Research Article

A Novel Adaptive Probabilistic Nonlinear Denoising Approach for Enhancing PET Data Sinogram

Musa Alrefaya[1,2] and Hichem Sahli[1,3]

[1] Department of Electronics and Informatics (ETRO-IRIS), Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
[2] Faculty of Information Technology and Computer Engineering, Palestine Polytechnic University, Hebron, Palestine
[3] Interuniversity Microelectronics Centre (IMEC), Leuven, Belgium

Correspondence should be addressed to Musa Alrefaya; musa.alrefaya@vub.ac.be

We propose filtering the PET sinograms with a constraint curvature motion diffusion. The edge-stopping function is computed in terms of edge probability under the assumption of contamination by Poisson noise. We show that the Chi-square is the appropriate prior for finding the edge probability in the sinogram noise-free gradient. Since the sinogram noise is uncorrelated and follows a Poisson distribution, we then propose an adaptive probabilistic diffusivity function where the edge probability is computed at each pixel. The filter is applied on the 2D sinogram prereconstruction. The PET images are reconstructed using the Ordered Subset Expectation Maximization (OSEM). We demonstrate through simulations with images contaminated by Poisson noise that the performance of the proposed method substantially surpasses that of recently published methods, both visually and in terms of statistical measures.

1. Introduction

Positron Emission Tomography (PET) is an in vivo nuclear medicine imaging method that provides functional information of the body tissues. The PET image results from reconstructing very noisy, low resolution raw data, that is, the sinogram, in which important features are shaped as curved structures. Enhancing the PET image spurred a wide range of denoising models and algorithms. Some methodologies focus on enhancing the reconstructed PET image directly, where others prefer enhancing the sinogram prior to reconstruction.

Although developing an appropriate denoising method for filtering the PET images has been widely studied in the last two decades, this problem is still receiving high attention from researchers trying to improve the diagnosis accuracy of this image. Existing methods may suffer drawbacks such as the careful selection of a high number of parameters, smoothing of the important features boundaries, or prohibitive computation. Recently, nonlinear diffusion techniques have been investigated for PET images. Many researchers did explore the application of the well-known Perona and Malik anisotropic diffusion [1] in combination with diverse diffusivity functions on PET images [2–6], as well as on sinograms [7–9].

In [5], authors introduced a postreconstruction adaptive nonlinear diffusion (Perona and Malik) filter based on varying the diffusion level according to a local estimation of the image noise. Applying the nonlinear anisotropic filtering method on the whole body, PET image and the sinogram were presented in [4, 8, 10], respectively. Results showed that the anisotropic diffusion filtering algorithm helps improve the quantitative aspect of PET images.

In the study of [9], combining the anisotropic diffusion method with the coherence enhancing diffusion method for filtering the sinogram as a preprocessing step was proposed. However, the considered cascading approach is time consuming, and the results are highly dependent on the parameters selection criteria. Zhu et al. [11] built the diffusivity function using fuzzy rules that were expressed in a linguistic form.

The mean curvature and the Gauss curvature diffusion algorithms for filtering the PET images during reconstruction were investigated in [12]. An anatomically driven anisotropic diffusion filter (ADADF) as a penalized maximum likelihood expectation maximization optimization framework was proposed in [2]. This filter enhances a single-photon emission computed tomography images using anatomical information from magnetic resonance imaging as a priori knowledge about the activity distribution. Authors in [3] proposed a reconstruction algorithm for PET with thin plate prior combined with a forward-and-backward diffusion algorithm. The main drawback of the above filter, with respect to sinogram images is that the diffusion produces important oscillations in the gradient. This leads to a poorly smoothed image [11, 12]. Moreover, the adopted diffusivity functions do not consider the special properties of the sinogram, which are high level of Poisson noise and curved-shape features.

Happonen and Koskinen [13] proposed filtering the sinogram in a new domain, that is, stackgram domain, where the signal along the sinusoidal trajectories can be filtered separately. They applied this method using the Gaussian and nonlinear filters. Radon transform is applied for inverse mapping of the filtered data from the stackgram domain to the sinogram domain. The above described method is impractical for noise reduction of the medical images such as PET, since the corresponding 3D stackgram requires an enormous space of computer memory. Furthermore, it is a complex method not suitable for clinical applications where timely reconstructien of the PET image is a very important issue.

Most of the above studies considered a global image noise level. Local or varying noise level has been used in [14], where a nonlinear diffusion method for filtering MR images with varying noise levels was presented. The authors assumed that the MR image can be modeled as a piecewise constant (slowly varying) function and corrupted by additive zero-mean Gaussian noise. Pizurica et al. [15] proposed a wavelet domain denoising method for subband-adaptive and spatially adaptive image denoising approach. The latter approach is based on the estimation of the probability that a given co-efficient contains a significant noise-free component called "signal of interest." The authors of [15] found that the spatially adaptive version of their proposed method yielded better results than the existing spatially adaptive ones.

In this work, based on the following PET sinogram characteristics:

  (i) the important features in the sinogram are curved structures with high contrast values and these represent the regions of interest in the reconstructed PET image, for example, tumor,

 (ii) the weak edges in the sinogram are the edges that contain low contrast values,

(iii) the noise in the sinogram is a priori identified as a Poisson noise,

we propose filtering the sinogram by means of a curvature constrained filter. The amount of diffusion is modulated according to a probabilistic diffusivity function that suits images contaminated with Poisson noise, being the known noise distribution of PET sinograms. Further, considering the singoram characteristics, we propose a probabilistic edge-stopping function based on Chi-square prior for the ideal sinogram gradient with a spatially adaptive algorithm for calculating the prior odd at each pixel. We show that this method is improving and denoising sinogram data which leads to enhanced reconstructed PET image.

The remainder of the paper is organized as follows, Section 2 briefly reviews the notions of curvature motion, edge affected diffusion filtering, and self-snakes. The proposed adaptive filtering scheme is introduced in Section 3. Finally, Section 4 discusses the experimental results, while the conclusions and future work are given in Section 5.

2. Geometry Driven Scale-Space Filtering

This section reviews the formulations for Mean Curvature Motion (MCM), Edge Affected Variable Conductance Diffusion (EA-VCD) and self-snakes. Also we recall the Gauge Coordinates notions. Extensive discussion can be find in [16]. Let f be a scalar image defined on the spatial image domain Ω, then the family of diffused versions of f is given by

$$U(f) : f(\cdot) \longrightarrow u(\cdot, t) \quad \text{with } u(\cdot, 0) = f(\cdot), \quad (1)$$

where U is referred to as the scale-space filter, u is denoted by the scale-space image, and the scale $t \in \mathbb{R}^+$ [6]. The denoised or enhanced version of f is a given $u(\cdot, t)$ that is the closest to the unknown noise-free version of f. In the following we denote $u(:; t)$ by u_t.

2.1. Curvature Motion. Mean Curvature Motion (MCM) is considered as the standard curvature evolution. MCM allows diffusion solely along the level lines. In Gauge coordinates (see Section 2.4) the corresponding PDE formulation is (k is Euclidian curvature) as follows:

$$u_t(\cdot, t) = u_{vv} = k |\nabla u| = \text{div} \left[\frac{\nabla u}{|\nabla u|} \right] |\nabla u|. \quad (2)$$

Hence diffusion solely occurs along the v-axis.

2.2. Edge Affected Variable Conductance Diffusion. Variable Conductance Filtering (VCD) is based on the diffusion with a variable conduction coefficient that controls the rate of diffusion [16]. In the case of Edge Affected-VCD (EA-VCD), the conductance coefficient is inversely proportional to the edgeness. Consequently it is commonly referred to as the edge-stopping function (g), in which the edgeness is typically measured by the gradient magnitude. The EA-VCD is governed by

$$u_t = \text{div} \left[g(|\nabla u|) \nabla u \right]. \quad (3)$$

The above PDE system together with the initial condition given in (1) is completed with homogenous von Neumann boundary condition on the boundary of the image domain. Note that the Perona and Malik's antitropic diffusion [1] is an EA-VCD.

2.3. Self-Snakes. Self-snakes are a variant of the MCM where an edge-stopping function is introduced [17]. The main goal is preventing further shrinking of the level-lines once they have reached the important image edges. For scalar images, self-snakes are governed by

$$u_t = |\nabla u| \operatorname{div} \left[g\left(|\nabla u|\right) \frac{\nabla u}{|\nabla u|} \right]. \qquad (4)$$

This equation adopts the same boundary condition as (3). Furthermore, it can be decomposed into two parts [16, 17]:

$$\begin{aligned} u_t &= g\,|\nabla u| \operatorname{div} \left[\frac{\nabla u}{|\nabla u|} \right] + (\nabla g) \cdot \nabla u \\ &= gk\,|\nabla u| + (\nabla g) \cdot \nabla u. \end{aligned} \qquad (5)$$

The first part describes a degenerate forward diffusion along the level lines that is orthogonal to the local gradient; it allows preserving the edges. Additionally, the diffusion is limited in areas with high gradient magnitude and encouraged in smooth areas. Actually the first term is the constraint curvature motion. The second term can be viewed as a shock filter since it pushes the level-lines towards valleys of high gradient, acting as Osher's shock filter.

2.4. Gauge Coordinates. An image can be thought of as a collection of curves with equal value, the isophotes. At extrema an isophote reduces to a point, at saddle points the isophote is self-intersecting. At the noncritical points; a Gauge coordinate system [18] is defined by two orthogonal axes v and w, which correspond to the directions of the tangent and normal at the isophote. The second order Gauge derivatives of the image in the vv and ww directions have the following second-order structures:

$$\begin{aligned} u_{vv} &= \frac{u_{xx}u_y^2 - 2u_x u_y u_{xy} + u_{yy}u_x^2}{\left(u_x^2 + u_y^2\right)}, \\ u_{ww} &= \frac{u_{xx}u_y^2 + 2u_x u_y u_{xy} + u_{yy}u_x^2}{\left(u_x^2 + u_y^2\right)}. \end{aligned} \qquad (6)$$

These gauge derivatives can be expressed as a product of gradients and a 2×2 matrix with second-order derivatives [18]:

$$\begin{aligned} u_{ww}u_w^2 &= \left(u_x, u_y\right) \begin{pmatrix} u_{xx} & u_{xy} \\ u_{xy} & u_{yy} \end{pmatrix} \begin{pmatrix} u_x \\ u_y \end{pmatrix}, \\ u_{vv}u_w^2 &= \left(u_x, u_y\right) \begin{pmatrix} u_{yy} & -u_{xy} \\ -u_{xy} & u_{xx} \end{pmatrix} \begin{pmatrix} u_x \\ u_y \end{pmatrix}. \end{aligned} \qquad (7)$$

For u_{ww} the matrix equals the Hessian H. For u_{vv} it is det $H \cdot H^{-1}$. Note that the expressions are invariant with respect to the spatial coordinates. The above expression can also be obtained as follows [16]:

$$\begin{aligned} u_{vv} &= u_{xx}\sin\theta^2 + u_{yy}\cos\theta^2 - 2u_{xy}\sin\theta\cos\theta, \\ u_{ww} &= u_{xx}\cos\theta^2 + u_{yy}\sin\theta^2 + 2u_{xy}\sin\theta\cos\theta, \end{aligned} \qquad (8)$$

where θ is given as $\theta = \arctan(u_y/u_x)$ and $u_v = -u_x\sin\theta + u_y\cos\theta = 0$.

The two expressions of (6) can be combined linearly in a PDE setting. In this way, an image u_0 is evolved according to a weighted combination of these two invariants:

$$u_t = g_1\left(|\nabla u|\right) u_{vv} + g_2\left(|\nabla u|\right) u_{ww} \qquad (9)$$

with $u(t = 0) = u_0$.

Equation (9) comprises a diffusion modulated by g_1 along the image edges vv (a smoothing term) and a diffusion adjustable by g_2 across the image edges ww (a sharpening term). Careful modeling of these terms allows efficiently denoising the PET sinograms, whilst keeping their interesting features.

3. The Probabilistic Curvature Motion Filter

The proposed probabilistic curvature motion filters are based on the idea of the *probabilistic diffusivity function* [19], where the diffusivity function is expressed as the probability that the observed gradient presents no edge of interest under a suitable marginal prior distribution for the noise-free gradient. In [19], the probabilistic diffusivity function has been defined as

$$g_{\mathrm{pr}}(x) = A\left(1 - P\left(H_1 \mid x\right)\right), \qquad (10)$$

where the normalizing constant A is set to $A = 1/(1 - P(H_1 \mid 0))$ to ensure that $g_{\mathrm{pr}}(0) = 1$, the hypothesis H_1 describes the notion whether an edge element of interest is present given the considered noise and H_0 an edge element of interest is absent. For a noisy gradient model $x = y + n$, we set

$$H_0 : y \leq \sigma_n, \qquad H_1 : y > \sigma_n \qquad (11)$$

with y being the ideal, noise-free, gradient magnitude, and σ_n being the standard deviation of the noise n. In [19] it has been demonstrated that

$$g_{\mathrm{pr}}(x) = \left(1 + \mu\eta(0)\right) \frac{1}{1 + \mu\eta(x)} \qquad (12)$$

with the prior odds defined as

$$\mu = \frac{P\left(H_1\right)}{P\left(H_0\right)} \qquad (13)$$

and the likelihood ratio

$$\eta(x) = \frac{P\left(x \mid H_1\right)}{P\left(x \mid H_0\right)}. \qquad (14)$$

Considering a Laplacian prior $p(y) = (\lambda/2)e^{-\lambda|y|}$, we have [19]:

$$\mu = \left(e^{\lambda\sigma_n} - 1\right)^{-1}, \qquad (15)$$

and the parameter λ can be estimated as

$$\lambda = \left[0.5\left(\sigma^2 - \sigma_n^2\right)\right]^{-1/2} \qquad (16)$$

with σ^2 denoting the variance of the noisy image and σ_n as defined above. Due to limited space, the reader is referred to [19] for the detailed expression of $\eta(x)$ in (14).

It has to be noted that the diffusivity function (12) fits in the cluster of backward-forward diffusivities, and it has no free parameters to be set. Moreover, for the considered PET sinograms, the noise standard deviation, σ_n, in (16) is estimated as $\sigma_n^2 = \mathrm{Var}(f_{\mathrm{Ln}})$, where the image noise, f_{Ln}, is reconstructed via wavelet decomposition of f, from the two finest resolution levels coefficients, using Daubechies (4) wavelet.

3.1. Probabilistic Self-Snakes (PSS).

It can be demonstrated that the diffusion of scalar images via EA-VCD can be decomposed into (5), which can be rewritten as follows [16]:

$$u_t = g\left(|\nabla u|\right) u_{vv} + \left[g\left(|\nabla u|\right) + g'\left(|\nabla u|\right)|\nabla u|\right] u_{ww} \quad (17)$$

consolidating the properties of both the self-snakes and the EA-VCD into a single diffusion schema.

Considering (9) and the probabilistic diffusivity function, if we set $g_1(x) \doteq g_{\mathrm{pr}}(x)$, and the sharpening term, $g_2(x) \doteq g_{\mathrm{pr}}(x) + x g'_{\mathrm{pr}}(x)$, we obtain the probabilistic self-snakes (PSS) [7]. By its nature, the PSS enhances the weak edges and/or features in the sinogram. The second term in (32) allows the sharpening. In this way, weak but important edges are enhanced whilst the noise is removed efficiently.

PSS proved to be very effective and flexible for the sinogram image where the high contrast regions, which represent a tumor in the reconstructed PET, should be smoothed wisely without blurring the poor edges [7]. Like EA-VCD, the main advantage of this filter is that the average gray value of the image is not altered during the diffusion process which is a significant issue in the sinogram.

3.2. Adaptive Probabilistic Diffusivity Function Based on a Chi-Square Prior for a Sinogram.

The probabilistic diffusivity function (12) does not take into account the spatially varying noise levels such as the case for the sinograms. It has a global threshold parameter, which is related to the image noise standard deviation $T = \sigma_n$. Hence, in such formulation, if two pixels/voxels have equal gradient magnitude, they will have the same $g_{\mathrm{pr}}(x)$ values, no matter the noise level at these pixels. In this work, the probabilistic diffusivity function is improved by considering the local statistical noise at each element. We adopt the ideas of [15] and adapt the estimator to the local spatial context in the image.

3.2.1. Spatially Adaptive Probabilistic Diffusivity Function.

The most appropriate way to achieve a spatial adaptation is to estimate the prior probability of signal presence $P(H_1)$ adaptively for each pixel instead of fixing it globally. This can be achieved by conditioning the hypothesis H_1 on a local spatial activity indicator such as the locally averaged magnitude or the local variance of the observed gradient.

To estimate the probability that "signal of interest" is present at the position i, we consider a local spatial activity indicator, z_i, at each position. z_i is defined as the locally averaged gradient magnitude within a relatively small window, $w(i)$, with a fixed size, N, around the position i in the gradient image.

Starting from the prior odd, (13), we replace the ratio of "global" probabilities by a locally adaptive prior, $P(H_1 \mid z_i)/P(H_0 \mid z_i)$; that is, $P(H_1)$ and $P(H_0)$ are conditioned on the local spatial indicator

$$\frac{P\left(H_1 \mid z_i\right)}{P\left(H_0 \mid z_i\right)} = \frac{P\left(z_i \mid H_1\right)}{P\left(z_i \mid H_0\right)} \cdot \frac{P\left(H_1\right)}{P\left(H_0\right)} = \xi_i\left(z_i\right) \cdot \mu, \quad (18)$$

where (ξ) is the local likelihood ratio:

$$\xi_i\left(z_i\right) = \frac{P\left(z_i \mid H_1\right)}{P\left(z_i \mid H_0\right)} \qquad \mu = \frac{P\left(H_1\right)}{P\left(H_0\right)}. \quad (19)$$

Considering Bayes' rule, the probability that an "edge of interest" is present at position i, $P(H_1 \mid x_i)$, is given as

$$P\left(H_1 \mid x_i\right) = \frac{\mu \eta_i\left(x_i\right) \xi_i\left(z_i\right)}{1 + \mu \eta_i\left(x_i\right) \xi_i\left(z_i\right)}. \quad (20)$$

The spatially adaptive probabilistic diffusivity function can then be formulated as

$$g_{\mathrm{apr}}\left(x_i, z_i\right) = A\left(1 - P\left(H_1 \mid x_i, z_i\right)\right),$$

$$g_{\mathrm{apr}}\left(x_i, z_i\right) = \left(1 + \mu \eta\left(0\right)\right)\left(\frac{1}{1 + \mu \eta_i\left(x_i\right) \xi_i\left(z_i\right)}\right) \quad (21)$$

with

$$\eta_i\left(x_i\right) = \frac{P\left(x_i \mid H_1\right)}{P\left(x_i \mid H_0\right)}. \quad (22)$$

We ensure that $g(0) = 1$, because the minimum of $P(H_1 \mid x)$ is at $x = 0$ and thus $(1 - P(H_1 \mid x))$ peaks at $x = 0$.

Intuitively, the proposed method considers an "observed gradient" at a given location as how probable this location presents useful information compared to its neighborhood, based on

(1) the likelihood ratio via $\eta_i(x_i)$,

(2) a measurement from the local surrounding via $\xi_i(z_i)$,

(3) the global prior odd via μ.

The local spatial activity indicator z_i is defined as

$$z_i = \frac{1}{N} \sum_{l \in w(i)} x_l, \quad (23)$$

where x_l is the gradient magnitude at location $l \in w(i)$. Assume that all the elements within the small window are equally distributed and conditionally independent. With these simplifications, the conditional probability of z_i given H_1 in a window $w(i)$ of size N, which is denoted as $P_N(z_i \mid H_1)$, is given by N convolutions of $P(x_i \mid H_1)$ with itself as follows:

$$P_N\left(z_i \mid H_1\right) = P\left(x_i \mid H_1\right) \mathrm{Conv}_N P\left(x_i \mid H_1\right), \quad (24)$$

while the conditional probability $P_N(z_i \mid H_0)$ of z_i given H_0 is given by N convolutions of $P(x_i \mid H_0)$ with itself:

$$P_N\left(z_i \mid H_0\right) = P\left(x_i \mid H_0\right) \mathrm{Conv}_N P\left(x_i \mid H_0\right). \quad (25)$$

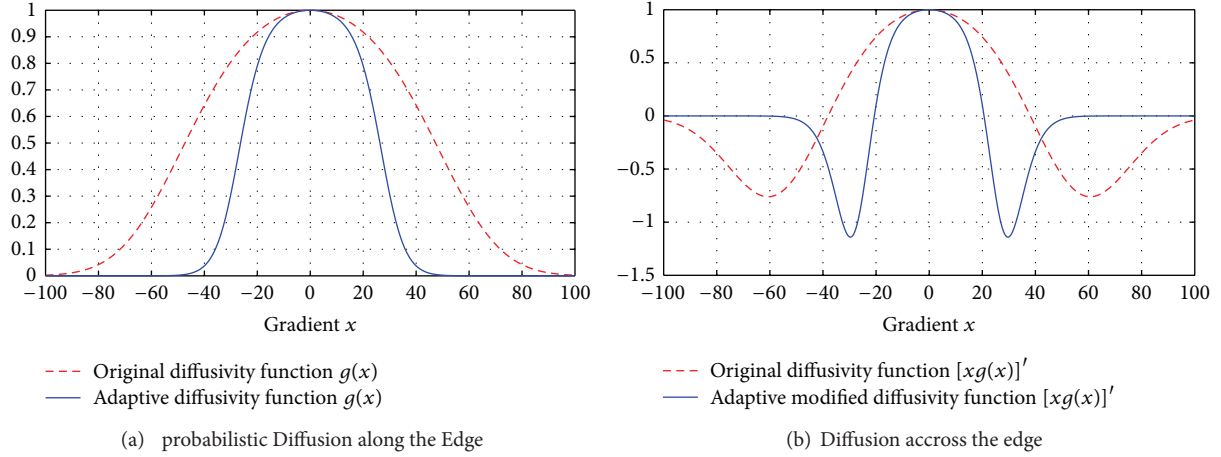(a) probabilistic Diffusion along the Edge

(b) Diffusion accross the edge

FIGURE 1: The Adaptive probabilistic diffusivity function $g_{\mathrm{apr}}(\ )$ versus the original probabilistic diffusivity function $g_{\mathrm{pr}}(\ )$. (a) Diffusivity functions along the edge ($g(x)$). (b) Diffusivity functions across the edge ($[xg(x)]'$).

Due to limited space, the reader is referred to [19] for the detailed expression of $P(x_i \mid H_0)$, $P(x_i \mid H_1)$, and $\eta(x_i)$.

Figure 1 illustrates the original probabilistic diffusivity function and adaptive probabilistic diffusivity function. The adaptive function has lower values, as shown in Figure 1(a), which allows better preservation of important edges than the original function. The negative peaks of the proposed function occur usually at larger gradient magnitude than the original one, which indicates stronger edge enhancement capability. This property indicates a quick smoothing of the nearly uniform areas while maintaining and enhancing the strong edges.

3.2.2. A Chi-Square Prior for Noise-Free Sinogram Gradient. The noise in the sinogram is a priori identified as a Poisson noise [12, 20]. The magnitude of Poisson noise varies across the image, as it depends on the image intensity. This makes removing such noise very difficult. In the original probabilistic diffusivity function (Section 3), the Laplacian prior was imposed for the ideal image gradient that is contaminated with Gaussian noise. In order to take into account the sinogram's noise distribution in the filtering scheme, in this section we aim at redefining the diffusivity function for handling Poisson noise. This can be accomplished by finding an appropriate prior for the ideal noise-free sinogram gradient.

In the following, we argue that we can represent Poisson distribution by a Gaussian distribution as $\mathrm{Gauss}(0, \sqrt{2m})$. Afterwards, we demonstrate that the gradient magnitude of the noise-free sinogram follows a Chi-square distribution, and finally we reformulate the probabilistic diffusivity function based on a Chi-square prior for the noise-free sinogram gradient.

In the literature, several studies demonstrated that the Poisson distribution (of probability mass $P(S) = m^S e^{-m}/S!$, with S being the number of occurrences of an event and m being the expected number of occurrences during a given

interval) approaches a Gaussian density function in the case of high number of counts [21, 22]. Moreover, Miller et al. [21] showed that the Gaussian approximation is surprisingly accurate, even for a fairly small number of counts. To illustrate this, we use the logarithmic function to simplify the proof:

$$\ln P(S) = \ln\left(\frac{m^S e^{-m}}{S!}\right). \tag{26}$$

Using Stirling's formula (for large S as we are assuming here) $S! \approx S^S \cdot e^{-S}\sqrt{2\pi S}$, we have

$$P(S) \approx \frac{e^{-(S-m)^2/2m}}{\sqrt{2\pi m}}. \tag{27}$$

Assuming that the sinogram gradient is approximated by absolute difference of neighboring pixel values on a 2-connected grid, we demonstrate that the gradient of Poisson random variables follows a Skellam distribution. Then, we show that the Skellam distribution can be approximated as a Gaussian distribution.

Let $s1$ and $s2$ be two statistically independent adjacent pixels in the observed sinogram be with a Gaussian distribution $s1 \sim \mathrm{Gauss}(m_1, \sigma_{s1})$ and $s2 \sim \mathrm{Gauss}(m_2, \sigma_{s2})$, respectively. The distribution of the difference, $a = s1 - s2$, of two statistically independent random variables $s1$ and $s2$, each having Poisson distributions with different expected values m_1 and m_2, is denoted as the Skellam distribution [23] and can be given as

$$\mathrm{PD}(a; m_1, m_2) = e^{-(m_1+m_2)}\left(\frac{m_1}{m_2}\right)^{a/2} I_{|a|}\left(2\sqrt{m_1 m_2}\right), \tag{28}$$

where $I_k(Z)$ is the modified Bessel function of the first kind.

The difference between two Poisson variables has the following properties: (i) $\sigma_{s1s2}^2 = \sigma_{s1}^2 + \sigma_{s2}^2 = 2\sigma^2$ and (ii) $m = m_{s1s2} = m_{s1} - m_{s2} = 0$. Considering these properties, the cross-correlation, and the delta function, the

(a) Noise-free histogram

(b) Noisy gradient histogram

■ Noise-free histogram
— Estimated χ^2 prior
- - - Estimated Laplacian prior

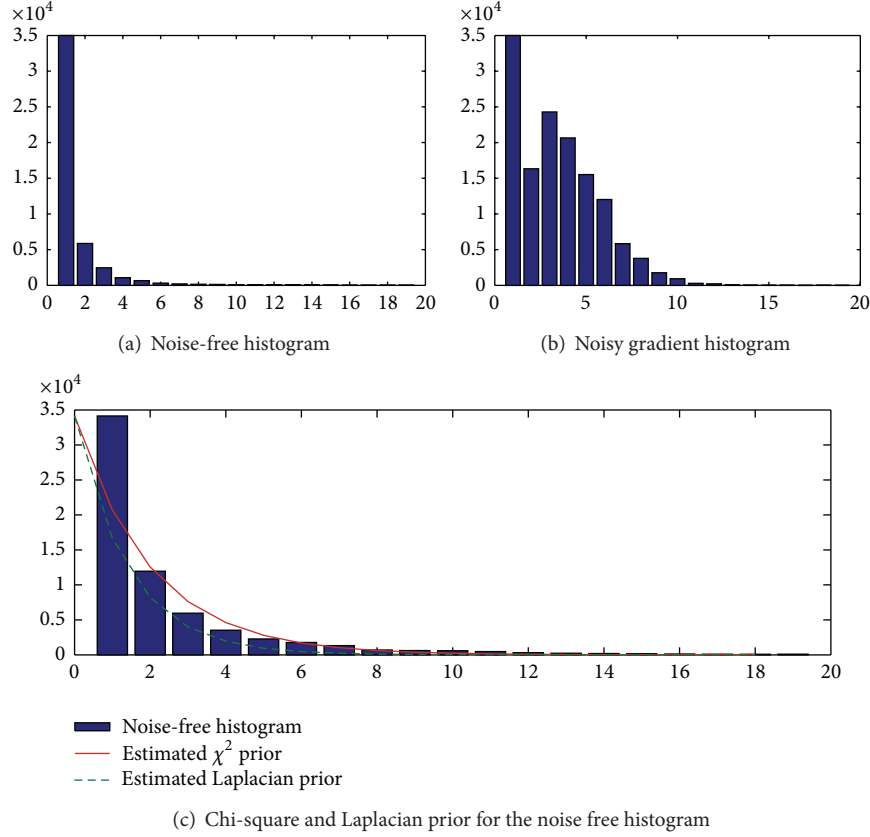(c) Chi-square and Laplacian prior for the noise free histogram

FIGURE 2: Chi-square prior versus Laplacian prior. (a) Histogram of the noise-free gradient. (b) Histogram of the noisy gradient magnitude. (c) Laplacian and Chi-square priors for the noisy-free gradient, estimated from the noisy data.

approximated distribution of the sinogram gradient can be given as Gauss$(0, \sqrt{2m})$.

Based on the assumption that the sinogram gradient follows the Gauss$(0, \sqrt{2m})$ distribution, we can show that the distribution of this gradient leads a Chi-square distribution as follows:

$$\nabla S_{(i,j)} \sim \text{Gauss}\left(0, \sqrt{2m}\right),$$
$$\frac{\left|\nabla S_{(i,j)}\right|}{\sqrt{2m}} \sim \text{Gauss}(0, 1), \qquad (29)$$
$$\frac{\left|\nabla S_{(i,j)}\right|^2}{2m} \sim \chi^2.$$

The Chi-square distribution is defined by the following probability density function:

$$P(y) = \frac{y^{\zeta/2-1} \cdot e^{-y/2}}{2^{\zeta/2}\gamma(\zeta/2)}, \qquad (30)$$

where $\gamma(\zeta/2)$ denotes the Gamma function and ζ is a positive integer that specifies the number of degrees of freedom. For the noise gradient model $x = y + n$, the Chi-square with 2

degrees of freedom (2 degrees because we are dealing with 2D images) is given as

$$P(y) = \frac{1}{2} \cdot e^{-(1/2)|y|}. \qquad (31)$$

Based on the above, the prior odd, (15), can be reformulated considering Chi-square prior instead of Laplacian prior and the noise $n \sim \text{Gauss}(0, 2\sigma_n^2)$. Note that the Chi-square with 2 degrees of freedom is almost a special case of Laplacian prior with a rate parameter (scale parameter) $\lambda = 1/2$. This parameter determines the "scale" or statistical dispersion of the probability distribution. It indicates that the distribution of the ideal gradient is independent of a rate parameter since the Poisson noise is a pixel dependent (i.e., depends on the number of counts). Therefore, it is more natural to use Chi-square prior for estimating the ideal gradient of Poisson data rather than Laplacian prior with a parameter, λ, which is based on the variances of the image and noise gradients as indicated by (16).

In Figure 2, we illustrate the estimation of the Laplacian and Chi-square priors for the noise-free gradient. Figures 2(a) and 2(b) show the histograms of the gradient magnitudes for the noise-free and noisy images, respectively. The noise-free gradient histogram is typically sharply peaked at zero since the noise-free images typically contain large portions

of relatively uniform regions that produce negligible gradient values. Sharp edges and textured regions produce some relatively large gradients, building in this way long tails of the gradient histogram. From the noisy histogram of Figure 2(b), we estimate the parameter of the prior for the noise-free sinogram gradient. The results are illustrated in Figure 2(c), which shows the estimated Laplacian (dotted curve) and Chi-square (continues curve) priors in comparison to the ideal, noise-free histogram.

3.2.3. Algorithm: The Adaptive Probabilistic Curvature Motion Filter Based on Chi-Square Prior. In summary the general equation of the proposed Adaptive Modified Probabilistic Curvature Motion (AMPCM) filter is given by

$$u_t = g_{\mathrm{apr}}\left(|\nabla u|\right) u_{vv} + \left[g_{\mathrm{apr}}\left(|\nabla u|\right) + g'_{\mathrm{apr}}\left(|\nabla u|\right)|\nabla u|\right] u_{ww}. \tag{32}$$

Let $g_1(x_i, z_i) \doteq g_{\mathrm{apr}}(x_i, z_i)$, as given by (21), and the sharpening term $g_2(x) \doteq g_1(x_i, z_i) + xg'_1(x_i, z_i)$; the overall algorithm is given as shown in Algorithm 1.

Algorithm 1. Adaptive Modified Probabilistic Curvature Motion Filter Algorithm

(i) Build the Probabilistic Diffusivity Function:

    (a) compute the prior odd based on the Chi-Square prior:

$$P\left(y\right) = \frac{1}{2} \cdot e^{-(1/2)|y|},$$
$$\mu = \left(e^{(1/2)\sigma_n} - 1\right)^{-1}. \tag{33}$$

(ii) Build the spatially adaptive diffusivity function $g_{\mathrm{apr}}(x_i, z_i)$, for each pixel $i = 1, \ldots, \mathrm{size}_f$:

    (a) compute the local spatial activity indicator

$$z_i = \frac{1}{N} \sum_{l \in w(i)} M_l, \tag{34}$$

    (b) compute the likelihood ratio for each window:

$$P_N\left(z_i \mid H_1\right) = P\left(M_i \mid H_1\right) \mathrm{Conv}_N P\left(M_i \mid H_1\right),$$
$$P_N\left(z_i \mid H_0\right) = P\left(M_i \mid H_0\right) \mathrm{Conv}_N P\left(M_i \mid H_0\right), \tag{35}$$

    (c) compute the normalizing constant A:

$$A = 1 + \mu\eta\left(0\right) \tag{36}$$

    (d) compute $P(H_1 \mid x_i)$ as defined by (20)

(iii) compute the diffusivity function $g(x_i, z_i)$:

$$g_1\left(x_i, z_i\right) = A\left(1 - P\left(H_1 \mid x_i, z_i\right)\right),$$
$$g_2\left(x_i, z_i\right) = g_1\left(x_i, z_i\right) + xg'_1\left(x_i, z_i\right). \tag{37}$$

(iv) Filter the sinogram, $u_0 = f$, based on (8) recursion, for $t = 0, \ldots,$ (*number of iterations 1*):

$$
\begin{aligned}
&u\left(x, y, t + \Delta t\right) \\
&= u\left(x, y, t\right) \\
&\quad + \frac{1}{\Delta t}\Big[g_1 \cdot \sin^2\theta \cdot u\left(x - 1, y, t\right) \\
&\qquad + g_2 \cdot \cos^2\theta \cdot u\left(x - 1, y, t\right) \\
&\qquad + g_1 \cdot \sin^2\theta \cdot u\left(x + 1, y, t\right) \\
&\qquad + g_2 \cdot \cos^2\theta \cdot u\left(x + 1, y, t\right) \\
&\qquad + g_1 \cdot \cos^2\theta \cdot u\left(x, y + 1, t\right) \\
&\qquad + g_2 \cdot \sin^2\theta \cdot u\left(x, y + 1, t\right) \\
&\qquad + g_1 \cdot \cos^2\theta \cdot u\left(x, y - 1, t\right) \\
&\qquad + g_2 \cdot \sin^2\theta \cdot u\left(x, y - 1, t\right) \\
&\qquad - 2 \cdot g_1 \cdot u\left(x, y, t\right) - 2 \cdot g_2 \cdot u\left(x, y, t\right) \\
&\qquad + \frac{g_2}{2} \cdot \sin\theta\cos\theta \cdot u\left(x + 1, y + 1, t\right) \\
&\qquad - \frac{g_1}{2} \cdot \sin\theta\cos\theta \cdot u\left(x + 1, y + 1, t\right) \\
&\qquad + \frac{g_2}{2} \cdot \sin\theta\cos\theta \cdot u\left(x - 1, y - 1, t\right) \\
&\qquad - \frac{g_1}{2} \cdot \sin\theta\cos\theta \cdot u\left(x - 1, y - 1, t\right) \\
&\qquad - \frac{g_2}{2} \cdot \sin\theta\cos\theta \cdot u\left(x + 1, y - 1, t\right) \\
&\qquad + \frac{g_1}{2} \cdot \sin\theta\cos\theta \cdot u\left(x + 1, y - 1, t\right) \\
&\qquad - \frac{g_2}{2} \cdot \sin\theta\cos\theta \cdot u\left(x - 1, y + 1, t\right) \\
&\qquad + \frac{g_1}{2} \cdot \sin\theta\cos\theta \cdot u\left(x - 1, y + 1, t\right)\Big].
\end{aligned} \tag{38}
$$

4. Experiments and Discussion

For the analysis of the proposed denoising approaches, we use a simulated thorax PET phantom, containing three hot regions of interest (tumors) of activities 1.18, 1.8, and 1.57. 100 realizations (noisy sinograms) with added noise of 1×10^6 coincident events have been generated. Each sinogram has a size of 256×256 pixels and its spacing is 2×2 (mm/pixel), with 128 detectors and 128 projection angles. Figure 3(a) illustrates the noise-free sinogram, and Figure 3(b) illustrates a noisy realization.

For reconstructing the PET images, we adopt the Ordered Subset Expectation Maximization (OSEM) reconstruction
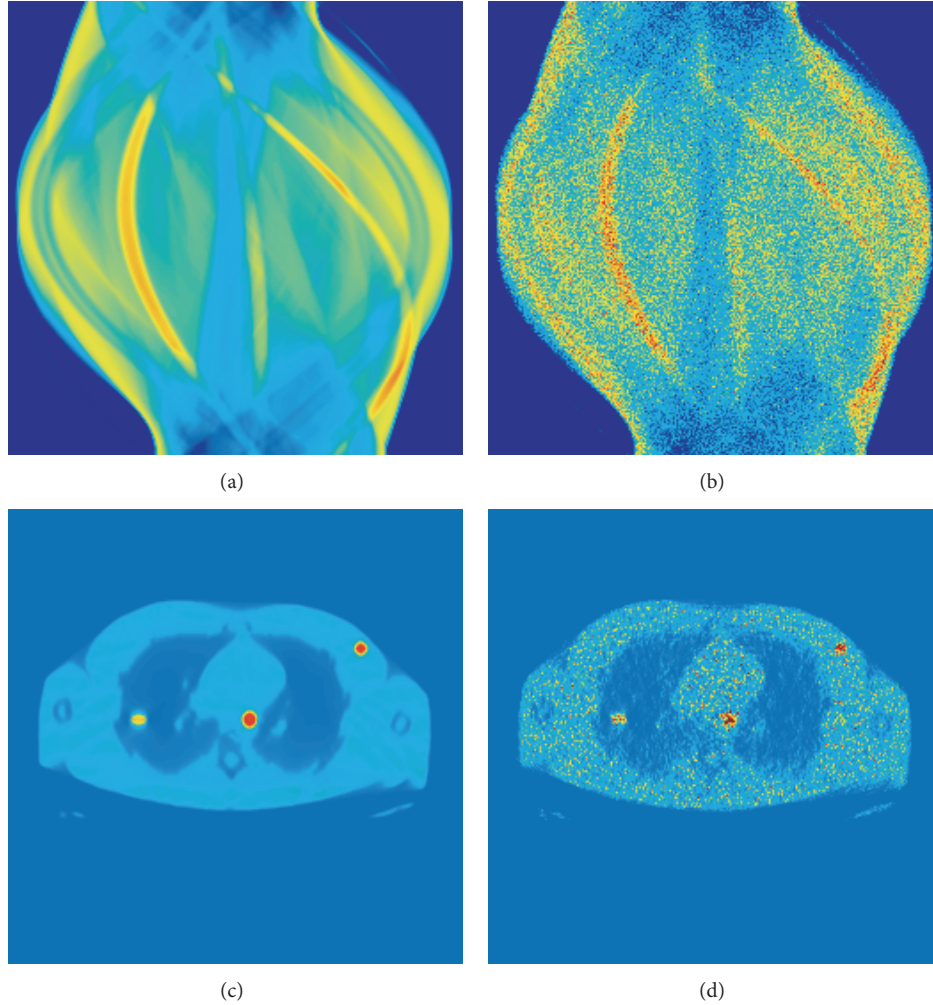
(a)

(b)

(c)

(d)

FIGURE 3: Experimental dataset. (a) Noise-free simulated sinogram. (b) Noisy sinogram. ((c)-(d)) Corresponding reconstructions.

algorithm [24]. One way to represent the imaging system is with the following linear relationship:

$$f = Hf_{\text{PET}} + n, \tag{39}$$

where f is the set of observations (sinogram), H is the known system model (projection matrix), f_{PET} is the unknown image, and n is the noise. The goal of reconstruction is to use the data values f (projections through the unknown object) to find the image f_{PET}. Under the Poisson assumption, counts of all the detectors around the object are considered as independent Poisson variables. OSEM groups projection data into an ordered sequence of subsets and progressively processes every subset of projections in each iteration process [24]. The OSEM method results are highly affected by the number of iterations and subsets used. The iterations should be ended before the noise becomes too dominant and not too early to avoid losing important information. In our experiments, the parameters of the OSEM algorithm are set as follows: we use 16 subsets and run them for 4 iterations. Such PET reconstructions are illustrated in Figures 3(c) and 3(d) for the noise-free and noisy sinograms, respectively.

For the experimental assessment of the proposed diffusion filtering, we use two sets of evaluation measures. The first set stems from measuring the quality of the filtering techniques on the sinogram, whilst the second set originates from validating the quality of the PET reconstruction, after filtering the sinogram observations. As ground-truth information, the former uses the noise-free sinogram, while the latter needs prior identification of the important areas by a medical professional.

A fundamental issue with scale-spaces induced by diffusion processes, as the ones proposed in this paper, is the automatic selection of the most salient scale. For PET sinogram denoising application, we use an earlier proposed optimal scale selection approach [25], where the maximum correlation method has been adopted:

$$
\begin{aligned}
t_{\text{opt}} &= \operatorname{argmax}\left[\widehat{\mathscr{C}}_{\text{mp}}(t)\right] \\
&= \operatorname{argmax}\left[\sigma\left[u_t\right] + \frac{\sigma\left[n_o\left(t_0\right)\right]}{\sigma\left[u_{t_0}\right]}\sigma\left[n_o(t)\right]\right]
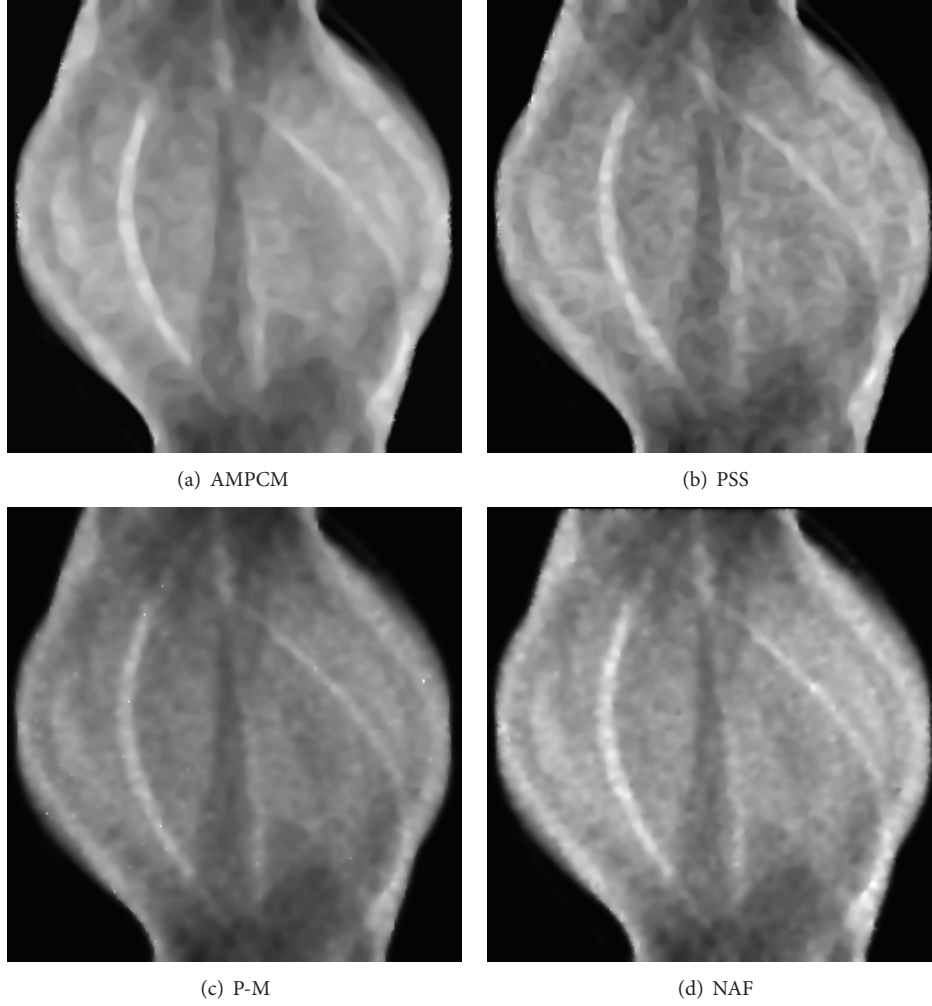\end{aligned} \tag{40}
$$

(a) AMPCM

(b) PSS

(c) P-M

(d) NAF

FIGURE 4: Optimal enhanced scale $(u_{t_{\mathrm{opt}}})$ of the noisy sinogram of Figure 3(b).

with n_o being the so-called outlier noise estimated using wavelet-based noise estimation. t_0 is the zeroth scale; thus $u_{t_0} = f$ and $n(t_0)$ represents the initial amount of noise.

4.1. Sinogram Denoising Evaluation.

In this work, we are interested in comparing the proposed AMPCM filter with filtering approaches from the literature: the Probabilistic self-snakes (PSS) [7], Perona and Malik filter (P-M) [1], and the Noise-Adaptive Nonlinear Diffusion Filtering (NAF) [14].

The Lorentzian diffusivity function $g(X) = 1/(1 + X^2/k^2)$ is used for applying the P-M filter. This function was proposed by Perona and Malik [1] and widely used in the literature. The contrast parameter k is calculated based on the value given by a percentage of the cumulative histogram of the gradient magnitude [16]. The same diffusivity constant is used for all filters ($d_t = 0.2$).

Figure 4 illustrates the optimal enhanced scale of the sinograms for all the considered filtering methods. The filtered sinogram by AMPCM and PSS shows better enhanced results especially the curvy shape features.

For the qualitative assessment of the denoised sinogram, $u_{t_{\mathrm{opt}}}$, with respect to the noise-free image I, we adopt the following measures [25]:

DQ1 The Peak Signal to Noise Ratio (PSNR) is a statistical measure of error, used to determine the quality of the filtered images. It represents the ratio of a signal power to the noise power corrupting. Obviously, one sees that the higher the PSNR, the better the quality:

$$\mathrm{PSNR}\left(t_{\mathrm{opt}}\right) = 10 \log_{10} \frac{\mathrm{Card}\,(\Omega)}{\sum_{p \in \Omega} \left| I\,(p) - u_{t_{\mathrm{opt}}}\,(p) \right|}. \quad (41)$$

DQ2 The correlation ($\mathscr{C}_{m\rho}$) between the noise-free and the filtered image. The higher this correlation is, the better the quality is:

$$\mathscr{C}_{m\rho}\left(t_{\mathrm{opt}}\right) = \rho\left[I, u_{t_{\mathrm{opt}}}\right]. \quad (42)$$

(a) AMPCM
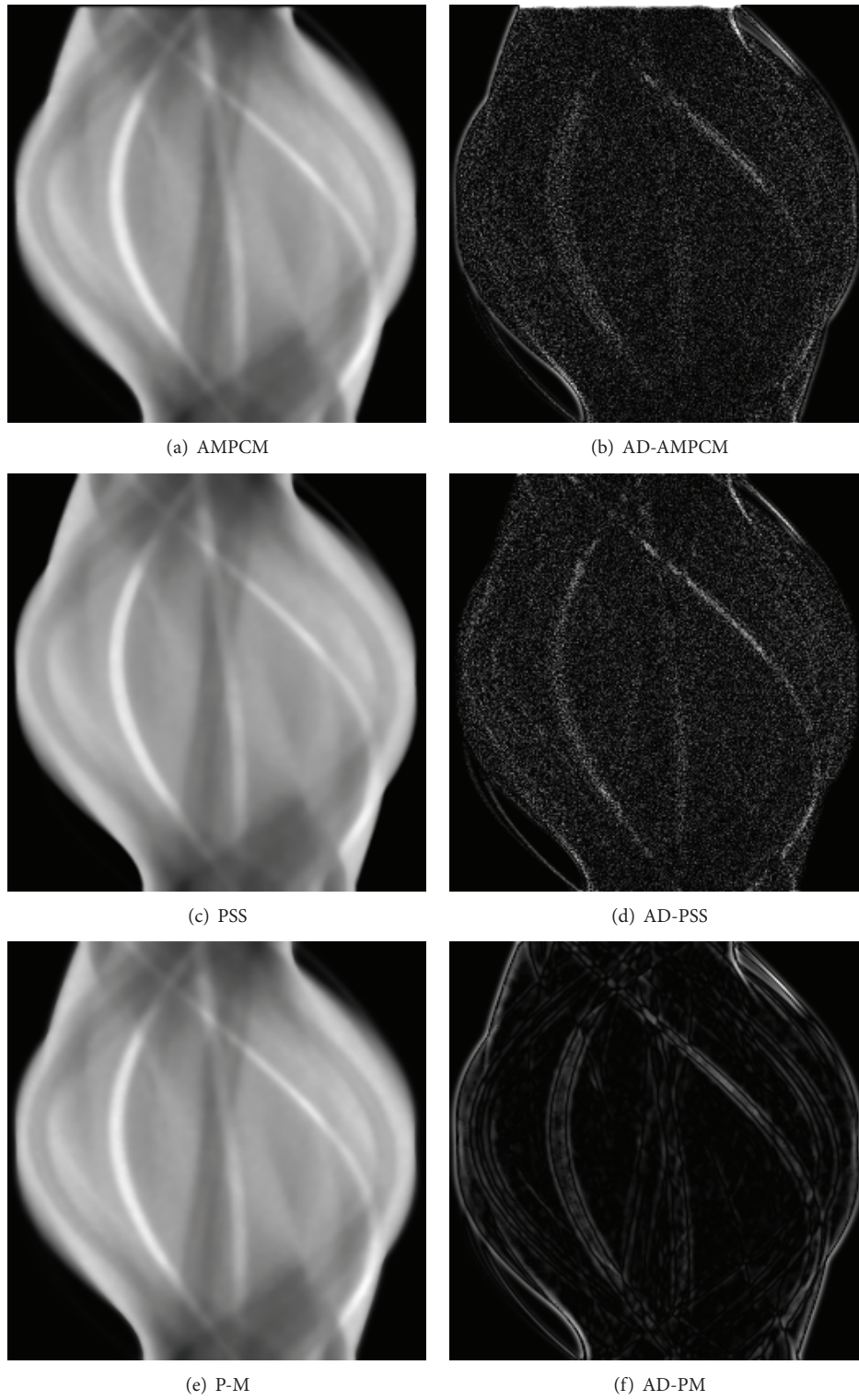
(b) AD-AMPCM

(c) PSS

(d) AD-PSS

(e) P-M

(f) AD-PM

Figure 5: Mean of the filtered sinograms (left column). Absolute difference between the mean of the filtered sinograms and the noise-free image (right column).

TABLE 1: Denoising quality measures.

Method	f	AMPCM	PSS	P-M	NAF
PSNR (t_{opt})	15.1	**30.8**	29.73	22.8	29.1
NV (t_{opt})	0.11	**0.02**	0.0223	0.225	0.0261
$\mathscr{C}_{m\rho}$ (t_{opt})	0.92	**0.9955**	0.9950	**0.9956**	0.9911

DQ3 The calculated variance of the noise (NV) describes the remaining noise level. Therefore, it should be as small as possible:

$$\text{NV}\left(t_{\text{opt}}\right) = \text{Var}\left(\left| I - u_{t_{\text{opt}}}\right|\right). \tag{43}$$

We are interested in comparing the maximum of each measure for different filtering approaches.

Table 1 lists the above measures for each of the considered filtering approaches. The best performance (maximum of the measure) is displayed in bold. As it can be seen the best performing filtering is achieved when using the AMPCM and PSS. Furthermore, we notice that for all measures, AMPCM and PSS outperform the NAF and (P-M) filters. The performance of the smoothing algorithms proposed in this paper is remarkable in discriminating a true edge from image noise (see Figure 4). We also notice the improved performance of the probabilistic curvature motion algorithms as compared to the standard diffusion algorithms. Table 2 gives the number of iterations to reach the optimal scale t_{opt} as defined in (40).

The absolute difference between the mean of the filtered sinograms and the noise-free one is another indication of the behavior and stability of the filtering methods. The mean results of the filtered realizations by the PSS and AMPCM filters are close to the ideal image, as shown in Figure 5. Comparing the mean results between the AMPCM and PSS with the other methods clearly demonstrates the effect of the sharpening/enhancing term which yields a better enhanced edges. On the other hand, P-M keeps edges unsmoothed, which is appearing as sharper edges in Figure 5. Figure 5 shows more noise remained, represented as larger values in the absolute difference images of P-M and NAF. Furthermore, using AMPCM and PSS, the absolute difference approaches zero (black image), indicating that the noise has been effectively and adaptively reduced from the noisy sinograms.

The heavy noise is clearly eliminated without destroying the texture (curves) by the probabilistic curvature motion filter. From the above and other various examples, we have observed that the proposed algorithm has ability to eliminate the Poisson noise. No stair-casing has been observed, nor severe blurring is introduced. Figure 4(c) shows that P-M filter performs well for relatively low levels of noise, while it results in poor quality of images for high noise levels. However, Perona's operator does not work well when applied to very noisy images because the noise introduces important oscillations of the gradient.

4.2. PET Reconstruction Evaluation. In this section, we use the smoothed sinograms, u_t, for PET reconstruction via the OSEM algorithm. The reconstructed PET image is denoted as

TABLE 2: Optimal number of Iterations.

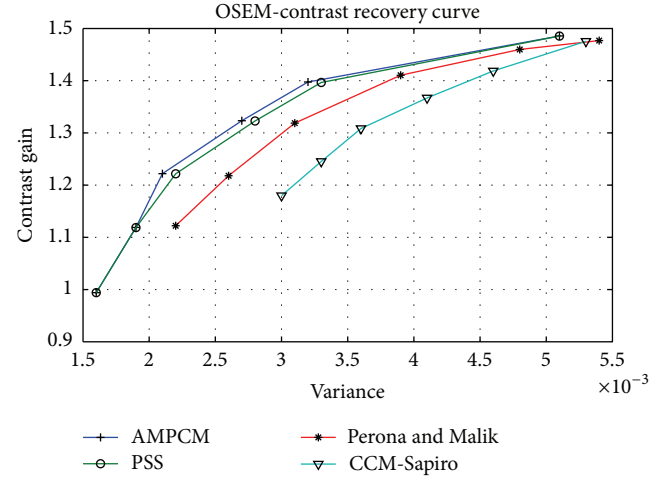| AMPCM	PSS	P-M	NAF
16	16	22	21

FIGURE 6: The contrast recovery curves for reconstructed PET by OSEM.

$f_{\text{PET}}(t)$. For evaluating the effect of sinogram predenoising on the PET reconstruction, we use the *contrast recovery method* which aims to measure the quality of $f_{\text{PET}}(t)$ by measuring the contrast recovery of the interesting objects (ROIs) in the image. The contrast recovery is computed based on the contrast gain. The latter measures how much the ROIs (i.e., tumor) in the PET image are discriminated from the background by sharp edges and on the variance of the contrast gain for different realizations. Further, the contrast gain evaluates the stability of the applied algorithm. The contrast recovery curve is estimated using the set of regions of interest that were identified by a medical professional. This curve is widely used in the literature for evaluating the quality of the reconstructed PET images [26].

In Figure 3(c), there are 3 white spots that represent tumors. We calculate the contrast gain CG_i for each realization f_{PET}^i, $i \in [1, N]$ ($N = 100$ denotes the number of realizations), and its overall variance σ_{CG}^2. Let $R = \{r_1, r_2, \ldots, r_{\text{no}}\}$ (no = 3 in our case) be the set of identified ROIs, and let B be a representative background tissue area then
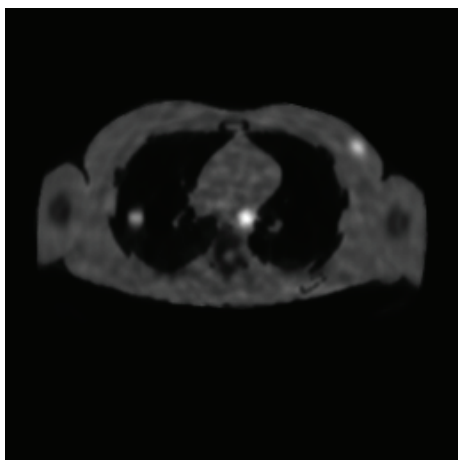
$$\text{CG}_i(t) = \frac{1}{\text{no}} \sum_{r \in R} \left[\frac{1}{\text{Card}(r)} \sum_{p \in r} f_{\text{PET}}^{(i)}(p, t) - C_{\text{bkg}} \right],$$

$$\sigma_{\text{CG}}^2(t) = \frac{1}{N} \sum_{i=1}^{N} \left(\text{CG}_i(t) - \overline{\text{CG}} \right),$$
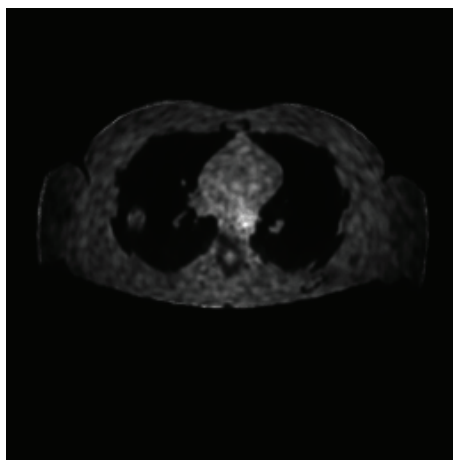
$$\tag{44}$$

where p denotes a pixel, t the scale number, and $\overline{\text{CG}}$ the contrast mean, $\overline{\text{CG}} = (1/N) \sum_{j=1}^{N} \text{CG}_j(t)^2$. C_{bkg} represent
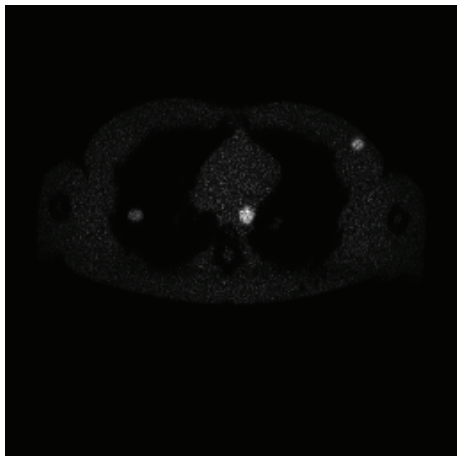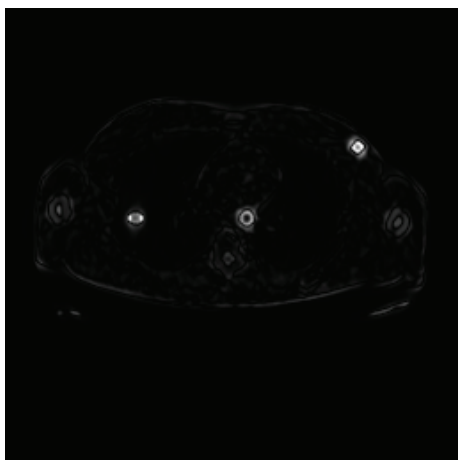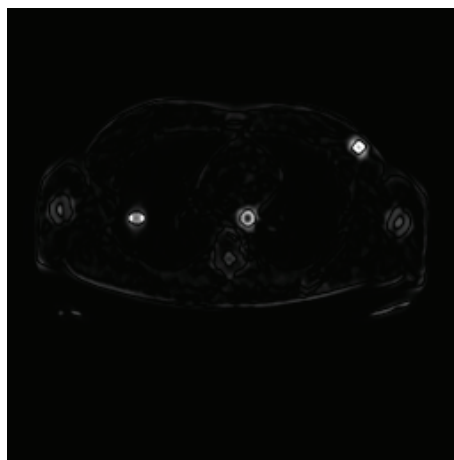
(Var-image (noisy PET))

(a) Reconstructed after AMPCM

(b) Variance image of (a).

(c) Mean results of the filtered reconstruction

(d) Difference between (c) and noise-free image

Figure 7: Evaluation of the PET reconstruction using enhanced sinograms.

the mean intensities inside background, defined as $C_{\text{bkg}} = (1/\text{Card}(B)) \sum_{p \in B} f_{\text{PET}}^{(i)}(p, t)$.

Plotting the contrast gain in function of the variance given the smoothness parameter, which is in this work, the scale parameter t yields the contrast recovery curve [26].

In order to perform, under the same conditions, the comparison of the contrast curves of the different diffusion scheme, one should identify the scales, t, emanating from different diffusion schemes, having similar information content. For this, we use the scale synchronization proposed by

Niessen et al. [27], being a Shannon-Wiener entropy used to compare nonlinear scale-space filters:

$$\Re\left[f_{\text{PET}}(t)\right] \triangleq \frac{\sigma^2\left(f_{\text{PET}}(0)\right) - \sigma^2\left(f_{\text{PET}}(t)\right)}{\sigma^2\left(f_{\text{PET}}(0)\right) - \sigma^2\left(\mathscr{M}\left(f_{\text{PET}}\right)\right)}, \qquad (45)$$

where $\mathscr{M}$ represents the averaging operator that maps an image f_{PET} onto a constant image, where each pixel equals the average feature vector.

The contrast gain and the variance were computed for 6 scales, for each filtering methods, with similar information content, selected based on Niessen et al. [27] approach. Figure 6 shows the contrast recovery curves for the considered dif-fusion schemes AMPCM, PSS, P-M, and CCM-Sapiro [17]. The best quality PET reconstruction is situated in the upper, that is, high contrast gain, left, that is, high stability, area of the plot. The figure shows that AMPCM has better contrast recovery of the interesting objects (ROIs) in the reconstructed PET images.

Figure 7 illustrates the reconstructed PET starting from the enhanced sinogram, by the AMPCM, along with the variance images of the reconstructed PET, the mean of the reconstructed PET images, and the absolute difference between the mean of the filtered PET and the noise-free image. We immediately notice that the background seems much flatter when adopting the OSEM reconstruction.

Experiments results show that combining the probabilistic diffusivity function with the curvature motion diffusion produces a powerful nonlinear filtering method that is appropriate for the unique characteristics of the PET images. The AMPCM filter preserves the boundaries of the curvy shape features and wisely smooths the regions of interest as well as the other regions. Figure 7 demonstrates that the edge enhancement around image data is significantly improved. The effect of edge strengthening is even more pronounced for weaker edges in PET images or in image areas affected by a high level of noise.

5. Conclusions

Adaptive probabilistic curvature motion filter (AMPCM) for enhancing PET images is developed and discussed in this work. The filter is applied on the 2D sinogram pre-reconstruction. For considering the special characteristics of the sinogram data, a Chi-square is used as a marginal prior for noise-free sinogram gradient in the diffusivity function.

The qualitative evaluation results results show that, among other diffusivity functions, the probabilistic adaptive function seems more effective in smoothing all the homogenous regions that contain high level of noise. Further, AMPCM retain in an accurate way the location of the edges that defines the shape of the represented structures in the sinogram. It preserves the boundaries of the curvy shape features and wisely smooths the regions of interest as well as the other regions. The application of the proposed diffusion scheme results in well-smoothed images and preserves the edges. It gains the advantages of the curvature motion diffusion and the shock filter. Further, it deals better with the problem of poor and discontinuous edges which are common in PET sinograms.

Applied as a preprocessing step before PET reconstruction, we demonstrated via qualitative measures (the contrast curve, the variance, and the mean images) that the adaptive probabilistic diffusion function has a great capability and stability to detect and enhance the important features in the reconstructed PET image, which make it a reliable and practical approach as it has no free parameters to be optimized. All parameters are image based and are automatically estimated and proved to give the best results.
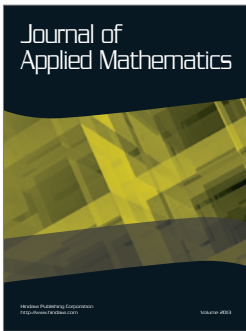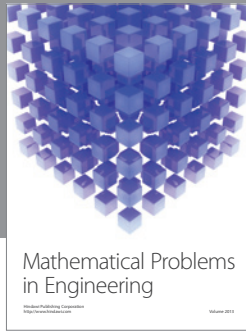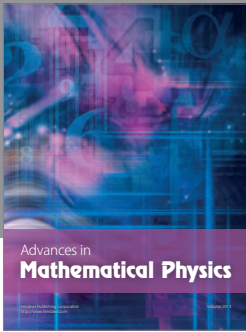
Acknowledgment

References

[1] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[2] D. Kazantsev, S. R. Arridge, S. Pedemonte et al., "An anatomically driven anisotropic diffusion filtering method for 3D SPECT reconstruction," *Physics in Medicine and Biology*, vol. 57, no. 12, pp. 3793–3810, 2012.

[3] Z. Quan and L. Yi, "Image reconstruction algorithm for positron emission tomography with Thin Plate prior combined with an anisotropic diffusion filter," *Journal of Clinical Rehabilitative Tissue Engineering Research*, vol. 15, no. 52, 2011.

[4] O. Demirkaya, "Post-reconstruction filtering of positron emission tomography whole-body emission images and attenuation maps using nonlinear diffusion filtering," *Academic Radiology*, vol. 11, no. 10, pp. 1105–1114, 2004.

[5] D. R. Padfield and R. Manjeshwar, "Adaptive conductance filtering for spatially varying noisein PET images," *Progress in Biomedical Optics and Imaging*, vol. 7, no. 3, 2006.

[6] J. Weickert, *Anisotropic Diffusion in Image Processing*, European Consortium for Mathematics in Industry, B. G. Teubner, Stuttgart, Germany, 1998.

[7] M. Alrefaya, H. Sahli, I. Vanhamel, and D. Hao, "A nonlinear probabilistic curvature motion filter for positron emission tomography images," in *Scale Space and Variational Methodsin Computer Vision*, vol. 5567 of *Lecture Notes in Computer Science*, pp. 212–2223, 2009.

[8] O. Demirkaya, "Anisotropic diffusion filtering of PET attenuation data to improve emission images," *Physics in Medicine and Biology*, vol. 47, no. 20, pp. N271–N278, 2002.

[9] W. Wang, "Anisotropic diffusion filtering for reconstruction of poisson noisy sinograms," *Journal of Communication and Computer*, vol. 2, no. 11, pp. 16–23, 2005.

[10] C. Negoita, R. A. Renaut, and A. Santos, "Nonlinear anisotropic diffusion in positron emission tomography kinetic imaging," in *SIAM Conference on Imaging Science*, Salt Lake City, Utah, USA, 2004.

[11] H. Zhu, H. Shu, J. Zhou, C. Toumoulin, and L. Luo, "Image reconstruction for positron emission tomography using fuzzy nonlinear anisotropic diffusion penalty," *Medical and Biological Engineering and Computing*, vol. 44, no. 11, pp. 983–997, 2006.

[12] H. Zhu, H. Shu, J. Zhou, X. Bao, and L. Luo, "Bayesian algorithms for PET image reconstruction with mean curvature and Gauss curvature diffusion regularizations," *Computers in Biology and Medicine*, vol. 37, no. 6, pp. 793–804, 2007.

[13] A. P. Happonen and M. O. Koskinen, "Experimental investigation of angular stackgram filtering for noise reduction of SPECT projection data: study with linear and nonlinear filters," *International Journal of Biomedical Imaging*, vol. 2007, Article ID 38516, 2007.

[14] A. A. Samsonov and C. R. Johnson, "Noise-adaptive nonlinear diffusion filtering of MR images with spatially varying noise levels," *Magnetic Resonance in Medicine*, vol. 52, no. 4, pp. 798–806, 2004.

[15] A. Pizurica, P. Scheunders, and W. Philips, "Multiresolution multispectral image denoisingbased on probability of presence of features of interest," in *Proceedings of Advanced Concepts for Intelligent Vision Systems (Acivs '04)*, 2004.

[16] I. Vanhamel, *Vector valued nonlinear diffusion and its application to image segmentation [Ph.D. thesis]*, Vrije Universiteit Brussel, Faculty of Engineering Sciences, Electronics and Informatics (ETRO), Brussels, Belgium, 2006.

[17] G. Sapiro, *Geometric Partial Differential Equations and Image Analysis*, Cambridge University Press, Cambridge, UK, 2001.

[18] A. Kuijper, "Geometrical PDEs based on second-order derivatives of gauge coordinates in image processing," *Image and Vision Computing*, vol. 27, no. 8, pp. 1023–1034, 2009.

[19] A. Pižurica, I. Vanhamel, H. Sahli, W. Philips, and A. Katartzis, "A Bayesian formulation of edge-stopping functions in nonlinear diffusion," *IEEE Signal Processing Letters*, vol. 13, no. 8, pp. 501–504, 2006.

[20] A. Teymurazyan, T. Riauka, H. S. Jans, and D. Robinson, "Properties of noise in positron emission tomography images reconstructed with filtered-back projection and row-action maximum likelihood algorithm," *Journal of Digital Imaging*, vol. 26, no. 3, pp. 447–456, 2013.

[21] G. Miller, H. F. Martz, T. T. Little, and R. Guilmette, "Using exact poisson likelihood functions in Bayesian interpretation of counting measurements," *Health Physics*, vol. 83, no. 4, pp. 512–518, 2002.

[22] H. A. Gersch, "Simple formula for the distortions in a Gaussian representation of a Poisson distribution," *American Journal of Physics*, vol. 44, no. 9, pp. 885–886, 1976.

[23] J. G. Skellam, "The frequency distribution of the difference between two Poisson variates belonging to different populations," *Journal of the Royal Statistical Society A*, vol. 109, no. 3, p. 296, 1946.

[24] X. Lei, H. Chen, D. Yao, and G. Luo, "An improved ordered subsets expectation maximization reconstruction," in *Advances in Natural Computation*, vol. 4221, pp. 968–971, 2006.

[25] I. Vanhamel, C. Mihai, H. Sahli, A. Katartzis, and I. Pratikakis, "Scale selection for compact scale-space representation of vector-valued images," *International Journal of Computer Vision*, vol. 84, no. 2, pp. 194–204, 2009.

[26] C. Comtat, P. E. Kinahan, J. A. Fessler et al., "Clinically feasible reconstruction of 3D whole-body PET/CT data using blurred anatomical labels," *Physics in Medicine and Biology*, vol. 47, no. 1, pp. 1–20, 2002.

[27] W. J. Niessen, K. L. Vincken, J. A. Weickert, and M. A. Viergever, "Nonlinear multiscale representations for image segmentation," *Computer Vision and Image Understanding*, vol. 66, no. 2, pp. 233–245, 1997.