ASTROPHYSICAL IMAGE DENOISING USING BIVARIATE ISOTROPIC CAUCHY DISTRIBUTIONS IN THE UNDECIMATED WAVELET DOMAIN

Alin Achim, Diego Herranz and Ercan E. Kuruoğlu

Signals & Images Lab
ISTI-CNR
Via G. Moruzzi 1, 56124 Pisa, Italy
Alin.Achim@isti.cnr.it

ABSTRACT

Within the framework of wavelet analysis, we describe a novel technique for removing noise from astrophysical images. We design a Bayesian estimator, which relies on a particular member of the family of isotropic α-stable distributions, namely the bivariate Cauchy density. Using the bivariate Cauchy model we develop a noise-removal processor that takes into account the interscale dependencies of wavelet coefficients. We show through simulations that our proposed technique outperforms existing methods both visually and in terms of root mean squared error.

1. INTRODUCTION

The study of the intensity fluctuations of the Cosmic Microwave Background (CMB) is one of the milestones of modern Cosmology. The CMB radiation was emitted shortly after the Big Bang and therefore it is regarded as the most ancient image of the Universe. It carries information on the fundamental cosmological parameters, such as the age of the Universe, its geometry, its content of matter and radiation, its equation of state and its future evolution. Currently there are a number of experiments aimed at the detection of this weak radiation. Among them, the satellite missions WMAP (launched in 2002 by NASA) and Planck (to be launched in 2007 by ESA) will allow us to study the CMB with unprecedented sensitivity and resolution.

Detecting the CMB radiation is a very difficult task. First, its intensity is very low. Also, a number of foreground signals as well as instrumental noise appear on CMB images. Among the different foregrounds, extragalactic point sources (EPS) are specially interesting for the understanding of the processes that control galaxy formation and evolution. EPS are distant galaxies that at the angular resolutions of the present generation CMB experiments appear as point-like objects. The study of these astrophysical "foregrounds" has a great scientific value on itself and therefore it is desirable not just to clean CMB data from them but also to separate the different astrophysical foregrounds for their study [1]. Removing instrumental noise from astrophysical images is thus a critical pre-processing step for further image analysis, provided that the individual components are not altered.

Achim was supported by an ERCIM (European Research Consortium in Informatics and Mathematics) fellowship.

Starting with Donoho's work [2], the wavelet transform has become an important tool for recovering signals from noisy data. For the purpose of astrophysical image processing, Sanz *et al.* [3] used a variant of Donoho's soft thresholding approach to separate the noise from the original image. However, wavelet shrinkage techniques have come a long way in recent years due to the better statistical models adopted for the wavelet coefficients. For example, in a number of recent publications [4, 5, 6], it has been shown that alpha-stable distributions, a family of heavy-tailed densities, are sufficiently flexible and rich to appropriately model wavelet coefficients of images in various applications. On the other hand, algorithms that exploit dependencies between coefficients could achieve better results compared with the ones based on the independence assumption [7, 8]. In this context, Sendur and Selesnick [8] have developed a *maximum a posteriori* (MAP) estimator based on a circular-symmetric Laplacian model for a coefficient and its parent.

Our approach for astrophysical image enhancement essentially consists of an wavelet-domain Bayesian denoising processor, which is based on a particular member of the isotropic alpha-stable family of distributions, namely the bivariate Cauchy density function. The outlines of the proposed method are given in the following.

2. BIVARIATE ISOTROPIC STABLE DISTRIBUTIONS

We start with a brief overview of the bivariate stable family of heavy-tailed distributions. Much like univariate stable distributions, the bivariate stable distributions are characterized by the stability property and the generalized central limit theorem [9]. However, they are much more difficult to describe because they form a nonparametric set. An exception is the family of multidimensional isotropic stable distributions.

The bivariate isotropic α-stable distribution is best defined by its characteristic function

$$\varphi(\omega_1, \omega_2) = \exp(\jmath(\delta_1\omega_1 + \delta_2\omega_2) - \gamma|\boldsymbol{\omega}|^\alpha) \qquad (1)$$

where $\boldsymbol{\omega} = (\omega_1, \omega_2)$ and $|\boldsymbol{\omega}| = \sqrt{\omega_1^2 + \omega_2^2}$. Here, α is the *characteristic exponent*, taking values $0 < \alpha \leq 2$ and γ ($\gamma > 0$) is the *dispersion* of the distribution. The parameters δ_1, δ_2 are the location parameters. The distribution

0-7803-8554-3/04/$20.00 ©2004 IEEE.

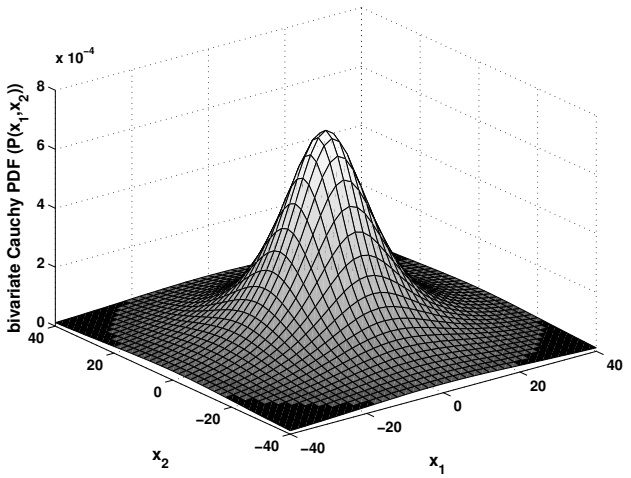

Fig. 1. Two dimensional Cauchy density surface ($\alpha = 1$, $\gamma = 15$).

is isotropic with respect to the point (δ_1, δ_2). Note that the two marginal distributions of the isotropic stable distribution are symmetric α-stable ($S\alpha S$) with parameters $(\delta_1, \gamma, \alpha)$ and $(\delta_2, \gamma, \alpha)$. Since our further developments are in the framework of wavelet analysis, in the following we will assume that $(\delta_1, \delta_2) = (0, 0)$. The bivariate *isotropic Cauchy* and *Gaussian* distributions are special cases for $\alpha = 1$ and $\alpha = 2$, respectively. The bivariate probability density functions (PDF) in these two cases can be written as:

$$P_{\alpha,\gamma}(x_1, x_2) = \begin{cases} \frac{\gamma}{2\pi(x_1^2+x_2^2+\gamma^2)^{3/2}} & \text{for } \alpha = 1 \\ \frac{1}{4\pi\gamma} \exp\left[-\frac{x_1^2+x_2^2}{4\gamma}\right] & \text{for } \alpha = 2. \end{cases} \quad (2)$$

Figure 1 shows an example of a two-dimensional Cauchy surface. The important observation regarding this plot is that although the bivariate Cauchy density behaves approximately like a Gaussian density near the origin, its tails decay at a lower rate compared with a Gaussian density.

As in the case of the univariate $S\alpha S$ density function, when $\alpha \neq 1$ or $\alpha \neq 2$, no closed form expressions exist for the density function of the bivariate stable random variable.

3. THE ALGORITHM

Our approach for image denoising uses the classical 3-steps technique generally employed in wavelet-based methods: i) analysis of raw data by means of wavelet transform, ii) noisy coefficients shrinkage using an appropriately designed algorithm, and iii) synthesis of the denoised image from the processed wavelet coefficients through the inverse wavelet transform. One should note at this point that in our implementation we use a redundant wavelet transform based on Mallat's pyramidal algorithm [10]. Redundant means here that the down-sampling after each stage of the forward transform has been omitted. The advantage of doing so is twofold: First, by using an undecimated decomposition, we are away from unwanted effects like pseudo-Gibbs phenomena. Also, the number of coefficients in every adjacent subbands is kept the same and thus the bivariate analysis is made easier.

The wavelet transform is a linear operation. Consequently, after decomposing an image we get, in each of the three orientations and for every two adjacent levels, sets of noisy wavelet coefficients written as the sum of the transformations of the signal and of the noise:

$$\begin{aligned} y_j^i &= x_j^i + n_j^i \\ y_{j+1}^i &= x_{j+1}^i + n_{j+1}^i \end{aligned} \quad (3)$$

where $1 < j < J$ refers to the decomposition level or scale and $i = 1, 2, 3$ refers to the three spatial orientations. The above set of equations can be written in vectorial form as:

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \quad (4)$$

where $\mathbf{y} = (y_1, y_2)$, $\mathbf{x} = (x_1, x_2)$, $\mathbf{n} = (n_1, n_2)$, and for simplicity we dropped the index i. The MAP estimator of $\mathbf{x}$ given the noisy observation $\mathbf{y}$ can be easily derived as being:

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg\max_{\hat{\mathbf{x}}} P_{\mathbf{x}|\mathbf{y}}(\mathbf{x}\,|\mathbf{y}) \quad (5)$$

Bayes' theorem gives the *a posteriori* PDF of $\mathbf{x}$ based on the measured data:

$$P_{\mathbf{x}|\mathbf{y}}(\mathbf{x}\,|\mathbf{y}) = \frac{P_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\mathbf{x})\, P_{\mathbf{x}}(\mathbf{x})}{P_{\mathbf{y}}(\mathbf{y})}, \quad (6)$$

where $P_{\mathbf{x}}(\mathbf{x})$ is the *prior* PDF of the signal component of the measurements and $P_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\mathbf{x})$ is the *likelihood* function. Substituting (6) in (5), we get:

$$\begin{aligned} \hat{\mathbf{x}}(\mathbf{y}) &= \arg\max_{\hat{\mathbf{x}}} P_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\mathbf{x})\, P_{\mathbf{x}}(\mathbf{x}) = \arg\max_{\hat{\mathbf{x}}} P_{\mathbf{n}}(\mathbf{y}\text{-}\mathbf{x}) P_{\mathbf{x}}(\mathbf{x}) \\ &= \arg\max_{\hat{\mathbf{x}}} P_{\mathbf{n}}(\mathbf{n}) P_{\mathbf{x}}(\mathbf{x}) \end{aligned} \quad (7)$$

In the above equation we use a bivariate Cauchy model for the signal component while we use a zero-mean bivariate Gaussian model for the noise component of the wavelet coefficients. In other words, the observed signal is a mixture of Cauchy signal and Gaussian noise. Figure 2 illustrates the numerically computed MAP surface using these two prior density functions with (7). The behavior of this processor incorporates the main property of the ones proposed in [5, 6], namely that large-amplitude observations are essentially preserved while small-amplitude values are suppressed. In addition, the shrinkage of a coefficient is also conditioned on the value of the corresponding coefficient at the next decomposition level (parent value): The smaller the parent value, the greater the shrinkage [8].

Naturally, in order for the processor in Eq. (7) to be of any practical use, one should be able to estimate the parameters γ_x for the signal and σ_n of the noise from the observed data. For this purpose we propose a method that is based on empirical characteristic functions. Specifically, since the PDF of the measured coefficients ($\mathbf{y}$) is the convolution between the PDFs of the signal ($\mathbf{x}$) and noise components ($\mathbf{n}$), the associated characteristic function of the measurements is given by the product of the characteristic functions of the signal and noise:

$$\varphi_y(\omega) = \exp\left(-\gamma_x|\omega|\right) \cdot \exp\left(-\frac{\sigma_n^2}{2}|\omega|^2\right) \quad (8)$$

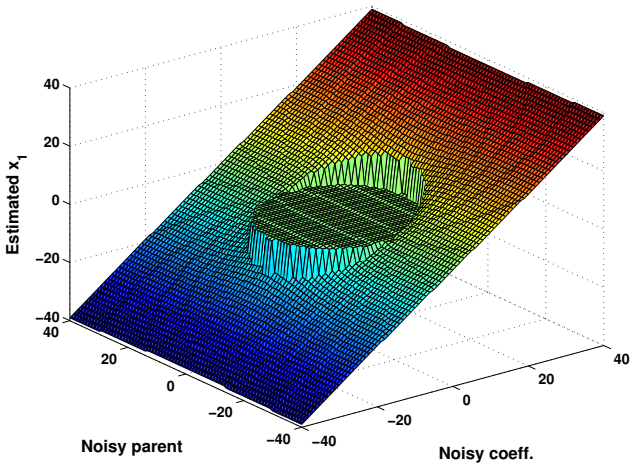

Fig. 2. Numerically computed MAP surface from Eq. (7) for bivariate Cauchy signal ($\alpha_x = 1$, $\gamma_x = 2$) and Gaussian noise ($\alpha_n = 2$, $\gamma_n = 12.5$) prior distributions.

After some straightforward manipulations one gets:

$$\gamma_x = -\frac{\log \varphi_y^2(\omega) - \sigma_n^2 |\omega|^2}{2|\omega|} \tag{9}$$

First, as proposed in [2], a robust estimate of the noise standard deviation, σ_n, is obtained using the median absolute deviation (MAD) of coefficients at the first decomposition level

$$\hat{\sigma}_n = \frac{MAD(y_1)}{0.6745}, \tag{10}$$

Finally, we solve (9) for the parameter γ_x. In principle, γ_x could be estimated using any nonzero value of ω. However, as suggested in [11] in order to reduce the overall variance of the estimate, we chose to average the results from the estimates corresponding to many possible choices of ω.

4. SIMULATION RESULTS

We tested our technique using realistic simulations of CMB, EPS and instrumental noise in accordance with the technical specifications for the future ESA's Planck mission. Planck will observe the CMB sky at nine different frequencies (30, 44, 70, 100, 143, 217, 353, 545 and 857 GHz) with angular resolutions that range from 33 arcmin (for the 30GHz channel) to 5 arcmin (for the 857 GHz channel). The CMB simulations represent the best cosmological model available to-date. The EPS samples were generated using Toffolatti's et. al model [12], which takes into account all the main observational data available up to now. Instrumental noise levels are the ones expected for the Planck detectors.

To quantify the denoising performance of the proposed algorithm we employed the peak signal to noise ratio (PSNR) defined as:

$$PSNR = 20 \log_{10}(256/\sqrt{\frac{1}{N^2}\sum_i (\hat{s}_i - s_i)^2}) \tag{11}$$

Table 1. Image enhancement measures obtained using 2 denoising methods. The PSNR is given in dB.

Channel	Noisy	Soft	Proposed
44GHz	30.06	33.07	33.71
70GHz	24.79	30.92	31.86
100GHz	26.23	30.80	31.99
143GHz	34.81	34.81	35.86
217GHz	25.42	31.92	33.10
353GHz	15.71	24.00	25.27
545GHz	15.55	24.77	26.98
857GHz	22.43	29.12	30.65

where s and $\hat{s}$ denote the noise-free and the denoised images respectively, and N^2 is the total number of pixels.

Before showing results for the application at hand, let us make the following remark: In [8] the authors have applied their technique to the standard "Lena" image for a $\sigma = 25.5$ noise level. They reported a $PSNR = 28.83dB$ for the denoised image. Using our technique and repeating their experiment under the same setting we got a $PSNR = 29.91dB$. Although conceptually similar, we attribute the better performance of our algorithm to the better ability of the Cauchy model in fitting the tails of the underlying distribution of the signal wavelet coefficients on the one hand, and to the redundant wavelet decomposition scheme used in the present work on the other hand.

The results of applying our method to astrophysical images are summarized in Table 1. For comparison we have also included the denoising results obtained using soft thresholding [2, 3]. Both schemes were implemented using Daubechies' length-8 "mother wavelet". From the table it is clear that our proposed Bayesian approach exhibits a very good noise mitigation performance, consistently outperforming the soft thresholding approach.

Remember however, that in astrophysical applications, we are interested in suppressing instrumental noise while at the same time preserving the features of the original image that will be used for further source separation tasks. Having in mind this observation, in Fig. 3 we show for visual comparison a result from the processing of one of our test images. The image corresponds to the 353GHz channel. As it can be seen, the soft thresholding technique achieves a good noise removal performance, but at the expense of an oversmoothed resulting image (see. Fig. 3(c)). Instead, the bivariate Cauchy-based Bayesian processor removes noise equally well retaining also the features that are clearly distinguishable in the noisy data (cf. Fig. 3(d)).

5. CONCLUSIONS

We proposed a new statistical representation for modeling the interscale dependencies of wavelet coefficients in astrophysical images. We designed and tested a MAP processor, based on a bivariate Cauchy prior distribution for the signal, which we found to be more effective than existing methods in terms of noise mitigation performance as well as in preserving fine signal details.

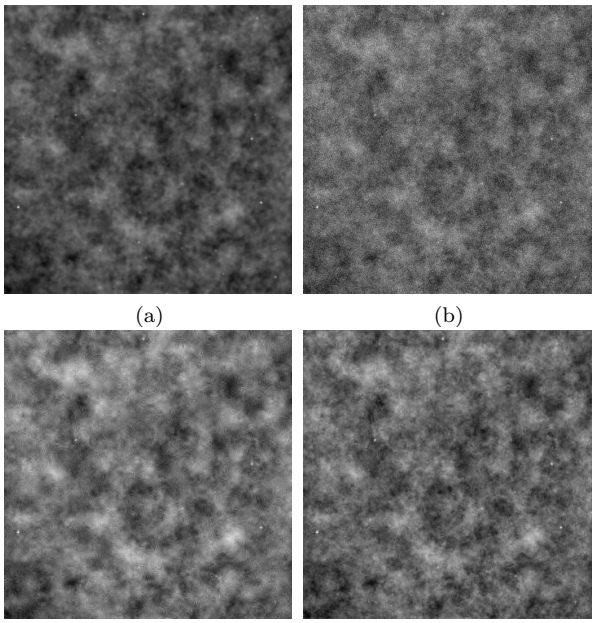

Fig. 3. Results of various denoising methods: (a) Original image; (b) Noisy image ($PSNR = 15.71dB$); (c) Soft thresholding; (d) Proposed method.

We are currently addressing several issues related to the work presented in this paper. A major issue is the extension of the algorithm to the case of general bivariate α-stable distribution. While the development of a method employing isotropic stable densities should be somehow straightforward, considering general α-stable densities would be a much more challenging task, presupposing accurate estimation of the spectral measure from noisy observations.

6. REFERENCES

[1] D. Herranz, E. E. Kuruoğlu, and L. Toffolatti, "Using alpha-stable distributions to model the P(D) distribution of point sources in CMB sky maps," *Astronomy & Astrophysics*, 2004. accepted.

[2] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, Aug. 1994.

[3] J. L. Sanz, F. Argueso, L. Cayon, E. Martinez-Gonzalez, R. B. Barreiro, and L. Toffolatti, "Wavelets applied to cosmic microwave background maps: a multiresolution analysis for denoising," *Mon. Not. R. Astron. Soc.*, vol. 309, pp. 672–680, 1999.

[4] P. Tsakalides, P. Reveliotis, and C. L. Nikias, "Scalar quantization of heavy-tailed signals," *IEE Proc. - Vision, Imag. Sign. Proc.*, vol. 147, pp. 475–484, Oct. 2000.

[5] A. Achim, A. Bezerianos, and P. Tsakalides, "Novel Bayesian multiscale method for speckle removal in medical ultrasound images," *IEEE Trans. Med. Imag.*, vol. 20, pp. 772–783, Aug. 2001.

[6] A. Achim, P. Tsakalides, and A. Bezerianos, "SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling," *IEEE Trans. Geosc. and Remote Sensing*, vol. 41, pp. 1773–1784, Aug. 2003.

[7] J. K. Romberg, H. Choi, and R. G. Baraniuk, "Bayesian tree-structured image modeling using wavelet-domain hidden Markov models," *IEEE. Tran. Image. Proc.*, vol. 10, pp. 1056–1068, July 2001.

[8] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE. Tran. Sig. Proc.*, vol. 50, pp. 2744–2756, Nov. 2002.

[9] M. Shao and C. L. Nikias, "Signal processing with fractional lower order moments: Stable processes and their applications," *Proc. IEEE*, vol. 81, pp. 986–1010, July 1993.

[10] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–692, July 1989.

[11] J. Ilow and D. Hatzinakos, "Applications of the empirical characteristic function to estimation and detection problems," *Sig. Proc.*, vol. 65, pp. 199–219, March 1998.

[12] L. Toffolatti, F. A. Gomez, G. DeZotti, P. Mazzei, A. Franceschini, L. Danese, and C. Burigana, "Extragalactic source counts and contributions to the anisotropies of the cosmic microwave background: predictions for the Planck Surveyor mission," *Mon. Not. R. Astron. Soc.*, vol. 297, pp. 117–127, 1998.