ELectronics EXpress

IEICE Electronics Express, Vol.7, No.18, 1409–1415

A wavelet-domain non-parametric statistical approach for image denoising

Jing Tian[1a)], **Li Chen**[1b)], **and Lihong Ma**[2c)]

[1] *School of Computer Science and Technology, Wuhan University of Science and Technology, P. R. China, 430081*

[2] *Guangdong Key Lab. of Wireless Network and Terminal, School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China*

a) *jingtian@ieee.org*

b) *chenli@ieee.org*

c) *eelhma@scut.edu.cn (correspondence)*

Abstract: The challenge of conventional parametric model-based wavelet image denoising approaches is that the efficiency of these methods greatly depends on the accuracy of the prior distribution used for modeling the wavelet coefficients. To tackle the above challenge, a *non-parametric* statistical model is proposed in this paper to formulate the marginal distribution of wavelet coefficients. The proposed non-parametric model differs from conventional parametric models in that the proposed model is automatically adapted to the observed image data, rather than imposing an assumption about the distribution of the data. Furthermore, the proposed non-parametric model is incorporated into a Bayesian inference framework to derive a *maximum a posterior* (MAP) estimation-based image denoising approach. Experiments are conducted to not only demonstrate that the proposed non-parametric statistical model is more suitable than conventional models to formulate the marginal distribution of wavelet coefficients, but also show that the proposed image denoising approach outperforms the conventional approaches.

Keywords: image denoising, statistical modeling, wavelet

Classification: Science and engineering for electronics

References

[1] Q. Guo and S. Yu, "Image denoising using a multivariate shrinkage function in the curvelet domain," *IEICE Electron. Express*, vol. 7, pp. 126–131, 2010.

[2] H. Rabbani, "Image denoising in steerable pyramid domain based on a local Laplace prior," *Pattern Recognition*, vol. 42, pp. 2181–2193, 2009.

DOI: 10.1587/elex.7.1409
Received August 27, 2010
Accepted August 30, 2010
Published September 25, 2010

IEICE Electronics Express, Vol.7, No.18, 1409–1415

[3] S. G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. Image Process.*, vol. 9, pp. 1522–1531, 2000.
[4] I. Prudyus, S. Voloshynovskiy, and A. Synyavskyy, "Wavelet-based MAP image denoising using provably better class of stochastic i.i.d. image models," *Proc. Int. Conf. Telecommun.*, Modern Satellite, Cable and Broadcasting Service, pp. 583–586, 2001.
[5] A. Achim and E. E. Kuruoglu, "Image denoising using bivariate α-stable distributions in the complex wavelet domain," *IEEE Signal Process. Lett.*, vol. 12, pp. 17–20, 2005.
[6] M. K. Mihcak, I. Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Process. Lett.*, vol. 6, pp. 300–303, 1999.
[7] L. Kaur, S. Gupta, and R. C. Chauhan, "Image denoising using wavelet thresholding," *Proc. Int. Conf. Computer Vision, Graphics and Image Process.*, pp. 1–4, 2002.
[8] L. Sendur and I. W. Selesnick, "Bivariate shrinkage with local variance estimation," *IEEE Signal Process. Lett.*, vol. 9, pp. 438–441, 2002.
[9] A. Pizurica and W. Philips, "Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising," *IEEE Trans. Image Process.*, vol. 15, pp. 654–665, 2006.
[10] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Interscience, 2000.
[11] L. Wasserman, *All of Statistics*, Springer, 2004.
[12] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Amer. Statist. Assoc.*, vol. 90, pp. 1200–1224, 1995.
[13] F. Luisier, T. Blu, and M. Unser, "A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding," *IEEE Trans. Image Process.*, vol. 16, pp. 593–606, 2007.

1 Introduction

Images are often corrupted with noise during image acquisition and image transmission. Wavelet-based algorithms that exploit parametric models have been proved to be effective for tackling the image denoising problem [1]. Their basic idea is to perform wavelet decomposition on the input noisy image, then estimate the noise-free wavelet coefficients by employing a Bayesian estimator, which is developed by using a suitable *probability density function* (pdf) as a prior for modeling the wavelet coefficients of the image. Finally, the denoised image is reconstructed by performing an inverse wavelet transform. For that, several statistical models have been developed in the literature, which include Gaussian model, Laplacian model [2], generalized Gaussian model [3], Student-t model [4], α-stable model [5]. However, the performance of the above-mentioned parametric model-based approaches greatly depends on the effectiveness of the prior pdf for modeling the wavelet coefficients.

In contrast to aforementioned conventional *parametric* model-based approaches, a *non-parametric* statistical model is proposed in this paper to

DOI: 10.1587/elex.7.1409
Received August 27, 2010
Accepted August 30, 2010
Published September 25, 2010

IEICE Electronics Express, Vol.7, No.18, 1409–1415

formulate the marginal distribution of wavelet coefficients. Since the proposed non-parametric wavelet coefficients model automatically adapts to the observed image data, it is expected to yield superior performance to that of the conventional approaches using parametric models that are fixed in advance. Furthermore, a *maximum a posteriori* (MAP) estimation-based image denoising approach is derived by incorporating the proposed model into a Bayesian inference framework.

The paper is organized as follows. Section 2 proposes a non-parametric statistical model to formulate the distribution of wavelet coefficients, followed by the derivation of the proposed MAP estimation-based image denoising approach. Experimental results are presented in Section 3 to justify that the proposed non-parametric statistical model is more suitable than conventional models to formulate the marginal distribution of wavelet coefficients. Furthermore, experimental results are provided to show that the proposed MAP estimation-based image denoising algorithm outperforms the conventional algorithms. Finally, Section 4 concludes this paper.

2 Proposed image denoising approach

A statistical approach is exploited in this paper to perform image denoising. Given the noisy wavelet coefficient of the noisy image (denoted as y_i, where i is the index), the aim is to recover the noise-free wavelet coefficient (denoted as s_i) via its MAP estimator (denoted as $\hat{s}_i$) as

$$\hat{s}_i = \arg\max p(s_i|y_i). \tag{1}$$

According to the Bayes rule, (1) can be rewritten as

$$p(s_i|y_i) = \frac{p(s_i, y_i)}{p(y_i)} \propto p(s_i, y_i) = p(y_i|s_i)p(s_i). \tag{2}$$

The formulation of (2) boils down to the formulations of its two product terms $p(y_i|s_i)$ and $p(s_i)$, respectively.

Firstly, since the noise is assumed to be additive white Gaussian noise with a zero mean and a standard deviation σ_n, the term $p(y_i|s_i)$ is formulated as [7, 6, 8, 9]

$$p(y_i|s_i) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(y_i - s_i)^2}{\sigma_n^2}}. \tag{3}$$

Secondly, to formulate the term $p(s_i)$, a non-parametric statistical model is proposed in this paper using the *kernel density estimation* (KDE) technique [10]. The pdf $p(s_i)$ is determined by the summation of a number of kernel functions, which are centered at each observed noisy coefficient in the neighborhood Ω_i at the position s_i. Furthermore, in this paper, the above-mentioned kernel function is suggested to be a Gaussian function with a mean, which is the observed noisy coefficient (say, y_j), and a standard deviation h_j, which is the standard deviation of all coefficients covered by a 7×7 square window Ω_i. To summarize, the pdf $p(s_i)$ is defined as

DOI: 10.1587/elex.7.1409
Received August 27, 2010
Accepted August 30, 2010
Published September 25, 2010

ELectronics EXpress

IEICE Electronics Express, Vol.7, No.18, 1409–1415

$$p(s_i) \propto \sum_{y_j \in \Omega_i} \frac{1}{\sqrt{2\pi} h_j} e^{-\frac{(s_i - y_j)^2}{2h_j^2}}. \tag{4}$$

The proposed non-parametric model (4) is fundamentally different from the conventional parametric models, since the proposed model automatically adapts to the observed image data (through parameters y_j and h_j in (4)). Therefore, the proposed non-parametric model is potential to outperform conventional parametric models for formulating the distribution of wavelet coefficients; that will be verified in Section 3.

The next goal is to incorporate the proposed wavelet coefficients model (4) into the Bayesian inference framework (2) to derive the MAP estimator of the noise-free coefficient. Since the proposed model (4) consists of a number of components (i.e., kernel functions), our idea is to use each component of (4) as the prior model to calculate its corresponding MAP estimator of the noise-free coefficient, and then combine all MAP estimators to be the final estimator of the noise-free coefficient.

To be more specific, substituting (3) and one component (say j-th) of (4) into (2), we get

$$p(s_i|y_i) \propto \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(y_i - s_i)^2}{\sigma_n^2}} \frac{1}{\sqrt{2\pi} h_j} e^{-\frac{(s_i - y_j)^2}{2h_j^2}}. \tag{5}$$

By setting the derivative of (5) to be zero with respect to s_i, we can obtain the MAP estimator of the noise-free coefficient (denoted as $\hat{s}_i^j$) as

$$\hat{s}_i^j = \frac{\sigma_n^2 y_j + h_j^2 y_i}{\sigma_n^2 + h_j^2}. \tag{6}$$

Finally, the estimated noise-free coefficient can be obtained by averaging all MAP estimators, which are obtained by using each component of (4) as the prior image model, respectively; that means

$$\hat{s}_i = \frac{1}{|\Omega_i|} \sum_{y_j \in \Omega_i} \hat{s}_i^j = \frac{1}{|\Omega_i|} \sum_{y_j \in \Omega_i} \frac{\sigma_n^2 y_j + h_j^2 y_i}{\sigma_n^2 + h_j^2}, \tag{7}$$

where $|\Omega_i|$ represents the number of coefficients in the neighborhood Ω_i.

3 Experimental results

Experiments are conducted to explore the performance of the proposed approach using three test images, *Barbara*, *Window* and *Lighthouse*, which are downloaded from http://www.site.uottawa.ca/ edubois/demosaicking. These test images serve as the ground truth and compared with the denoised images for performance comparison. The noisy images are generated by adding the ground truth image with an additive white Gaussian noise with a zero mean and a standard deviation σ_n, respectively.

The proposed approach first performs a 2-D discrete wavelet decomposition (a five-level decomposition using a *Daubechies*'s wavelet with eight vanishing moments) on a noisy image to get the noisy wavelet coefficients. The

DOI: 10.1587/elex.7.1409
Received August 27, 2010
Accepted August 30, 2010
Published September 25, 2010

wavelet decomposition is implemented via a five-level decomposition using a *Daubechies*'s wavelet with eight vanishing moments. Then, the proposed approach uses (7) to estimate each noise-free coefficient excluding those of the LL subband. Finally, the inverse wavelet transform is applied to obtain the denoised image.

The first experiment is to justify the proposed non-parametric statistical model by comparing it with other five models: Gaussian model, Laplacian model [2], generalized Gaussian model [3], Student-t model [4], α-stable model [5], using three statistical metrics. The smaller metrics values indicate better performance, and their definitions are presented as follows. The first criterion is *Chi-square* (C-S) metric [11]

$$d_{cs} = \sum_{w \in \mathbf{R}} \frac{p_h(w) - p_e(w)}{p_h(w)}, \tag{8}$$

where $p_h(w)$ and $p_e(w)$ denote prior and empirical pdfs, respectively. The second criterion is *Kolmogorov-Smirnov* (K-S) metric [11]

$$d_{ks} = \max_{w \in \mathbf{R}} |F_h(w) - F_e(w)|, \tag{9}$$

where $F_h(w)$ and $F_e(w)$ denote the *cumulative density function* (cdf) of the prior pdf and the empirical cdf, respectively. The third criterion is *Kullback-Liebler* (K-L) metric [11]

$$d_{kl} = \sum_{w \in \mathbf{R}} p_h(w) \ln\left(\frac{p_h(w)}{p_e(w)}\right), \tag{10}$$

where $p_h(w)$ and $p_e(w)$ denote prior and empirical pdfs, respectively.

The *Barbara* image is applied on a two-level decomposition using a Daubechies's wavelet with eight vanishing moments. Then, the distributions of the second-level detail subbands are fitted with the above-mentioned models, which are further compared with the histogram of the wavelet coefficients in the sense of the above-mentioned two statistical metrics. The comparison is presented in Table I, where the proposed model is more consistent to the histogram of the wavelet coefficients than other five conventional models, by providing the best objective statistical performance.

The second experiment is to compare the proposed approach with other seven denoising methods [12, 7, 3, 6, 8, 13, 9], which are implemented using

Table I. Objective comparison of various models for formulating the distribution of wavelet coefficients.

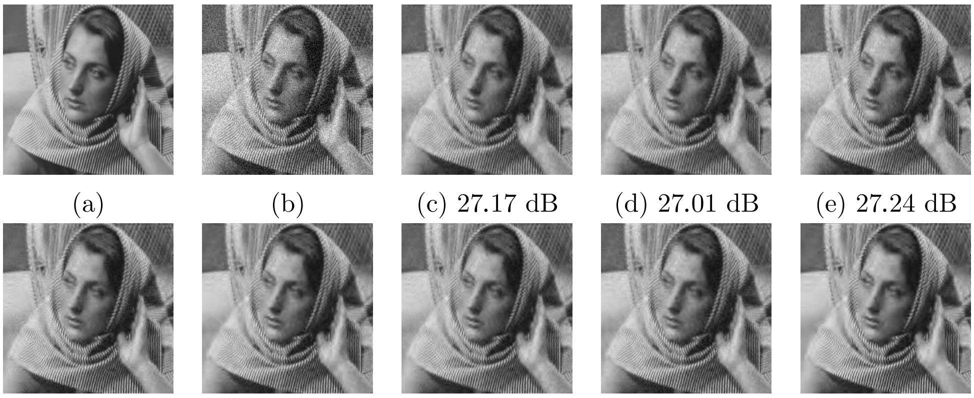
Criterion	Subband	Gaussian model	Laplacian model [2]	Generalized Gaussian model [3]	Student-t model [4]	α-stable model [5]	Proposed model (4)
C-S	HL2	0.6569	0.1300	0.0149	0.2553	0.0316	**0.0012**
metric	LH2	0.6063	0.2089	0.0313	0.4009	0.0630	**0.0014**
(8)	HH2	0.9893	0.3751	0.0455	0.1602	0.0283	**0.0011**
K-S	HL2	0.1558	0.0633	0.0212	0.0822	0.0154	**0.0099**
metric	LH2	0.1568	0.0815	0.0235	0.1143	0.0285	**0.0031**
(9)	HH2	0.1948	0.1117	0.0376	0.0869	0.0212	**0.0039**
K-L	HL2	0.3832	0.0869	0.0084	0.1031	0.0150	**0.0008**
metric	LH2	0.3705	0.1133	0.0160	0.1603	0.0288	**0.0008**
(10)	HH2	0.5669	0.2163	0.0233	0.0762	0.0131	**0.0006**

DOI: 10.1587/elex.7.1409
Received August 27, 2010
Accepted August 30, 2010
Published September 25, 2010

IEICE Electronics Express, Vol.7, No.18, 1409–1415

Table II. The PSNR performance comparison of various image denoising approaches.

Method	*Barbara*		*Window*		*Lighthouse*	
	$\sigma_n = 10$	$\sigma_n = 20$	$\sigma_n = 10$	$\sigma_n = 20$	$\sigma_n = 10$	$\sigma_n = 20$
Ref. [12]	30.98	27.11	30.36	26.10	32.03	28.03
Ref. [7]	30.97	27.01	30.15	25.90	31.98	27.92
Ref. [3]	30.85	27.03	30.30	26.08	31.77	27.85
Ref. [6]	32.49	28.14	31.20	26.74	33.14	28.70
Ref. [8]	32.06	28.10	30.81	26.58	33.16	29.25
Ref. [13]	31.55	27.73	30.58	26.56	32.91	28.92
Ref. [9]	31.48	27.64	29.92	26.26	32.88	29.00
Proposed approach	32.58	28.57	31.20	26.84	33.30	29.30



(a) (b) (c) 27.17 dB (d) 27.01 dB (e) 27.24 dB

(f) 28.02 dB (g) 28.04 dB (h) 27.67 dB (i) 27.75 dB (j) 28.46 dB

Fig. 1. A close-up comparison of various results of *Barbara*: (a) ground truth; (b) noisy image ($\sigma_n = 20$); (c)-(i) are results of Ref. [12, 7, 3, 6, 8, 13, 9], respectively; and (j) proposed method that outperforms the above seven methods.

programs provided by their respective authors online. Table II presents the PSNR performance comparison, where the output PSNRs have been averaged over ten noise realizations. Due to limited space of this paper, a close-up comparison of various denoised images of *Barbara* is presented in Figure 1. As seen from the above Table and Figure, the proposed approach always outperforms seven conventional approaches.

4 Conclusions

A new image denoising algorithm using a non-parametric statistical model of wavelet coefficients has been proposed in this paper. The proposed non-parametric statistical model is more suitable than conventional models to formulate the marginal distribution of wavelet coefficients, as shown in our experiments using three statistical metrics. Furthermore, the proposed MAP estimation-based image denoising algorithm, which exploits the proposed non-parametric model as the prior image model, has been shown to be superior to conventional algorithms, as verified in our experiments.

DOI: 10.1587/elex.7.1409
Received August 27, 2010
Accepted August 30, 2010
Published September 25, 2010

IEICE Electronics Express, Vol.7, No.18, 1409–1415

Acknowledgments

This work was supported by National Natural Science Foundation of China (Grant No. 60705012, 60972133).

DOI: 10.1587/elex.7.1409
Received August 27, 2010
Accepted August 30, 2010
Published September 25, 2010