EM-BASED ESTIMATION OF SPATIALLY VARIANT CORRELATED IMAGE NOISE

Bart Goossens, Aleksandra Pižurica, Wilfried Philips*

Ghent University - TELIN - IPI - IBBT
Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium
bart.goossens@telin.ugent.be

ABSTRACT

In image denoising applications, noise is often correlated and the noise energy and correlation structure may even vary with the position in the image. Existing noise reduction and estimation methods are usually designed for stationary white Gaussian noise and generally work less efficient in this case because of the noise model mismatch. In this paper, we propose an EM algorithm for the estimation of spatially variant (nonstationary) correlated image noise in the wavelet domain. In particular, we study additive white Gaussian noise filtered by a space-variant linear filter. This general noise model is applicable to a wide variety of practical situations, including noise in Computed Tomography (CT). Results demonstrate the effectiveness of the proposed solution and its robustness to signal structures.

***Index Terms*—** Noise estimation, Image restoration, Correlated noise

I. INTRODUCTION

Gaussian noise processes are characterized entirely by their second order statistical moments [1]. On the other hand recent studies (e.g. [2], [3]) have shown that signal features in the bandpass and highpass subbands of a given multiresolution representation are not Gaussian and require the specification of the fourth order moment, the kurtosis. This property can be exploited to distinguish signal information from noise and this has succesfully been applied to the estimation of stationary correlated noise [4]. However, in practice, we encounter many situations where the noise energy and correlation structure depends on the position in the image (nonstationary noise). Even for *local* stationary Gaussian noise processes, that have properties that change slowly in space, the estimation is still difficult because only local information can be used. Therefore it is useful to estimate the noise properties in well structured bases that approximately diagonalize noise covariance matrices, such that fewer observations are needed. An example are the local cosine bases in [1].

In this work, we assume an additive noise process, that is generated by sending white Gaussian noise through a linear spatially variant filter. We employ a wavelet basis that has similar "sparsifying" properties as the local cosine bases, but that are better suited in representing nonstationary *signal* features like edges and textures. Wavelet bases provide a non-uniform partitioning of the time-frequency plane which allows retrieving information both in specific frequency bands and at spatial positions. We propose an Expectation-Maximization (EM) algorithm for the wavelet domain estimation of the noise covariance function. The estimated noise properties can be directly plugged in into recent wavelet domain denoising methods (e.g. [2], [5], [6], [7]). On the other hand, this allows us to study noise properties in regions where we have no signal-free patches, e.g. in medical images. First we treat the case where the noise Power Spectral Density (PSD) is the same throughout the image but where the local noise energy is allowed to vary (we will call this separable *space-varying spectrum*, see further). Next, we study the more general case where the local PSD is position-dependent (denoted as *nonseparable space-varying spectrum*).

The remainder of this paper is as follows: in Section II we introduce some basic concepts that are used throughout this paper. In Section III we explain the EM algorithm that is used in the wavelet domain, for both separable and nonseperable space-varying noise spectra. Implementation aspects are discussed in Section IV. Results and a discussion are given in Section V. Finally, Section VI concludes this paper.

II. BASIC CONCEPTS

II-A. Local stationarity and space-varying spectra

Let $Y(\mathbf{t}), \mathbf{t} \in \mathbb{Z}^2$ be a real-valued zero-mean random process with covariance function ($\mathbf{r} \in \mathbb{Z}^2$):

$$R(\mathbf{t}, \mathbf{r}) = \mathrm{E}\left\{Y(\mathbf{t})Y(\mathbf{t}+\mathbf{r})\right\} \tag{1}$$

If the process is *stationary* then the covariance only depends on the distance between two points and not on their absolute positions: $R(\mathbf{t}, \mathbf{r}) = R(\mathbf{0}, \mathbf{r})$. Furthermore, we say that a process is *locally* stationary, if in the neighbourhood of any $\mathbf{v} \in \mathbb{Z}^2$, there exists a square window $\delta(\mathbf{v})$ of size $l(\mathbf{v})$, centered at position $\mathbf{v}$, where the process can be approximated by a stationary one : for $\mathbf{t} \in \delta(\mathbf{v})$ and for $|\mathbf{r}| \leq l(\mathbf{v})/2$, the covariance is well approximated by [1]

$$\mathrm{E}\left\{Y(\mathbf{t})Y(\mathbf{t}+\mathbf{r})\right\} \approx \mathrm{E}\left\{Y(\mathbf{v})Y(\mathbf{v}+\mathbf{r})\right\} = R(\mathbf{v}, \mathbf{r}) \tag{2}$$

We define the space-varying spectrum (SVS) of $Y(\mathbf{t})$ as the Discrete Time Fourier transform (DTFT) of $R(\mathbf{v}, \mathbf{r})$ with respect to $\mathbf{r}$:

$$S(\mathbf{v}, \boldsymbol{\omega}) = \sum_{\mathbf{r} \in \mathbb{Z}^2} R(\mathbf{v}, \mathbf{r}) \exp(-j\langle \mathbf{r}, \boldsymbol{\omega}\rangle) \tag{3}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. For stationary processes, the SVS reduces to the Power Spectral Density (PSD). We say that the SVS is *separable* if it can be factored as $S(\mathbf{v}, \boldsymbol{\omega}) = S_0(\mathbf{v})S_1(\boldsymbol{\omega})$ with $\frac{1}{2\pi}\int_{-\pi}^{\pi} S_1(\boldsymbol{\omega})d\omega = 1$. The first component $S_0(\mathbf{v})$ represents the variance at position $\mathbf{v}$ while the second component $S_1(\boldsymbol{\omega})$ denotes the normalized Power Spectral Density (PSD).

II-B. Spatially variant filtering of White Noise

A specific class of locally stationary processes is obtained by the spatially variant filtering of white noise. Let $\epsilon(\mathbf{t})$ denote a white Gaussian noise process, then $Y(\mathbf{t})$ is obtained as:

$$Y(\mathbf{t}) = \sum_{\mathbf{v} \in \mathbb{Z}^2} \epsilon(\mathbf{v}) K(\mathbf{t}, \mathbf{t}-\mathbf{v}) \tag{4}$$

with $K(\mathbf{t}, \mathbf{r})$ the impulse response of a linear spatially variant filter with DTFT $A(\mathbf{t}, \boldsymbol{\omega})$. The covariance function of $Y(\mathbf{t})$ is then given

*A. Pižurica is a postdoctoral researcher of the Fund for the Scientific Research in Flanders (FWO) Belgium

by:

$$R(\mathbf{t},\mathbf{r}) = \mathrm{E}\{Y(\mathbf{t})Y(\mathbf{t}+\mathbf{r})\} = \frac{1}{2\pi}\int_{-\pi}^{+\pi} A^{\star}(\mathbf{t},\boldsymbol{\omega})A(\mathbf{t}+\mathbf{r},\boldsymbol{\omega})\exp(j\langle\boldsymbol{\omega},\mathbf{r}\rangle)d\boldsymbol{\omega} \quad (5)$$

The local stationarity assumption (2) imposes that $A(\mathbf{t},\boldsymbol{\omega})$ has to satisfy some smoothness conditions (see [1]). More specifically, if $|\langle\mathbf{r},\frac{\partial A(\mathbf{t},\boldsymbol{\omega})}{\partial\mathbf{t}}\rangle| \ll |A(\mathbf{t},\boldsymbol{\omega})|$, for $|\mathbf{r}| \leq l(\mathbf{t})/2$, we have approximately: $A^{\star}(\mathbf{t},\boldsymbol{\omega})A(\mathbf{t}+\mathbf{r},\boldsymbol{\omega}) \approx |A(\mathbf{t},\boldsymbol{\omega})|^2$. In practice, $R(\mathbf{t},\mathbf{r})$ may have many non-zero elements. Therefore, it is useful to use structural bases that compress the covariance function well. This has the advantage that spatial variant correlations can be efficiently estimated from relatively few observations. In this paper, we use an (overcomplete) undecimated wavelet basis. Let $H^{(s,o)}(\boldsymbol{\omega})$ denote the frequency response of the cascaded wavelet filters at scale s and orientation $o \in \{\mathrm{HL},\mathrm{LH},\mathrm{HH}\}$, i.e. $H^{(s,o)}(\boldsymbol{\omega}) = \prod_{i=0}^{s-1} H(2^i\boldsymbol{\omega})G(2^s\boldsymbol{\omega})$, with $G(\omega)$ and $H(\omega)$ respectively the scaling and wavelet filters of each decomposition stage, then the wavelet domain noise covariance function at scale s and orientation o is approximately given by:

$$R^{(s,o)}(\mathbf{t},\mathbf{r}) \approx \frac{1}{2\pi}\int_{-\pi}^{+\pi} |H^{(s,o)}(\boldsymbol{\omega})|^2|A(\mathbf{t},\boldsymbol{\omega})|^2\exp(j\langle\boldsymbol{\omega},\mathbf{r}\rangle)d\boldsymbol{\omega}$$

This means that in the wavelet domain, based on the local stationarity assumption, we also have white noise, but now filtered by the spatially variant filter $A^{(s,o)}(\mathbf{t},\boldsymbol{\omega}) = H^{(s,o)}(\boldsymbol{\omega})A(\mathbf{t},\boldsymbol{\omega})$.

III. WAVELET DOMAIN NOISE ESTIMATION

Our goal is to estimate the noise covariance function $R^{(s,o)}(\mathbf{t},\mathbf{v})$ in the wavelet domain, in the presence of signal structures. Consider one wavelet subband (s,o). For additive noise, we have the following relationship between the noisy wavelet coefficients $\mathbf{y}(\mathbf{t})$, the noise-free coefficients $\mathbf{x}(\mathbf{t})$ and the white noise $\boldsymbol{\epsilon}(\mathbf{t})$ at position $\mathbf{t} \in \mathcal{B}$:

$$\mathbf{y}(\mathbf{t}) = \mathbf{x}(\mathbf{t}) + \mathbf{K}(\mathbf{t})\boldsymbol{\epsilon}(\mathbf{t}) \quad (6)$$

The vectors $\mathbf{x}(\mathbf{t})$, $\epsilon(\mathbf{t})$ and $\mathbf{y}(\mathbf{t})$ are formed by column-stacking the wavelet coefficients in local $\sqrt{d}\times\sqrt{d}$ overlapping windows centered at position $\mathbf{t}$. $\mathbf{K}(\mathbf{t})$ is a spatially variant $d\times d$ matrix that correlates the noise $\boldsymbol{\epsilon}(\mathbf{t}) \sim N(\mathbf{0},\mathbf{I}_d)$. To distinguish noise from signal structures, we take prior knowledge about the noise-free signal $\mathbf{x}(\mathbf{t})$ into account. More specifically, we model $\mathbf{x}(\mathbf{t})$ as a Gaussian Scale Mixture (GSM) [2], [6] with discrete hidden multiplier $z \in \{z_1, z_2, ..., z_K\}$: $\mathbf{x} \overset{d}{=} z^{1/2}\mathbf{u}$. Here, $\overset{d}{=}$ denotes equality in distribution and $\mathbf{u}$ is Gaussian $N(\mathbf{0},\mathbf{C}_u)$. As a result, the covariance matrix of $\mathbf{x}(\mathbf{t})$ is given by $\mathbf{C}_x = \mathrm{E}\{z\}\mathbf{C}_u$. With this model, estimating $R^{(s,o)}(\mathbf{t},\mathbf{r})$ comes down to estimating $\mathbf{K}(\mathbf{t})\mathbf{K}^T(\mathbf{t})$, for which we can use statistical estimation techniques. In the following, we denote $\alpha_k = \mathrm{P}(z = z_k), k = 1,...,K$.

III-A. Noise with separable space-varying spectrum

In many denoising applications, the noise covariance matrix is constant for the whole image, up to a spatially varying scale factor $\sigma^2(\mathbf{t})$, representing the local noise variance. We have:

$$\mathbf{K}(\mathbf{t})\mathbf{K}^T(\mathbf{t}) = \sigma^2(\mathbf{t})\mathbf{C}_\epsilon \quad (7)$$

It is clear that $\mathbf{C}_\epsilon$ can be estimated for the *whole* subband, taking advantage of all the information in the whole image, while $\sigma^2(\mathbf{t})$ can only be obtained *locally*. Let $\boldsymbol{\theta}(\mathbf{t}) = \{\mathbf{C}_u,\mathbf{C}_\epsilon,\sigma^2(\mathbf{t})\} \cup \{\alpha_k, k = 1,...,K\}$ denote the model parameters related to position $\mathbf{t}$. To estimate the total set of model parameters $\boldsymbol{\Theta} = \bigcup_{\mathbf{t}\in\mathcal{B}}\boldsymbol{\theta}(\mathbf{t})$ with hidden variable k, the EM algorithm [8] can be used. Unfortunately, finding an exact solution for the noise covariance matrix, using the classical EM algorithm, has proven to be difficult in general, even for globally stationary noise [4]. Instead, we note that $\mathbf{y}(\mathbf{t})$ is locally (for $\mathbf{t} \in \delta(\mathbf{v})$) distributed according to a *finite* Gaussian Mixture with zero mean and covariances $\mathbf{C}_k(\mathbf{v}) = z_k\mathbf{C}_u + \sigma^2(\mathbf{v})\mathbf{C}_\epsilon$. We will call this the "Scale Mixture" constraint. Our approach then consists in updating the component covariances $\mathbf{C}_k(\mathbf{v})$ such that the "Scale Mixture" constraint is still satisfied. Given a set of model parameters $\boldsymbol{\Theta}^{(i)}$ at iteration i, we optimize the new parameters $\boldsymbol{\Theta}$ in order to increase the objective function:

$$Q(\boldsymbol{\Theta}^{(i)},\boldsymbol{\Theta}) = \mathrm{E}\left\{\log\prod_{\mathbf{v}\in\mathcal{B}}\prod_{\mathbf{t}\in\delta(\mathbf{v})} f(\mathbf{y}(\mathbf{t}),k|\boldsymbol{\theta}(\mathbf{v}))\big|\mathbf{y},\boldsymbol{\Theta}^{(i)}\right\} - \sum_{k=1}^{K}\sum_{\mathbf{v}\in\mathcal{B}}\lambda_k||\mathbf{C}_k(\mathbf{v}) - z_k\mathbf{C}_u - \sigma^2(\mathbf{v})\mathbf{C}_\epsilon||_F^2$$

with the first term the expected complete-data log-likelihood function. The second term denotes the "Scale Mixture" constraint added to the problem using Lagrangian multipliers $\lambda_k, k = 1,...,K$. $||\mathbf{A}||_F^2 = \mathrm{trace}(\mathbf{A}\mathbf{A}^T)$ is the matrix Frobenius norm. It can be shown that the EM update equations are given by:

$$\hat{\alpha}_k = \frac{1}{N}\sum_{\mathbf{v}\in\mathcal{B}}\frac{1}{l^2(\mathbf{v})}\sum_{\mathbf{t}\in\delta(\mathbf{v})}\mathrm{P}(k|\mathbf{y}(\mathbf{t}),\boldsymbol{\theta}(\mathbf{v})), k = 1,...,K$$

$$\hat{C}_k(\mathbf{v}) = \frac{\sum_{\mathbf{t}\in\delta(\mathbf{v})}\mathrm{P}(k|\mathbf{y}(\mathbf{t}),\boldsymbol{\theta}(\mathbf{v}))\mathbf{y}(\mathbf{t})\mathbf{y}^T(\mathbf{t})}{\sum_{\mathbf{t}\in\delta(\mathbf{v})}\mathrm{P}(k|\mathbf{y}(\mathbf{t}),\boldsymbol{\theta}(\mathbf{v}))}, k = 1,...,K$$

$$\begin{pmatrix}\hat{\mathbf{C}}_u\\ \hat{\mathbf{C}}_\epsilon\end{pmatrix} = \begin{pmatrix}N\mu_2 & \mu_1\nu_1\\ \mu_1\nu_1 & \nu_2\end{pmatrix}^{-1}\begin{pmatrix}\sum_{k=1}^K\hat{\alpha}_k z_k\sum_{\mathbf{v}\in\mathcal{B}}\hat{\mathbf{C}}_k(\mathbf{v})\\ \sum_{k=1}^K\hat{\alpha}_k\sum_{\mathbf{v}\in\mathcal{B}}\sigma^2(\mathbf{v})\hat{\mathbf{C}}_k(\mathbf{v})\end{pmatrix} \quad (8)$$

$$\widehat{\sigma^2}(\mathbf{v}) = \frac{\mathrm{trace}\left(\sum_{k=1}^K\hat{\alpha}_k\mathbf{C}_\epsilon(\hat{\mathbf{C}}_k(\mathbf{v}) - \mathbf{C}_u)^T\right)}{\mathrm{trace}\left(\mathbf{C}_\epsilon\mathbf{C}_\epsilon^T\right)} \quad (9)$$

with $\mu_b = \sum_{k=1}^K\hat{\alpha}_k z_k^b, b = 1,2$ and $\nu_b = \sum_{\mathbf{v}\in\mathcal{B}}\sigma^{2b}(\mathbf{v}), b = 1,2$ and with N the number of wavelet coefficients for the considered wavelet subband. We note that update equations (8) and (9) depend on each other and must be used alternatingly in subsequent EM iterations in order to maximize the likelihood. The formulas above must be iterated until convergence of the likelihood. In this iterative process, the *global* noise and signal covariance matrices $\mathbf{C}_u$, $\mathbf{C}_\epsilon$ as well as the *local* variance $\sigma^2(\mathbf{v})$ are estimated jointly. It can be shown that the above formulas are the *exact* classical EM formulas for two mixture components (i.e. $K = 2$). For $K > 2$, we obtain a practical approximation to the classical EM algorithm.

III-B. Noise with nonseparable space-varying spectrum

In a more general scenario, the noise covariance matrix varies spatially and has to be estimated *locally*: $\mathbf{K}(\mathbf{t})\mathbf{K}^T(\mathbf{t}) = \mathbf{C}_\epsilon(\mathbf{t})$. To facilitate this, we will still estimate the signal covariance matrix $\mathbf{C}_u$ *globally*. The objective function now becomes:

$$Q(\boldsymbol{\Theta}^{(i)},\boldsymbol{\Theta}) = \mathrm{E}\left\{\log\prod_{\mathbf{v}\in\mathcal{B}}\prod_{\mathbf{t}\in\delta(\mathbf{v})} f(\mathbf{y}(\mathbf{t}),k|\boldsymbol{\theta}(\mathbf{v}))\big|\mathbf{y},\boldsymbol{\Theta}^{(i)}\right\} - \sum_{k=1}^{K}\sum_{\mathbf{v}\in\mathcal{B}}\lambda_k||\mathbf{C}_k(\mathbf{v}) - z_k\mathbf{C}_u - \mathbf{C}_\epsilon(\mathbf{v})||_F^2$$

Maximizing this function yields the same update equations as in Section III-A, except that (8) and (9) have to be replaced by:

$$\begin{pmatrix}\hat{\mathbf{C}}_u\\ \hat{\mathbf{C}}_\epsilon(\mathbf{t}_1)\\ \hat{\mathbf{C}}_\epsilon(\mathbf{t}_2)\\ \vdots\\ \hat{\mathbf{C}}_\epsilon(\mathbf{t}_N)\end{pmatrix} = \begin{pmatrix}N\mu_2 & \mu_1 & \mu_1 & \dots & \mu_1\\ \mu_1 & 1 & & & \\ \mu_1 & & 1 & & \\ \vdots & & & \ddots & \\ \mu_1 & & & & 1\end{pmatrix}^{-1}\begin{pmatrix}\sum_{\mathbf{v}\in\mathcal{B}}\sum_{k=1}^K\hat{\alpha}_k z_k\hat{\mathbf{C}}_k(\mathbf{v})\\ \sum_{k=1}^K\hat{\alpha}_k\hat{\mathbf{C}}_k(\mathbf{t}_1)\\ \sum_{k=1}^K\hat{\alpha}_k\hat{\mathbf{C}}_k(\mathbf{t}_2)\\ \vdots\\ \sum_{k=1}^K\hat{\alpha}_k\hat{\mathbf{C}}_k(\mathbf{t}_N)\end{pmatrix} \quad (10)$$

where $\mathbf{t}_1, ..., \mathbf{t}_N$ enumerate all spatial positions. Due to the sparsity of the matrix in (10), the solution can be computed in $\mathcal{O}(N)$ operations:

$$\hat{\mathbf{C}}_u = \frac{1}{N} \sum_{k=1}^{K} \sum_{\mathbf{t}\in\mathcal{B}} \hat{\alpha}_k \left(\frac{z_k - \mu_1}{\mu_2 - \mu_1^2} \right) C_k(\mathbf{t}) \tag{11}$$

$$\hat{\mathbf{C}}_\epsilon(\mathbf{t}) = \sum_{k=1}^{K} \hat{\alpha}_k \hat{\mathbf{C}}_k(\mathbf{t}) - \mu_1 \hat{\mathbf{C}}_u, \quad \mathbf{t} \in \mathcal{B} \tag{12}$$

III-C. Parameter initialization

The EM algorithm requires estimates of the initial set of parameters $\mathbf{\Theta}^{(0)}$. Due to the nonstationary character of the noise, the initial signal and noise covariances $\mathbf{C}_u$ and $\mathbf{C}_\epsilon$ are nontrivial to estimate. Therefore, we simply use $\mathbf{C}_\epsilon^{(0)} = \mathbf{I}_d$ (or $\mathbf{C}_\epsilon^{(0)}(\mathbf{t}) = \mathbf{I}_d$ in Section III-B) and $\mathbf{C}_u^{(0)} = \hat{\mathbf{C}}_y - \mathbf{C}_\epsilon^{(0)}$, where $\hat{\mathbf{C}}_y$ is obtained using the Maximum Likelihood (ML) estimator ($\hat{\mathbf{C}}_y = \frac{1}{N}\sum_{\mathbf{v}\in\mathcal{B}} \mathbf{y}(\mathbf{v})\mathbf{y}^T(\mathbf{v})$). The local noise variance in Section III-A is initialized to one: $\sigma^{(0)}(\mathbf{t}) = 1, \mathbf{t} \in \mathcal{B}$.

For the discrete values of the hidden multiplier, we select equidistant samples on a logarithmic scale, similar to [2]: $z_k = \exp(-3 + 7\frac{k-1}{K-1}), k = 1, ..., K$, which probabilities $\alpha_k = z_k^{\tau-1}, k = 1, ..., K$. With this choice, the finite GSM mixture approximates the multivariate Bessel K Form prior with parameter τ (see [3], [6]).

Another important choice is the size of the window $\delta(\mathbf{t})$ where the local stationarity is assumed to be valid. When $\sigma^2(\mathbf{t})$ is sufficiently smooth, according to [1], $l(\mathbf{t})$ should be choosen proportional to $1/\max(\Delta_\mathbf{t}\sigma^2(\mathbf{t}))$, with $\Delta_\mathbf{t}$ the discrete derivative operator with respect to $\mathbf{t}$. This means that the window becomes smaller when the variations in $\sigma^2(\mathbf{t})$ are higher. On the other hand, the estimates (8) and (10) may become unreliable due to insufficient number of samples. For the results in this paper, we use a constant $l(\mathbf{t})$ for simplicity.

IV. IMPLEMENTATION ASPECTS

To speed up the likelihood computations in the EM update formulas, it is useful to apply an extra diagonalisation as in [2]. For noise with separable space-varying spectrum from Section III-A, we have:

$$\mathbf{C}_k(\mathbf{t}) = \mathbf{U}^{-1}(\sigma^2(\mathbf{t})\mathbf{\Lambda} + z_k\mathbf{I}_d)\mathbf{U}^{-T} \tag{13}$$

where $\mathbf{U} = (\mathbf{SQ})^{-1}$, $\mathbf{S}^{-1}\mathbf{C}_\epsilon\mathbf{S}^{-T} = \mathbf{Q\Lambda Q}^T$ and $\mathbf{C}_u = \mathbf{SS}^T$. As a consequence, the determinant of $\mathbf{C}_k(\mathbf{t})$ is given by $|\mathbf{C}_k(\mathbf{t})| = |\mathbf{C}_u||\sigma^2(\mathbf{t})\mathbf{\Lambda} + z_k\mathbf{I}_d| = |\mathbf{C}_u|\prod_{i=1}^{d}(\sigma^2(\mathbf{t})\mathbf{\Lambda}_{ii} + z_k)$. Note that for $\mathbf{t}\in\delta(\mathbf{v})$, we have $\mathbf{y}(\mathbf{t})|k \sim N(\mathbf{0}, \mathbf{C}_k(\mathbf{v}))$, such that:

$$\log f(\mathbf{y}(\mathbf{t})|k, \boldsymbol{\theta}(\mathbf{v})) = -\frac{d}{2}\log(2\pi) - \frac{1}{2}\log|\mathbf{C}_x| - \frac{1}{2}\sum_{i=1}^{d}\left[\log(\sigma^2(\mathbf{v})\mathbf{\Lambda}_{ii} + z_k) + \frac{[\mathbf{Uy}(\mathbf{t})]_i^2}{\sigma^2(\mathbf{v})\mathbf{\Lambda}_{ii} + z_k}\right]$$

Because $\mathbf{Uy}(\mathbf{t})$ only needs to be evaluated once for different k, the diagonalisation (13) yields an approximate speed up of a factor d. However, $\log f(\mathbf{y}(\mathbf{t})|k, \boldsymbol{\theta}(\mathbf{v}))$ still has to be computed for $\mathbf{v} \in \mathcal{B}$ and for $\mathbf{t} \in \delta(\mathbf{v})$. Based on the assumption of local stationarity, we linearize $\sigma^2(\mathbf{t})$ in the neighbourhood of $\mathbf{v}$ ($\mathbf{t} \in \delta(\mathbf{v})$), which means that we can write $\log f(\mathbf{y}(\mathbf{t})|k, \boldsymbol{\theta}(\mathbf{v}))$ in terms of $\log f(\mathbf{y}(\mathbf{v})|k, \boldsymbol{\theta}(\mathbf{v}))$ using the Taylor-series approximation:

$$\log f(\mathbf{y}(\mathbf{t})|k, \boldsymbol{\theta}(\mathbf{v})) \approx \log f(\mathbf{y}(\mathbf{v})|k, \boldsymbol{\theta}(\mathbf{v})) - \frac{(\sigma^2(\mathbf{t}) - \sigma^2(\mathbf{v}))}{2}\sum_{i=1}^{d}\frac{(\sigma^2(\mathbf{t})\mathbf{\Lambda}_{ii} + z_k - [\mathbf{Uy}(\mathbf{t})]_i^2)\mathbf{\Lambda}_{ii}}{(\sigma^2(\mathbf{v})\mathbf{\Lambda}_{ii} + z_k)^2}$$

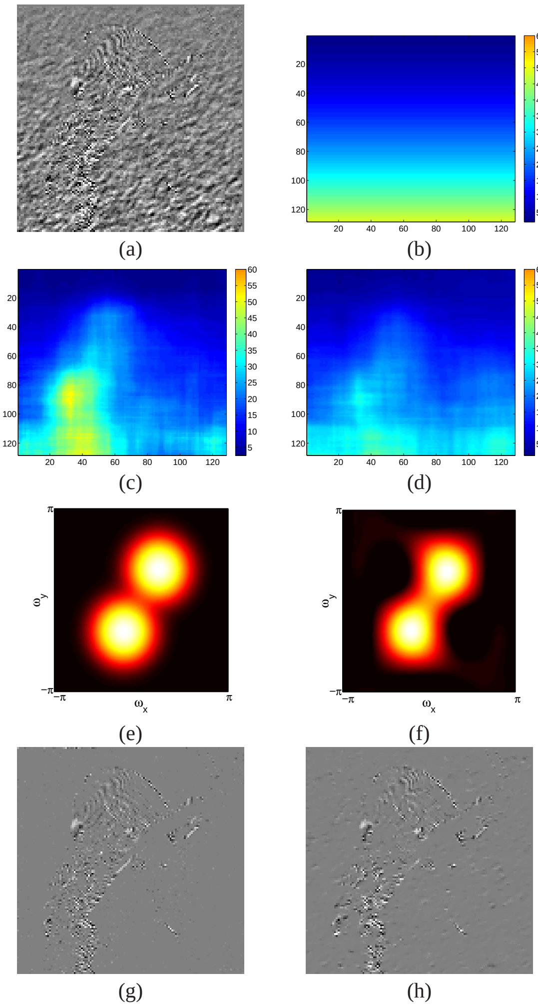

Fig. 1. (a) Wavelet subband of Lena with added artificial noise with separable SVS (b) True local noise variance $\sigma^2(\mathbf{t})$ (c) Estimated local noise variance $\widehat{\sigma^2}(\mathbf{t})$, using the MAD-estimator (MSE=0.682) (d) Estimated local noise variance $\widehat{\sigma^2}(\mathbf{t})$, using the proposed method (MSE=0.451) (e) True noise PSD (f) Estimated noise PSD, using the proposed method (g) Original noise-free wavelet subband of Lena (h) Denoised wavelet subband of (a) using the estimated noise PSD (f) and local variance (d)

The logarithm and the sum in this equation have to be computed once for every $\mathbf{v} \in \mathcal{B}$.

V. RESULTS AND DISCUSSION

In Fig. 1, visual results are given for the noise estimation of Section III-A. First, the noise-free wavelet subband of Fig. 1g is corrupted with additive noise, resulting from filtering white Gaussian noise by the space variant filter with spectrum $|A(\mathbf{t}, \boldsymbol{\omega})|^2 \sim t_y^2 \exp(-60((\omega_x - 0.34\pi)^2 + (\omega_y - 0.20\pi)^2))$, see Fig. 1a. Here ω_x and ω_y denote respectively the x- and y-components of $\boldsymbol{\omega}$ and t_y is the y-component of $\mathbf{t}$. We use $l(\mathbf{t}) = 32$ and $d = 9$, corresponding to a 3×3 window for local correlations. The local noise variance $\sigma^2(\mathbf{t}) \sim t_y^2$ is depicted in Fig. 1b. In Fig. 1c the local noise variance is estimated locally using the robust Median of Absolute Deviations (MAD) estimator in a 32×32-window. Fig. 1d shows the estimated $\widehat{\sigma^2}(\mathbf{t})$ using the proposed method with the same window size. The EM estimate is clearly much more robust to the presence of signal structures than the MAD estimate. This is mainly due to the fact that our method takes *signal correlations* into account whereas the MAD estimate does not. The estimated

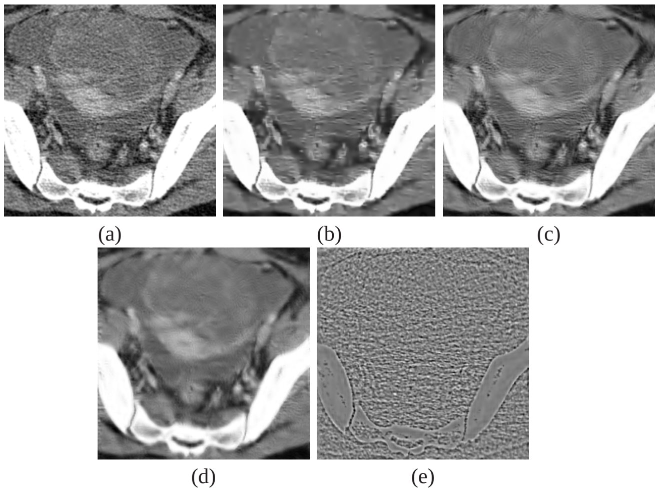

Fig. 2. (a) Pathological Thorax Computed Tomography (CT) image of a 15-year old female, source: Sophia Children's Hospital, EMC, Rotterdam, the Netherlands (b) Denoised version of (a), using [5] (c) Denoised version of (a), using [4] (d) Denoised version of (a), using the proposed method. (e) Difference image between (d) and (a) (contrast enhanced, intensity 128 corresponds to difference zero)

noise PSD in Fig. 1f. is obtained first by converting the estimated noise covariance matrix $\hat{\mathbf{C}}_\epsilon$ into an autocorrelation function of size $128{\times}128$ by averaging over correlations that correspond to the same difference in position, putting correlations that cannot be captured using a $\sqrt{d}{\times}\sqrt{d}$ window to zero and next by computing the Discrete Fourier Transform. Despite the small window size 3×3 used for estimating local correlations, there is a very good resemblance with the original noise PSD in Fig. 1e. Next, the estimated noise parameters from Fig. 1d and Fig. 1f are used to denoise the wavelet subband, with an extension of the algorithm presented in [6] (such that it can deal with nonstationary noise, similar to the extension presented in [9]), which results in Fig. 1h. Due to the reliable noise estimation, the denoising algorithm reconstructs most of the signal structures present in Fig. 1a.

In Fig. 2, we give visual results for the noise estimation of Section III-B. Fig. 2a shows a low-dose Computed Tomography (CT) image, that suffers from noisy streaking artifacts. It can be shown that by the filtered backprojection in CT reconstruction, CT noise can be modeled as being filtered by a space-variant filter. We compare the proposed noise estimation combined with the extension of the algorithm of [6] (Fig. 2d) (see above), to the blind denoising methods of [5] (Fig. 2b) and [4] (Fig. 2c). The method of [5] assumes white stationary noise and estimates the noise variance from the highpass subband of the nondecimated spline wavelet transform. Due to the noise model mismatch, noise artifacts are left in the denoised image (Fig. 2b) in areas where the local noise variance exceeds the estimated noise variance, whereas the proposed method does not. The method of [4] [1] assumes stationary correlated noise and also because of the nonstationarity, not all parts of the noise are removed. Our method uses the Dual-Tree Complex wavelet transform from [10], with 3 scales, $d = 9$ and $l(\mathbf{t}) = 16$. Fig. 2d shows the difference image of Fig. 2c and Fig. 2a. It can be noticed that some signal structures are present in the difference image, for example at the edges of the bright areas in Fig. 2a. Here, due to the saturation in the scanner at intensity 255, there is a fast transition in the local noise variance. As a consequence, the local-stationarity assumption is violated and the local noise variance is slightly overestimated, causing oversmoothing of the edges. To deal with saturation effects from the scanner, we are currently investigating the use of a pre-segmentation, such that the local-stationarity assumption is still valid within each detected segment. On Pentium IV 2 GHz processor, denoising a 256×256 image in an unoptimized implementation takes 143 s, from which 110 s are spent to noise estimation.

[1]Because an implementation of the method from [4] is not yet available from the author, we developed and used our own implementation of [4].

VI. CONCLUSION

In many practical situations, stationary white Gaussian noise models are too restrictive and yield poor denoising performance due to the noise model mismatch. White Gaussian noise filtered by a spatially variant filter, as a specific class of locally stationary processes, offers a much broader applicability and its parameters can be estimated in a wavelet basis, which compresses the spatially variant noise autocorrelation function well. We presented a new algorithm for the noise parameter estimation of spatially variant correlated image noise, which is not possible yet using current existing techniques.

VII. REFERENCES

[1] S. Mallat, G. Papanicolaou, and Z. Zhang, "Adaptive covariance estimation of locally stationary processes," *Ann. Statist.*, vol. 26, no. 1, pp. 1–47, 1998.

[2] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Image denoising using Gaussian Scale Mixtures in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, pp. 1338–1351, 2003.

[3] A. Srivastava, X. Liu, and U. Grenander, "Universal Analytical Forms for Modeling Image Probabilities," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1200–1214, Sept. 2002.

[4] J. Portilla, "Full blind denoising through noise covariance estimation using Gaussian Scale Mixtures in the wavelet domain," *IEEE Int. Conf. on Image Process. (ICIP)*, vol. 2, pp. 1217–1220, Oct. 2004.

[5] A. Pižurica, W. Philips, I. Lemahieu, and M. Acheroy, "A versatile wavelet domain noise filtration technique for medical imaging," *IEEE Trans. Medical Imaging*, vol. 22, no. 3, pp. 323–331, Mar. 2003.

[6] B. Goossens, A. Pižurica, and W. Philips, "Removal of Correlated Noise by Modeling Spatial Correlations and Interscale Dependencies in the Complex Wavelet Domain," in *IEEE Int. Conf. on Image Processing (ICIP)*, San Antonio, Texas, USA, Sept. 2007, pp. 317–320.

[7] ——, "Wavelet domain image denoising for non-stationary and signal-dependent noise," in *IEEE International Conference on Image Processing (ICIP)*, Atlanta, GA, USA, Oct. 2006, pp. 1425–1428.

[8] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, vol. 39, no. 1, pp. 1–38, 1977.

[9] J. Portilla, "Image restoration using Gaussian Scale Mixtures in Overcomplete Oriented Pyramids (a review)," in *Proc. of the SPIE's 50th Annual Meeting: Wavelets XI*, vol. 5914, San Diego, CA, Aug. 2005, pp. 468–482.

[10] N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Journal of Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 234–253, May 2001.