Research Journal of Applied Sciences, Engineering and Technology 4(18): 3339-3343, 2012
ISSN: 2040-7467

Submitted: March 03, 2012 Accepted: May 18, 2012 Published: September 15, 2012

Bayesian Image Denoising by Local Singularity Detection

[1]Yanqiu Cui, [2]Tao Zhang and [1]Shuang Xu
[1]College of Information and Communication Engineering,
[2]College of Electromechanical and Information Engineering, Dalian Nationalities University
Dalian, 116600, China

Abstract: In this study, we present a wavelet-based method for removing noise from images and a Bayesian shrinkage factor was derived to estimate noise-free wavelet coefficients. This method took into account dependencies between wavelet coefficients. The interscale dependencies were measured from the local singularity and a conditional probability model was proposed. The intrascale dependencies were measured from the spatial clustering properties and a prior probability model was used. Based on these models in a Bayesian framework, each coefficient was modified separately. Experimental results demonstrate this method improves the denoising performance and preserves the details of the image.

Keywords: Image denoising, singularity, wavelet transform

INTRODUCTION

In image denoising, a number of methods based on fthe wavelet transform are proposed. The image can be represented in a sparse way by a small number of large wavelet coefficients. Due to this property of wavelet transform, additive noise can be removed efficiently by simple thresholding of the wavelet coefficients (Donoho and Johnstone, 1995). In the method of (Donoho and Johnstone, 1995), the wavelet coefficients with an absolute value below a threshold are replaced by zero. The other coefficients are reduced in absolute value. In the approach of Xu, Weaver and collaborators (Xu *et al*., 1994), the criterion to distinguish noise from meaningful signal is based on the observation that wavelet coefficients of noise have a much weaker correlation between scales than the coefficients of a noiseless image. The high-frequency contents are mostly suppressed, except where a highly correlated feature is detected. The method of Mallat and collaborators (Mallat and Hwang, 1992) is based on the assumption that the noiseless image is regular and the noise irregular. It exploits the fact that a function local regularity parameters can be derived from its wavelet coefficients.

The method based on the wavelet transform in a Bayesian framework is shown to be more effective for noise suppression (Yin *et al*., 2011; Firoiu *et al*., 2011; Hashemi and Beheshti, 2011), because the statistic property of wavelet coefficients is taken into account. A large class of Bayesian wavelet-based denoising methods approximates the wavelet coefficients as mutually independent. Advanced wavelet based denoising methods take into account dependencies between wavelet coefficients. The method of Maurits (Malfait and Roose, 1997) combines the inter-and intrascale dependencies. A bilevel Markov Random Field (MRF) model is used to encode prior knowledge about spatial clustering of wavelet coefficients. The interscale dependencies between wavelet coefficients are encountered via interscale ratios. The statistical properties of these significance measures are expressed in a conditional probability density model and combined with the prior model in a Bayesian framework; hence, the name eometrical Bayesian approach The authors of (Jansen and Bultheel, 2001) motivate the whole approach theoretically and develop practical algorithms from it. However, they only consider the magnitude of a wavelet coefficient as its significance measure and a different prior model is proposed. Aleksandra (Pižurica *et al*., 2002) further extend the geometrical Bayesian approach. A joint significance measure and an anisotropic MRF prior model are developed (Malfait and Roose, 1997; Jansen and Bultheel, 2001; Pižurica *et al*., 2002).

In this study, a new significance measure and its conditional density model are introduced in the geometrical Bayesian approach. The local singularity is used to measure the interscale dependencies between the wavelet coefficients. Its new conditional model is constructed. Experiment results demonstrate this method improves the denoising performance quantitatively and qualitatively.

Wavelet transform: Instead of the classical fast wavelet transform, we use the so-called non-decimated wavelet transform, which is also called redundant or stationary wavelet transform

In a wavelet decomposition of an image, a wavelet coefficient $w_{j,l}^{D}$ represents its bandpass content at.

Corresponding Author: Yanqiu Cui, College of Information and Communication Engineering, Dalian Nationalities University Dalian, 116600, China

resolution scale 2^j, spatial position l and orientation D. Typically, three orientation subbands are used, leading to three detail images at each scale, characterized by horizontal, vertical and diagonal directions. Whenever there can be no confusion, we omit the indices that denote the scale and the orientation. A given detail image is represented as vector w = $\{w_1, \ldots, w_n\}$. The set of indices L = $\{1, \ldots, n\}$ is a set of pixels on a regular rectangular lattice. In the spirit of early literature, we assign a measure of significance m_l and a binary label x_l to each wavelet coefficient w_l. The value $x_l = 1$ denotes that w_l is a "significant" coefficient and means that the corresponding noise-free coefficient is certainly larger than the noise deviation; the label value $x_l = -1$ denotes that w_l represents mainly noise. A set m = $\{m_1, \ldots, m_n\}$ is called significance map and m = |w|. A set x = $\{x_1, \ldots, x_n\}$ is called mask. We assume that the wavelet coefficients are corrupted with additive Gaussian white noise. An observed wavelet coefficient is given by $w_l = y_l + n_l$, where y_l is the noise-free wavelet coefficient and n_l is the noise coefficient.

PROPOSED METHODOLOGY

Geometrical bayesian approach: We shall extend the geometrical Bayesian approach with a new conditional density model. The geometrical Bayesian approach proceeds according to the following scheme.

- The wavelet decomposition of the image is computed.
- The obtained wavelet coefficients are modified. Each wavelet coefficients is multiplied with the marginal probability that is noise-free, given the computed significance map:

$$\hat{y}_l = P\left(X_l = 1 \middle| M = m\right) w_l \tag{1}$$

- The resulting image is reconstructed from the modified coefficients.

The exact computation of the marginal probability $P(X_l = 1|M = m)$ requires the summation of the posterior probabilities $P(X = x|M = m)$ of all possible configurations x for which $x_l = 1$. Since this is an intractable task, one typically alleviates it by using a relatively small, but epresentative subset of all possible configurations. Such a representative subset is obtained via an importance-type sampling: the probability that a given mask is sampled is proportional to its posterior probability. An estimate of $P(X_l = 1|M = m)$ is then obtained by computing the fractional number of all sampled masks for which $x_l = 1$. We apply the Metropolis sampler. An appropriate threshold generates the initial mask. From each configuration x, it generates a new, andidate configuration x^c by switching the binary label at a randomly chosen position l. The decision about accepting the change is based on the ratio p of the posterior probabilities of the two configurations. After applying the Bayesian rule $P(X = x|M = m) = P(X = x) f_{Mx}(m|x) f_M(m)$, assuming the conditional independence $f_{Mx}(m|x) = \prod_{l\in L} fM_l \left| X_l\left(m_l \middle| x_l\right)\right.$ and a MRF prior model form (3), the ratio p reduces to:

$$p = \frac{fM_l \left| X_l\left(m_l \middle| x_l^c\right)\right.}{fM_l \left| X_l\left(m_l \middle| x_l\right)\right.} \exp\left(V_{N_{l,D}}(x) - V_{N_{l,c}}\left(x^c\right)\right) \tag{2}$$

If p>1, the local change is accepted, because it has produced a mask with a higher posterior probability. If p<1, the change is accepted with probability p. After all labels in mask have been updated in this way, one iteration is completed. Ten such iterations suffice to estimate the marginal probabilities, provided that the initial mask is well chosen (Jansen and Bultheel, 2001).

MRF prior model: Let L/l denote the set of all pixels in except itself. The Markov property of a random field is:

$$P(X_l = x_l | X_{L/l} = x_{L/l}) = P(X_l = x_l | X\partial_{(l)} = x\partial_{(l)})$$

where, $\partial(l)$ is the neighborhood of the pixel l. Most often used are the first-order neighborhood (four nearest pixels) and the second-order neighborhood (eight nearest pixels). A set of pixels, which are all neighbors of one another is called a clique. The joint probability $P(X = x)$ of a MRF is a special case of the Gibbs distribution, exp[-H(x)]/z with partition constant Z, where the energy H(x) can be decomposed into contributions of clique potentials $V_c(x)$ over all possible cliques. The clique potential $V_c(x)$ is a function of only those labels x_l, for which $l \in C$. In practice, one chooses the appropriate clique potential functions to give preference to certain local spatial dependencies, e.g., to assign higher prior probability to edge continuity. A simple MRF model is used in this study that takes the form:

$$P(X = x) = \frac{1}{Z} \exp\left(-\sum_{l\in L} V_{Nl}(x)\right) \tag{3}$$

$$V_{Nl}(x) = -r \sum_{k\in N_l} x_l x_l \tag{4}$$

where, r is a positive constant. This model is the Ising model.

New conditional density model: For the geometrical Bayesian approach, the choice of the significance measure and the characterization of its conditional densities are very important. To distinguish between useful edges and noise, the significance measure m_l is defined as:

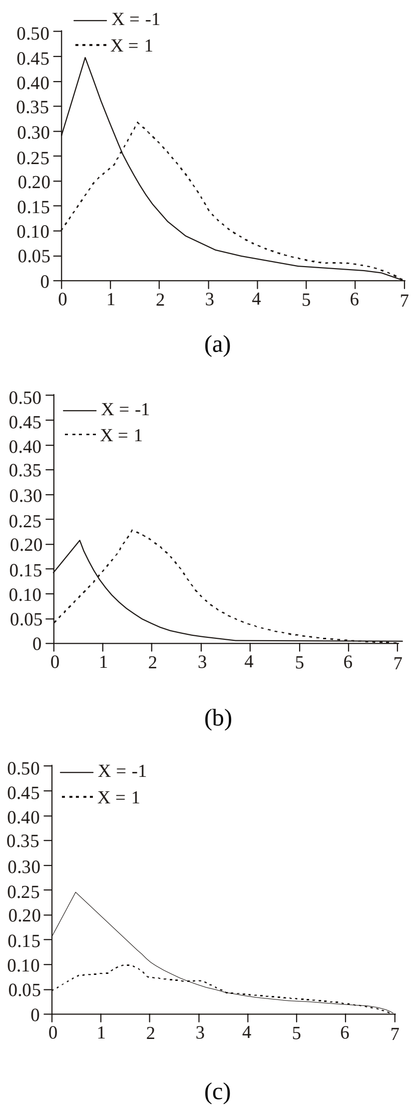

Fig. 1: Conditional densities of inerscale ratio, (a) interscale ratio of the magnitudes, (b) interscale ratio of the local maxima and (c) interscale ratio of the local non-maxima

$$m_l = \begin{cases} \left| \dfrac{\max_{i \in S(j+1,l)} |w_{j+1,i}|}{w_{j,l}} \right| \approx 2^{\alpha} & |w_{j,l}| > |w_{j,l} + 1| \, and \, |w_{j,l}| > |w_{j,l} - 1|, \\ \left| \dfrac{w_{j+1,l}}{w_{j,l}} \right|, & Otherwise \end{cases} \quad (5)$$

where, α is the local Lipschitz exponent, s(j, l) defines as the cone of influence of the point l.

This significance measure is used to estimate roughly the local singularity, which is different to useful edges and noise. The restrictive condition of local maximum value can be more accurate to estimate the singularity, which can more effectively distinguish between useful signal and noise. This can be verified by the statistical properties of inerscale ratio of wavelet coefficients. The conditional densities of inerscale ratio computed from the standard image ouse is illustrated in Fig. 1. With respect to Fig. 1a, the overlap of the two conditional probabilities in Fig. 1b is smaller, so the interscale ratio of local maxima provides more information for distinction between useful signal and noise. Most of the two conditions probability in Fig. 1c overlap and therefore the interscale ratio of the local non-maxima provides less information for distinguishing between useful signal and noise. In this study, the local maximum and local non-maximum are modeled separately, that is more effective to separate the useful signal and noise. But realistic densities of interscale ratio from experiments are not convenient to practical applications, a simplified heuristic model is proposed in this study, which is illustrated in Fig. 2.

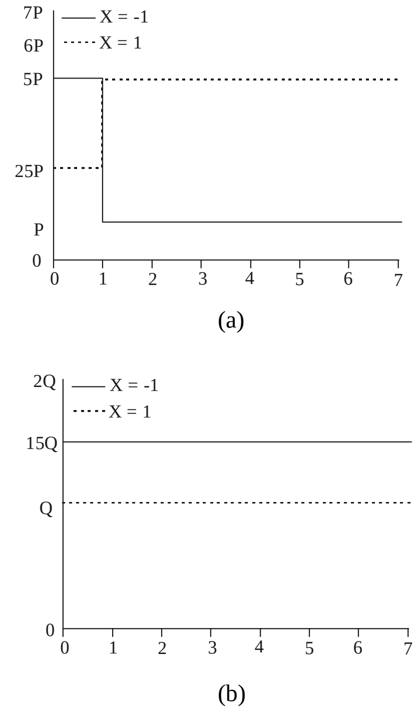

Fig. 2: Simplified conditional densities of inerscale ratio, (a) interscale ratio of the local maxima and (b) interscale ratio of the local non-maxima

EXPERIMENTAL RESULTS

The performance of the method with the new conditional density model will be illustrated with a quantitative and a qualitative performance measure. The

Fig. 3: Original house image (left) and the same image with artifical additive noise (SNR = 9 dB, right)

(a) The proposed method (b) MGB method

(c) GCV method (d) HMT method

Fig. 4: Result of noise suppression for noisy house image

qualitative measure is the visual quality of the resulting image. The Signal-to-Noise (SNR) is used as the quantitative measure. It is expressed in dB units and defined as:

$$SNR = 10 \log_{10}(P_{signal}/P_{noise})$$

where, P_{signal} is the variance of the image and P_{noise} is the variance of the noise.

The performance will be illustrated on the 256×256 House image with artificial noise, which is shown in Fig. 3. This image is mainly characterized by sharp edges and flat background. We apply a non-decimated wavelet transform using the variation on the CDF-(spline)-filters "with less dissimilar lengths" like in (Jansen and Bultheel, 2001). Primal and dual wavelets have four vanishing moments. These wavelets are rather popular in image processing. With this transform, the input image is decomposed over four levels. The stochastic sampling procedure is active on the finest three levels only. For comparison, the improvements in terms of SNR of our method and the three related methods are summarized in Table 1. All denoising methods obviously achieve a higher gain in SNR when the input image is noisier, but the method with the new conditional density model performs better than the other related methods.

Table 1: Comparison of quantitative result for different methods in SNR (dB)

Input	SNR = 3	SNR = 6	SNR = 9
The proposed method	14.408	15.793	17.484
MGB method (Jansen and Bultheel, 2001)	12.708	14.297	16.072
GCV method (Jansen *et al*., 1997)	13.273	14.863	16.494
HMT method (Romberg *et al*., 2001)	13.331	14.945	16.639

However, it is not sufficient to rely on quantitative measures only, since they cannot take into account all aspects of image quality for which the human eye is sensitive. The result of the new denoising algorithm is shown in Fig. 4. Image features such as the edges of roof and windows, the shadow contours, the water conduit and the chimney are well retained. For comparison, Fig. 4 also shows the result obtained with the other related methods. A lot of image details are removed compared with the result obtained with the proposed method

.

CONCLUSION

To improve the denoising performance, an image denoising method by detecting the local singularity is proposed in this study. The local singularity is used to measure the interscale dependencies between the wavelet coefficients. Its new conditional model is constructed. Experiment results demonstrate this method improves the denoising performance quantitatively and qualitatively.

ACKNOWLEDGMENT

This study was supported by Fundamental Research Funds for the Central Universities (No. DC110324 and No. DC110309), Foundation of Liaoning Educational Committee (No. L2010093), Research Foundation for Talents of Dalian Nationalities University (No. 20076105) and Research Projects of State Ethnic Affairs Commission (No. 10DL03).

REFERENCES

Donoho, D.L. and I.M. Johnstone, 1995. Adapting to unknown smoothness via wavelet shrinkage. J. Amer. Statist. Assoc., 90: 1200-1224.

Firoiu, I., C. Nafornita, D. Isar and A. Isar, 2011. Bayesian hyperanalytic denoising of SONAR images. IEEE Geosci. Remote Sens. Lett., 8: 1065-1069.

Hashemi, S. and S. Beheshti, 2011. Adaptive image denoising by rigorous Bayesshrink thresholding. IEEE Workshop on Statistical Signal Processing Proceedings, pp: 713-716.

Jansen, M., M. Malfait and A. Bultheel, 1997. Generalized cross validation for wavelet thresholding. Signal Proc., 56: 33-44.

Jansen, M. and A. Bultheel, 2001. Empirical Bayes approach to improve wavelet thresholding for image noise reduction. J. Amer. Statist. Assoc., 96(454): 629-639.

Mallat, S. and W.L. Hwang, 1992. Singularity detection and processing with wavelets. IEEE Trans. Inform. Theory, 38: 617-634.

Malfait, M. and D. Roose, 1997. Wavelet-bsed image denoising using a Markov Random Field a priori model. IEEE Trans. Image Proc., 6: 549-565.

Pižurica, A., W. Philips, I. Lemahieu and M. Acheroy, 2002. A joint inter- and intrascale statistical model for Bayesian wavelet based image denoising. IEEE Trans. Image Proc., 11: 545-557.

Romberg, J.K., H. Choi and R.G. Baraniuk, 2001. Bayesian tree-structured image modeling using wavelet-domain hidden Markov models. IEEE Trans. Image Proc., 10: 1056-1068.

Xu, Y., J.B. Weaver, D.M. Healy and J. Lu, 1994. Wavelet transform domain filters: A spatially selective noise filtration technique. IEEE Trans. Image Proc., 3: 747-758.