink and 0.58dB over DTCWT-BiShrink. In most cases, this improvement is more evident

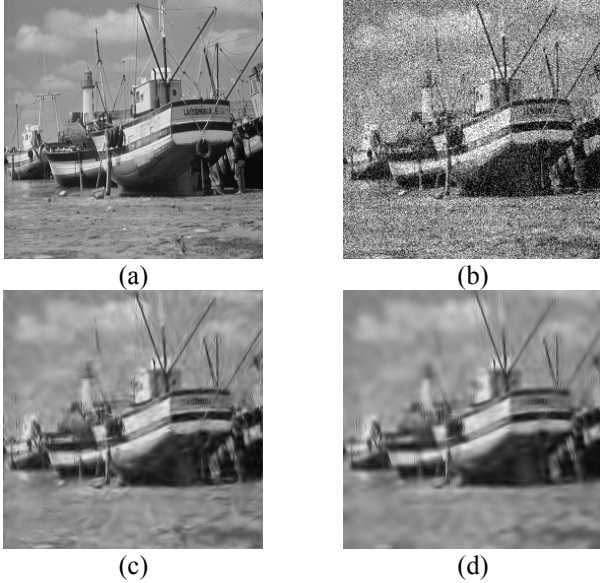


Fig. 3 (a)-(b) original and noisy image with $\sigma=35$; (c) DTCWT-salienShrink; (d) DTCWT-BiShrink

at higher level of noises as can be seen by the tables. The improvement in visual quality is obvious for most cases. Figs. 2, 3 show the results for two images and Fig. 4 shows 2 selected parts of the denoised version of image “boat” to illustrate further the gain in visual quality obtained by proposed method. Even when the PSNR difference between the two methods is relative low (e.g. “house”, Table 1), salienShrink results in more crispy edges and less blurring (Fig. 2).

4. CONCLUSIONS

In this paper, we presented a computationally simple approach to improve common image denoising algorithms based on the wavelet transform. A saliency map is computed for each of the wavelet subbands and a WTA network selects the most salient regions, which are then used to bias the estimate of noise and therefore preserve as much as possible the important coefficients corresponding to pure signal. Detailed experiments are presented that

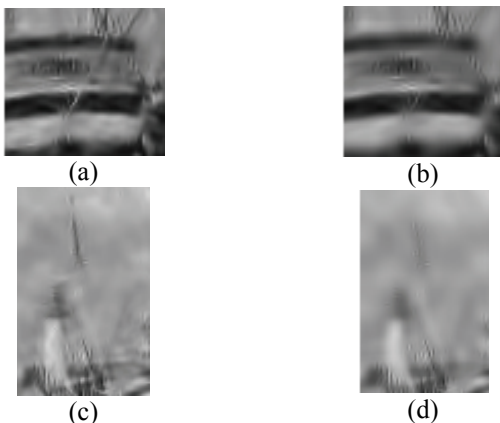


Fig. 4 Details of the denoised “boat” image; (a),(c) DTCWT-BiShrink; (b),(d) DTCWT-salienShrink

prove the improvements both in PSNR and visual quality of the denoised images. Future work will be mainly focused on evaluating the effect of the size of the inhibition area and an automatic way to determine the stopping criterion both involved in the WTA process.

4. REFERENCES

- [1] D.L. Donoho and I.M. Johnstone, “Ideal spatial adaptation via wavelet shrinkage”, *Biometrika*, vol. 81, pp. 425–455, 1994.
- [2] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, “A joint inter- and intra scale statistical model for Bayesian wavelet based image denoising”, *IEEE Trans. Image Process.*, vol. 11, no. 5, pp. 545–557, May 2002.
- [3] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, “Adaptive Wiener denoising using a Gaussian scale mixture model”, *ICIP*, 2001.
- [4] G.Y. Chen, T.D. Bui, and A. Krzyzak, “Image denoising using neighbouring wavelet coefficients”, *ICASSP’04*, vol. 2, pp. 917-920, 2004.
- [5] F. Faghieh, and M. Smith, “Combining spatial and scale-space techniques for edge detection to provide a spatially adaptive wavelet-based noise filtering algorithm”, *IEEE Trans. Image Process.*, vol. 11, no. 9, pp. 1062–1071, Sep 2002.
- [6] L. Sendur and I. W. Selesnick, “Bivariate shrinkage functions for wavelet-based denoising, exploiting interscale dependencies”, *IEEE Trans. on Signal Processing*, vol. 50, no. 11, pp. 2744-2756, Nov 2002.
- [7] L. Sendur and I. W. Selesnick, “Bivariate shrinkage with local variance estimation”, *IEEE Signal Processing Letters*, vol. 9, no. 12, pp. 438-441, Dec 2002. <http://taco.poly.edu/WaveletSoftware/denoise2.html>
- [8] E.P. Simoncelli, “Modeling the joint statistics of images in the wavelet domain”, *Proc. SPIE*, vol. 3813, pp. 188-195, 1999.
- [9] D.L. Donoho and I.M. Johnstone, “Adapting to unknown smoothness via wavelet shrinkage”, *J. Amer. Statist. Assoc.*, vol. 90, no. 432, pp. 1200–1224, 1995.
- [10] S. Chang, B. Yu, and M. Vetterli, “Adaptive wavelet thresholding for image denoising and compression”, *IEEE Trans. Image Processing*, vol. 9, pp. 1532–1546, Sep 2000.
- [11] J. Tsotsos, S. Culhane, W. Yan Kei Wai, Y. Lai, N. Davis, and F. Nuflo, “Modeling visual attention via selective tuning”, *Artificial Intelligence*, vol. 78, pp. 507--545, 1995
- [12] L. Itti, C. Koch, E. Niebur, “A Model of Saliency-Based Visual Attention for Rapid Scene Analysis”, *IEEE Trans. on Patt. Analysis and Mach. Intell.*, vol. 20, no. 11, pp. 1254-1259, Nov 1998
- [13] N.G. Kingsbury, “Complex wavelets for shift invariant analysis and filtering of signals”, *Applied Computational Harmonic Anal.*, vol. 10, no. 3, pp. 234-253, May 2001