

Parametric surface denoising

Ioannis A. Kakadiaris^a, Ioannis Konstantinidis^a, Manos Papadakis^b, Wei Ding^a, Lixin Shen^c,
Donald Kouri^d and David Hoffman^e

^aComputational Biomedicine Lab, Dept. of Computer Science, Univ. of Houston, Houston,
TX 77204;

^bDept. of Mathematics, Univ. of Houston, Houston, TX 77204;

^cDept. of Mathematics, Western Michigan University, Kalamazoo, MI 49008;

^dDepts. of Chemistry and Physics, Univ. of Houston, Houston, TX 77204;

^eDept. of Chemistry and Ames Laboratory, Iowa State University, Ames, IA 50011

ABSTRACT

Three dimensional (3D) surfaces can be sampled parametrically in the form of range image data. Smoothing/denoising of such raw data is usually accomplished by adapting techniques developed for intensity image processing, since both range and intensity images comprise parametrically sampled *geometry* and *appearance* measurements, respectively. We present a transform-based algorithm for surface denoising, motivated by our previous work on intensity image denoising, which utilizes a non-separable Parseval frame and an ensemble thresholding scheme. The frame is constructed from separable (tensor) products of a piecewise linear spline tight frame and incorporates the weighted average operator and the Sobel operators in directions that are integer multiples of 45° . We compare the performance of this algorithm with other transform-based methods from the recent literature. Our results indicate that such transform methods are suited to the task of smoothing range images.

Keywords: Surface denoising, tight frame, wavelets

1. INTRODUCTION

Range images represent surface *geometry* in a parametric fashion. Range data are a sampled set of measurements corresponding to *distances* from a specific viewpoint (the detector location) to surface points on the objects in a three dimensional (3D) scene. This distance is measured by means of a laser beam which is reflected by the object surface onto a monochromatic detector. For each sampled point on the object surface three separate measurements are recorded: the distance d and the deflection angles (ϕ, θ) of the laser beam that it corresponds to. These measurements constitute an explicit parametric description of the object surface $d = F(\phi, \theta)$, and can be transformed to position coordinates (x, y, z) with respect to an orthogonal axes system by trigonometric calibration.

Given a scene of 3D objects, its contents can also be represented by an image capturing intensity data. Intensity images therefore represent surface *appearance*. They also comprise a sampled set of measurements, this time corresponding to the intensity of light (usually forming a broad part of the visible spectrum) as it is reflected by points on the object surface. Thus, intensity images provide a parametric description $I = F(u, v)$ of the reflected light intensity of the object surface or background, with intensity I samples measured at pixel positions (u, v) .

It must be noted that the parametrization obtained using range data is neither conformal nor area preserving, and hence it is far from isometric, as the spacing of the samples does not depend on the incident angle, i.e., the angle between the laser beam and the surface normal. In contrast, the pointing error in estimating d does depend on the incident angle,¹ and this is not the case for intensity images. However, both types of images are also affected by the presence of additive Gaussian noise due to sensor electronics.

Corresponding author: I.A. Kakadiaris, E-mail: ioannisk@uh.edu

It is due to this affinity, and despite their differences, that range images do lend themselves to treatment using techniques originally developed for intensity images. We present in this paper the results of adapting algorithms developed during our previous work on intensity image denoising² to the task of smoothing surface data in the form of such range images, and compare performance with other similar transform-based algorithms.

The remainder of this paper is organized as follows. Section 2 provides an overview of denoising methods for range/intensity images. In Section 3, we review the construction of a frame-based filtering scheme that is tuned to detecting singularities of first order at multiple orientations. This filtering scheme is employed in Section 3.2 for the construction of an ensemble thresholding denoising algorithm. Our experimental results are presented and discussed in Section 4.

2. BACKGROUND

Most of the existing work regarding smoothing/denoising surface data centers around two main themes. First, the techniques applied are suitable modifications of algorithms developed for intensity images, and second, they involve the formulation of an elliptic PDE which is then solved via an iterative algorithm. In our work, we explore an alternative approach, motivated by the success in image denoising of recent transform-based algorithms (wavelets/frames).

Transform-based noise reduction has a long history as an area of interest in the (intensity) image processing community. The goal of denoising is accomplished via a combination of methods involving suitable filtering/transforms and statistical estimation. Typically, the image is transformed onto some domain where the noise component can be identified more easily, and a statistical estimation is performed to identify and remove its influence. This is accomplished by an estimation operator that suppresses noise while preserving the true signal. Finally, the transformation is inverted to obtain the denoised image.

2.1. PDE-based range image denoising

Saint-Marc *et al.*³ in their seminal work apply adaptive smoothing to the range image derivatives in order to achieve a robust detection of curvature features. Barash⁴ follows up examining the relationship between anisotropic diffusion, adaptive smoothing and bilateral filtering, explaining how adaptive smoothing can serve as the link between bilateral filtering and anisotropic diffusion. In this general vein, Umasuthan⁵ developed heat diffusion techniques, and Desbrun *et al.*⁶ use the mean curvature flow. Also, Boulanger *et al.*⁷ use a multi-scale filtering technique which produces a scale-space filtering analogous to Gaussian filtering, taking into account a physical model of the sensor to ensure optimum filtering of the signal.

2.2. Transform-based intensity image denoising

In recent years, a wide class of image denoising algorithms have been based on the discrete wavelet transform. The usefulness of the wavelet transform was first demonstrated by Donoho and Johnstone,⁸⁻¹⁰ when they proved that thresholding estimators in a wavelet basis have nearly minimax risk for sets of piecewise regular images. For the case of additive Gaussian noise they suggested two thresholding functions:

$$\begin{aligned} \text{soft-threshold} & : \eta_T^S(x) = \text{sgn}(x) \cdot \max(|x| - T, 0) \\ \text{hard-threshold} & : \eta_T^H(x) = \begin{cases} 0 & |x| \leq T, \\ x & \text{otherwise} \end{cases} \end{aligned}$$

where T is the threshold level and x is the wavelet coefficient of the underlying image. The threshold T is to be selected using VisuShrink⁸ or SureShrink.¹⁰

Coifman and Donoho¹¹ established that the use of non-decimated transforms minimizes artifacts in the denoised data; the translation invariant denoising scheme they proposed is equivalent to thresholding in the shift-invariant redundant representation implemented by a non-subsampled filter bank, or frame. In addition, it has been shown¹²⁻¹⁵ that a redundant representation is substantially superior to a non-redundant representation for image denoising in terms of mean-squared error and signal-to-noise ratio. Several redundant representations have been applied to image denoising. Simoncelli *et al.*¹⁶ introduced the “steerable pyramid”, a tight frame also

used by image denoising algorithms.¹⁷ The dual-tree complex wavelet transform introduced by Kingsbury¹⁸ is also redundant, and it has been used by Sendur and Selesnick.^{19,20}

In order to improve the selection criteria for the threshold T one needs to depart from the minimax framework, which is optimal when no *a priori* information about the signal itself is assumed, and move to a Bayesian approach, where both the noise and the true image signal coefficients in the wavelet domain can be modeled using some prior distribution. This approach was first used by Simoncelli and Adelson.^{12,21} This is also the approach used by Chang *et al.*,^{13,22} where *BayesShrink*, an adaptive, data-driven threshold, is derived in a Bayesian framework using context modeling of the global coefficients histogram. Another approach exploiting the local structure of wavelet coefficients was proposed by Mihçak *et al.*²³ Recently, Portilla *et al.*¹⁷ developed a local Gaussian Scale Mixture (GSM) model with excellent results, while Pizurica *et al.*²⁴ use an empirical model to estimate the prior.

Even better results can be obtained by exploiting the fact that wavelet coefficients are statistically dependent.^{25,26} For example, a binary Hidden Markov model (HMM) based on wavelet trees for denoising of one-dimensional (1D) signals was proposed by Crouse *et al.*²⁷ and extended to image denoising by Romberg *et al.*²⁸ It is interesting to note that the GSM model¹⁷ can be interpreted as a continuous form of this binary HMM. HMMs were also used by Fan.²⁹ Sendur and Selesnick^{19,20} propose a new bivariate model (BISHRINK) by considering dependencies between the wavelet coefficients and their parents. In general, redundant representations such as frames are preferable for denoising purposes. Designing frames also proves to be more efficient than designing bases, especially when one designs Riesz or orthonormal bases in multi-dimensions arising from scaling functions.³⁰

3. METHODS

In this section, we review the construction of a filter bank implementing a Parseval frame, as described in our previous work.² These framelet filters include the Sobel operators at directions which are integer multiples of 45° . We then describe the ensemble thresholding scheme used to denoise surface data parametrically represented as range images.

3.1. Notation

We start by briefly establishing our notation. The Hilbert space of digital signals we wish to work with in our applications is $\ell^2(\mathbb{Z}^d)$, where $d = 2$. An element K of $\ell^2(\mathbb{Z}^d)$ is a *digital filter* if its Fourier transform \widehat{K} is a bounded function. In order to eliminate any possible ambiguity we define the Fourier transform of $L^1([-\pi, \pi])$, by

$$\widehat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\omega) e^{-in\omega} d\omega, \quad n \in \mathbb{Z}.$$

This definition imposes a scaled version of the definition of the usual inner product in $L^2([-\pi, \pi])$ as follows: $\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\omega) \overline{g(\omega)} d\omega$, where $f, g \in L^2([-\pi, \pi])$.

A digital filter K acts on a digital signal $s \in \ell^2(\mathbb{Z}^d)$ by convolution (i.e., $s \rightarrow s * K$). A family $\{x_j : j \in J\}$ of elements of a Hilbert space \mathcal{H} , where J is an index set, is a *frame* for \mathcal{H} if there exist constants $0 < A \leq B < \infty$ such that for every $x \in \mathcal{H}$ we have that:

$$A\|x\|^2 \leq \sum_{j \in J} |\langle x, x_j \rangle|^2 \leq B\|x\|^2.$$

We refer to the optimal positive constants A, B as *frame bounds*. We refer to the frame as *tight* if $A = B$, and as *Parseval frame* (PF) if $A = B = 1$. The advantage of PF versus other types of frames is that the same set of vectors can be used for decomposition and reconstruction, and signal energy remains constant, just as in the case of orthonormal bases. Exact frames (frames having no redundancy) are Riesz bases and vice-versa. We also say that a finite set of filters $\{K_i : i = 0, 1, \dots, l\}$ *generates* a frame of $\ell^2(\mathbb{Z}^d)$ if the family $\{T_{\mathbf{n}}K_i : \mathbf{n} \in \mathcal{H}, i = 0, 1, \dots, l\}$ is a frame of $\ell^2(\mathbb{Z}^d)$.

3.2. Constructing new framelet digital filters

We begin by stating the following general result from our previous work.² Assume that K_i , with $i = 0, 1, \dots, l$ is a finite set of digital filters that generates a frame for $\ell^2(\mathbb{Z}^d)$. For a given positive integer p , let U be a $2\pi\mathbb{Z}^d$ -periodic $(p+1) \times (l+1)$ matrix-valued function whose entries $(U(\omega))_{q,r}$ are continuous (or more generally, measurable and bounded functions).

PROPOSITION 1. *If there exists $A_1 > 0$ such that for almost every $\omega \in [-\pi, \pi]^d$*

$$A_1 \|\mathbf{x}\| \leq \|U(\omega)\mathbf{x}\| \quad \text{for all } \mathbf{x} \in \mathbb{C}^{l+1},$$

then the matrix multiplication $U(\omega)(\widehat{K}_0(\omega), \widehat{K}_1(\omega), \dots, \widehat{K}_l(\omega))^t$, defines a new family of digital filters which generate a frame for $\ell^2(\mathbb{Z}^d)$. If, in particular, $U(\omega)$ is an isometry, for almost every $\omega \in [-\pi, \pi]^d$, then the resulting and the original frames have the same frame bounds.

The subject of three-channel filter banks based on tight frames that are derived from multiresolution analyses has an extensive bibliography.^{31–36}

We use Proposition 1 to lift the frame described by Ron and Shen³⁷ as being the simplest example of a compactly supported tight spline frame. We select splines of degree one in order to keep the number of wavelets to a minimum. The 1D Parseval frame is generated by the following low-pass h_0 , bandpass h_1 , and high-pass h_2 filters:

$$h_0 = \frac{1}{4} [1, 2, 1], \quad h_1 = \frac{\sqrt{2}}{4} [1, 0, -1], \quad \text{and} \quad h_2 = \frac{1}{4} [-1, 2, -1].$$

Their corresponding inverse Fourier transforms are, respectively:

$$m_0(\omega) = \cos^2 \frac{\omega}{2}, \quad m_1(\omega) = i \frac{\sqrt{2}}{2} \sin \omega, \quad \text{and} \quad m_2(\omega) = \sin^2 \frac{\omega}{2}.$$

It is easy to note that

$$|m_0(\omega)|^2 + |m_1(\omega)|^2 + |m_2(\omega)|^2 = 1, \quad \text{for } \omega \in [-\pi, \pi], \quad (1)$$

and therefore h_0, h_1 and h_2 generate a Parseval frame for $\ell^2(\mathbb{Z})$.

The tensor product of the 1D PF of Ron and Shen with itself is another PF arising from nine separable (tensor product) filters:

$$m_{p,q}(\omega_1, \omega_2) = m_p(\omega_1)m_q(\omega_2),$$

where $p, q \in \{0, 1, 2\}$. The following equality holds:

$$\sum_{p,q=0}^2 |m_{p,q}(\omega)|^2 = 1, \quad \text{for } \omega \in [-\pi, \pi]^2. \quad (2)$$

Thus, we view $m_{0,0}$ as a low-pass filter, and the remaining eight filters as band-pass and high-pass. We refer to these filters collectively as the UHF9 filter bank. The filter taps corresponding to it are given by the nine 3×3 matrices $M_{p,q} := h_p^t h_q$.

We note that the filters $M_{0,1}$ and $M_{1,0}$ in the UHF9 filter bank are the Sobel operators detecting vertical and horizontal edges. This motivates us to augment bank UHF9 with two diagonal first order singularity detectors. Equation (2) implies that for every ω in $[-\pi, \pi]^2$, the vector $v = v(\omega)$ defined by:

$$v := (m_{0,0}, m_{0,1}, m_{1,0}, m_{1,1}, m_{0,2}, m_{2,0}, m_{2,1}, m_{1,2}, m_{2,2})^t$$

is a unit vector in \mathbb{C}^9 . To construct the new filters, we first “clone” the pair of filters $(m_{0,1}, m_{1,0})$ into the quadruplet $(m_{0,1}, m_{1,0}, m_{0,1}, m_{1,0})$. To achieve this isometrically, we define the mapping D , where

$$D := \frac{1}{\sqrt{2}} \begin{bmatrix} I_{2 \times 2} \\ I_{2 \times 2} \end{bmatrix}.$$

Since the columns of D form an orthonormal subset of \mathbb{C}^4 , we obtain that D defines an isometry mapping from \mathbb{C}^2 into \mathbb{C}^4 . We then use the rotation matrix R , where

$$R := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix},$$

to entwine the two new copies while leaving the originals unchanged. In summary, we apply the isometry matrix E from \mathbb{C}^2 into \mathbb{C}^4 , where

$$E := \begin{bmatrix} I_{2 \times 2} & 0 \\ 0 & R \end{bmatrix} D = \frac{1}{2} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \\ 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Next, by considering the decompositions $\mathbb{C}^9 = \mathbb{C} \oplus \mathbb{C}^2 \oplus \mathbb{C}^6$ and $\mathbb{C}^{11} = \mathbb{C} \oplus \mathbb{C}^4 \oplus \mathbb{C}^6$, we define the matrix-valued function

$$U(\omega) := \begin{bmatrix} 1 & 0 & 0 \\ 0 & E & 0 \\ 0 & 0 & I_{6 \times 6} \end{bmatrix},$$

which is also unitary, since E and thus $U(\omega)$, are isometries for every ω . Thus, Proposition 1 applies to $v_1(\omega) := U(\omega)v(\omega)$, and $v_1(\omega)$ gives rise to an 11-band filter bank, with the integer translates of the filters defined by the coordinate functions of v forming a PF of $\ell^2(\mathbb{Z}^2)$. We call this the UHF11 filter bank.

In summary, the UHF11 filter bank is implemented by the following filters: $K_0 = M_{0,0}$ (low pass filter), $K_1 = \frac{\sqrt{2}}{2}M_{0,1}$, $K_2 = \frac{\sqrt{2}}{2}M_{1,0}$, $K_3 = \frac{1}{2}(M_{0,1} + M_{1,0})$, $K_4 = \frac{1}{2}(M_{0,1} - M_{1,0})$, $K_5 = M_{1,1}$, $K_6 = M_{0,2}$, $K_7 = M_{2,0}$, $K_8 = M_{1,2}$, $K_9 = M_{2,1}$, and $K_{10} = M_{2,2}$. The technique employed to construct the UHF11 filter bank relies on Proposition 1 in two ways. First, to construct isometries that increase the redundancy of an existing filter-based PF by producing scaled duplicates of a subset of those filters; and second, to apply unitary operators on selected collections of those filters to produce new filters with desirable characteristics (e.g., geometry). In both cases the resulting frame is a Parseval frame. This is a versatile technique that can be employed in a much more general setting than the two examples discussed above.

In addition, several key properties of the Ron and Shen frame are also lifted to the filter banks constructed in our examples. We note that the second and third coordinates of v_1 still define the Sobel operators (scaled by $\frac{1}{\sqrt{2}}$) in the horizontal and vertical directions, respectively. In addition, the impulse responses of the fourth and fifth coordinates of v_1 are given by:

$$\frac{\sqrt{2}}{8} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix} \quad \text{and} \quad \frac{\sqrt{2}}{8} \begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix},$$

respectively. These two filters act as derivatives parallel to the directions $-\frac{\pi}{4}$ and $\frac{\pi}{4}$ or, equivalently, as Sobel operators in these directions.

3.3. Denoising Algorithm

Let \mathbf{X} be the surface data in the form of a range image. We filter \mathbf{X} using the UHF11 filter bank constructed in the previous section. We stress that the output of each channel in this filter bank is undecimated. Let \mathbf{Y}_m be the output of the image \mathbf{X} through the m -th band (i.e., $\mathbf{Y}_m = \mathbf{X} * K_m$). We separate the ten high-pass subband outputs into two groups, $m = 1, \dots, 5$ and $m = 6, \dots, 10$, respectively. Note that filters in the first group can be used to detect first order singularities, while filters in the second group can be used to detect second order singularities. Accordingly, we choose different thresholding strategies for each of the two groups. Although we do not have a mathematical proof yet, these strategies have performed well in our denoising experiments.

For the first group, we modify the coefficients in the \mathbf{Y}_m , $m = 1, \dots, 5$ using the hard threshold operator η_T^H , where $T = \alpha \cdot \sigma_n \sqrt{2 \log N}$, α is a thresholding factor, N is the number of pixels in \mathbf{X} , and σ_n is the noise variance.

The threshold $\sigma_n \sqrt{2 \log N}$ is a good choice for large values of N when a unitary wavelet transform is used.⁸ However, the transforms induced by convolution with K_m are only isometric, and not unitary. This results in an overall reduction of the energy contribution of the noise in the transformed image.³⁸ Therefore, the threshold needs to be scaled by a factor α , where $0 < \alpha < 1$, which is selected experimentally. If σ_n is not known, it is estimated by the robust median estimator:

$$\hat{\sigma} = \frac{\text{Median}(|\mathbf{Y}_{\text{Haar}}[i, j]|)}{0.6745},$$

where \mathbf{Y}_{Haar} is the output of \mathbf{X} using 1-level Haar high-pass filtering.

Our proposed algorithm jointly thresholds \mathbf{Y}_1 and \mathbf{Y}_2 to obtain $\tilde{\mathbf{Y}}_1, \tilde{\mathbf{Y}}_2$. It should be noted that the proposed shrinkage of the wavelet coefficients is not the same as the classical wavelet shrinkage. For $m = 1, 2$:

$$\tilde{\mathbf{Y}}_m = \begin{cases} \mathbf{Y}_m, & \text{if } |\mathbf{Y}_m| > T_1 \text{ or } (|\mathbf{Y}_3| \text{ or } |\mathbf{Y}_4|) > T_2 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

where $T_1 = (1/2)T$ and $T_2 = (1/8)(2 + \sqrt{2})T$. The scaled thresholds T_1, T_2 are obtained by computing the maximum magnitude of the response of the filters K_1 and K_4 , respectively. Similarly,

$$\tilde{\mathbf{Y}}_m = \begin{cases} \mathbf{Y}_m, & \text{if } |\mathbf{Y}_m| > T_2 \text{ or } (|\mathbf{Y}_1| \text{ or } |\mathbf{Y}_2|) > T_1 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

for $m = 3, 4$. For \mathbf{Y}_5 , we use the hard thresholding operator directly:

$$\tilde{\mathbf{Y}}_5 = \eta_{T_1}^H(\mathbf{Y}_5). \quad (5)$$

Outputs in the second group are denoised by applying locally adaptive window-based denoising using maximum *a posteriori* estimates (LAWMAP), a method proposed by Mihçak *et al.*²³ We assume that the coefficients $\mathbf{Y}_m[i, j]$ are independent zero-mean Gaussian variables with unknown variance $\sigma^2[i, j]$. An estimate of $\sigma^2[i, j]$ is formed based on a local neighborhood $\mathcal{N}_{i, j}$ which is a square window of size M centered at $\mathbf{Y}_m[i, j]$. We postulate an exponential prior $f_\sigma(\sigma^2) = \lambda e^{-\lambda \sigma^2}$. Then the maximum *a posteriori* (MAP) estimator for $\sigma^2[i, j]$, using the exponential prior, is given by:

$$\tilde{\sigma}^2[i, j] = \frac{M}{4\lambda} \left[-1 + \sqrt{1 + (8\lambda/M^2) \sum \mathbf{Y}_m^2[p, q]} \right] - \sigma_n^2,$$

where the sum is over all $[p, q]$ in $\mathcal{N}_{i, j}$. We impose a positivity condition by setting all negative estimates equal to zero, as suggested by Mihçak *et al.*,²³ since it is possible to obtain negative values from the actual MAP estimate if M is too small. Thus, we use $\sigma^2[i, j] = \max(0, \tilde{\sigma}^2[i, j])$. With the estimated $\sigma[i, j]$ and σ_n , we apply a Wiener (least-squares fit) filter to all $\mathbf{Y}_m[i, j] \in \mathbf{Y}_m$, $m = 6, \dots, 10$:

$$\tilde{\mathbf{Y}}_m[i, j] = [(\sigma^2[i, j]) / (\sigma^2[i, j] + \sigma_n^2)] \mathbf{Y}_m[i, j]. \quad (6)$$

We can further decompose the output \mathbf{Y}_0 and denoise the wavelet outputs using the above described process. Our algorithm can be summarized as follows:

Parametric Surface Denoising Algorithm (PSDA)

1. Initialization: Input the raw range image \mathbf{X} , a threshold factor α , and the number of decomposition levels J .
2. Decompose the range image \mathbf{X} up to level J using the UHF11 filter bank to obtain \mathbf{Y}_m , $m = 0, \dots, 10$.
3. Compute $\tilde{\mathbf{Y}}_m$, $m = 1, \dots, 10$ using Equations (3)–(6).
4. Reconstruct the surface data from the denoised range image $\tilde{\mathbf{X}}$ from \mathbf{Y}_0 and $\tilde{\mathbf{Y}}_m$, $m = 1, \dots, 10$ by using the UHF11 filter bank.

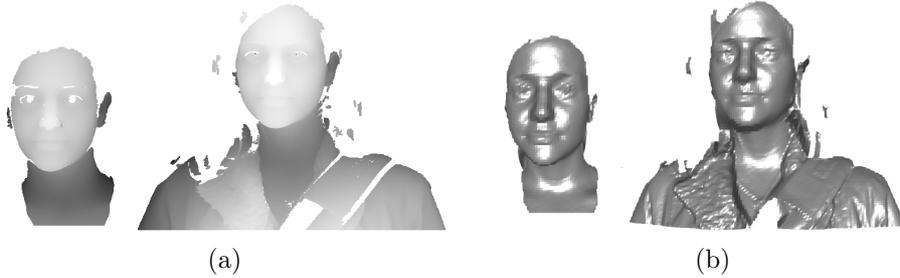


Figure 1. Sample face surface data from the FRGC data corpus; (a) range data and (b) reconstructed 3D meshes.

4. RESULTS AND DISCUSSION

We present in this section selected results of our experimentation on a large corpus of 3D range data of human faces. The range data were provided as part of the Face Recognition Grand Challenge (FRGC) data corpus.³⁹ Figure 1 presents examples obtained from the same individual at different times. All data were acquired using a Minolta Vivid 900/910 series sensor at the University of Notre Dame during the 2003-2004 academic year from 466 subjects in a total of 4007 subject sessions. The Vivid scanner is a structured light sensor that takes a 640 by 480 range sampling. Subjects were at an approximate distance of 1.5m from the sensor.

We tested PSDA against various other methods reported in the literature. Here, we present results comparing it to two other frame-based methods. The first is implemented by the BISHRINK algorithm described by Sendur and Selesnick,²⁰ and the second by the BLS-GSM algorithm described by Portilla *et al.*¹⁷ Again, we acknowledge the fact that these methods were developed and tuned to the realm of natural intensity images; results comparing our method to the above on intensity image data were reported in our previous work.² We note that all three methods exhibit comparable performance in the case of intensity images. We have performed numerous experiments. Due to space limitations, here we present two set of experiments to assess their performance on parametric surface data.

For the first set of experiments, the original surface data in the form of range images were smoothed using each of the three methods mentioned above (Fig. 2). We note that PSDA and BLS-GSM achieve the best reduction in noise, while preserving surface details, even in areas of high principal curvature fluctuation (e.g., around the eyes). For the second set of experiments, we used smoothed images obtained using PSDA as the ground truth. Gaussian noise of several levels was added and the resulting noised image was denoised using all three methods. Results are presented for visual inspection in Fig. 3. Again we note that results are significantly cleaner than the noised model.

We also performed a quantitative assessment using the normal error comparison metric introduced by Ohtake *et al.*⁴⁰ Normal error is defined as follows. Let T represent a triangle of area $A(T)$ in a triangle mesh M . The unit normal to T will be denoted by \mathbf{n}_T . Given a corresponding triangle mesh $f(M)$ (e.g., one obtained after adding noise or denoising), we define the normal error for a vertex $P \in M$ to be:

$$E(P, f(P)) = \frac{1}{3} \sum_{P \in T} A(T) \|\mathbf{n}_T - f(\mathbf{n}_T)\|^2,$$

and the total mesh normal error to be:

$$E(M, f(M)) = \sum_{P \in M} E(P, f(P)).$$

Notice that we have added Gaussian noise along unit normal of each mesh triangle T , thus the number of mesh triangles are not altered. Similarly, denoising methods, PSDA, BISHRINK and BLS-GSM, affect the location of vertices, but do not change mesh topology. Table 1 presents a quantitative assessment using this metric, and Fig. 4 presents the same data in diagram form. Our method exhibits the best performance among all three in

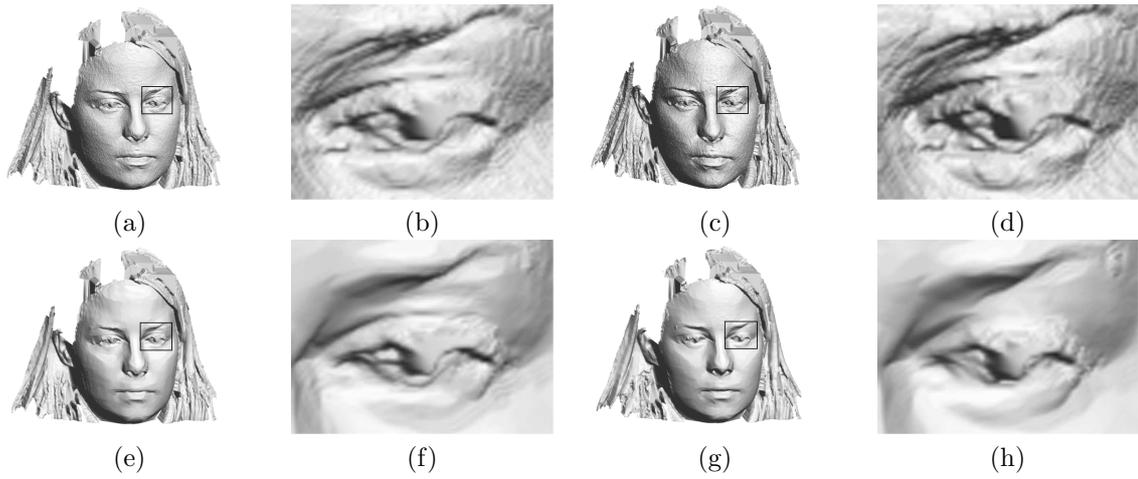


Figure 2. Data and zoomed window of (a)(b)original and result after smoothing using the (b)(c) BISHRINK, (e)(f) BLS-GSM and (g)(h) PSDA algorithms.

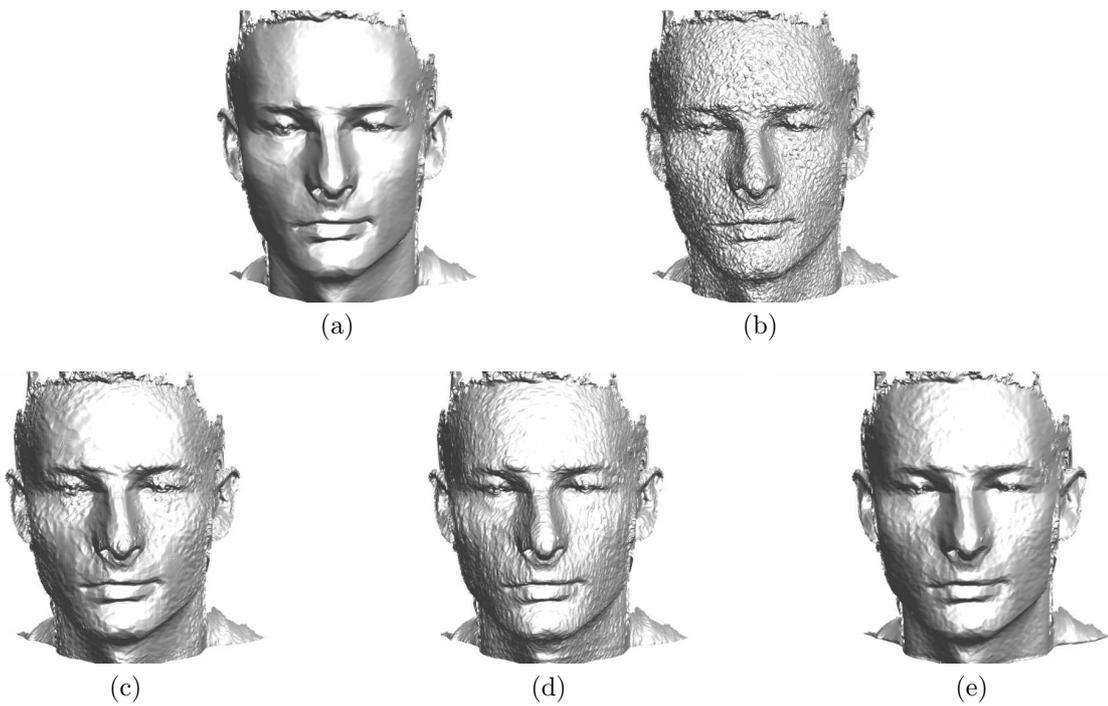


Figure 3. (a) Smoothed data (using the PSDA algorithm) and (b) Gaussian noise ($\sigma = 1.6$) added. Results after denoising using the (c) BISHRINK, (d) BLS-GSM and (e) PSDA algorithms.

σ	noised	BISHRINK	BLS-GSM	PSDA
0.5	0.230	0.127	0.196	0.142
1.0	0.310	0.178	0.233	0.164
1.2	0.336	0.194	0.247	0.172
1.4	0.361	0.209	0.256	0.180
1.6	0.384	0.223	0.264	0.188
2.0	0.427	0.246	0.288	0.213

Table 1. Performance evaluation; average (using 24 datasets) normal error (in mm^2) for BISHRINK, BLS-GSM and PSDA.

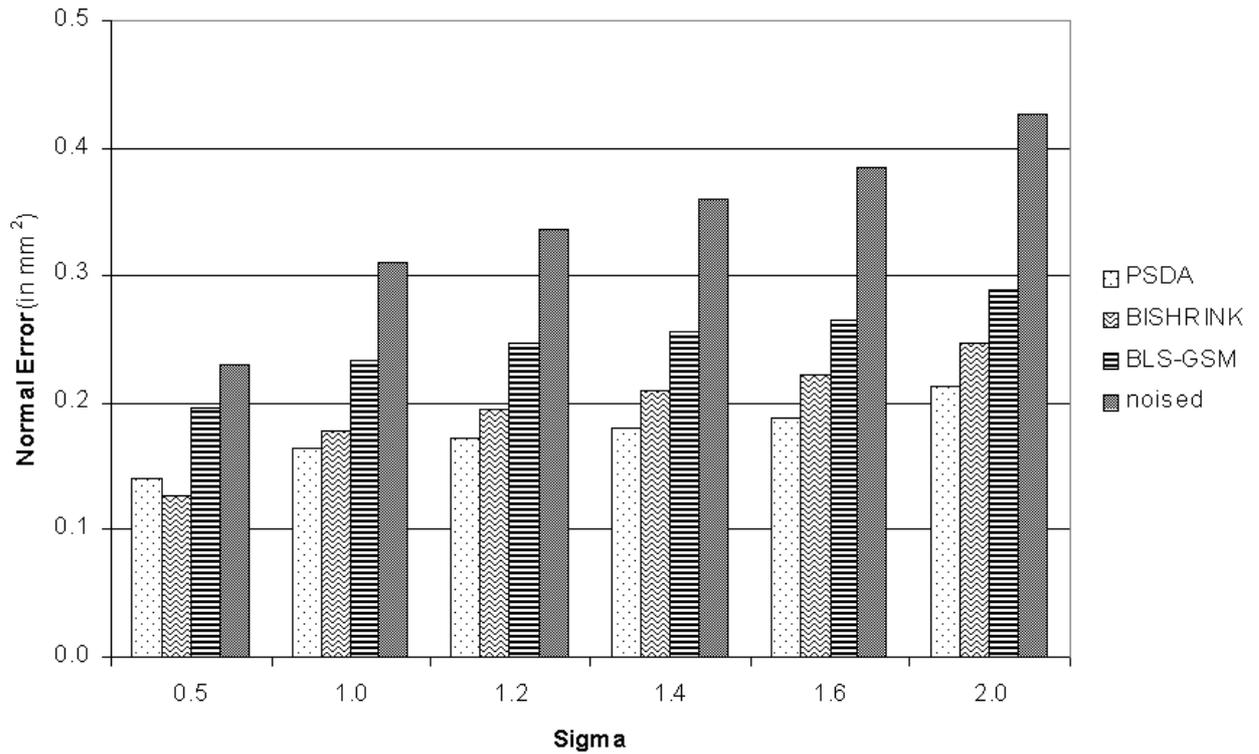


Figure 4. Average (using 24 datasets) normal error (in mm^2) for BISHRINK, BLS-GSM and PSDA.

for most levels of noise, with BISHRINK coming a close second. Our timing results indicate that PSDA requires twice the time needed by BISHRINK, but less than half the time needed by BLS-GSM.

Our results indicate that frame-based methods are well suited to the task of smoothing range images. Although the statistical assumptions about the type of noise are based on intensity image considerations, the results we obtained using PSDA are of high quality. In addition, this behavior is quite robust, despite the non-isometric nature of the parameterization used.

Acknowledgements

This work was supported in part by the following grants: NSF IIS-0431144, NSF IIS-0335578, NSF DMS-0406748, NSF DMS-1041539, NSF CHE-0074311, ARP 003652-0071-2001, and Welch Foundation E-0608. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the above mentioned sponsors.

REFERENCES

1. F. Prieto, T. Redarce, P. Boulanger, and R. Lepage, "Tolerance control with high quality 3D data," in *Proc. Conf. on 3D digital imaging and modeling (3DIM)*, (Quebec, Canada), May 2001.
2. L. Shen, M. Papadakis, I. A. Kakadiaris, I. Konstantinidis, D. Kouri, and D. Hoffman, "Image denoising using a tight frame," *IEEE Trans. on Image Processing*, 2005.
3. P. Saint-Marc, J.-S. Chen, and G. Medioni, "Adaptive Smoothing: A general tool for early vision," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **13**(6), pp. 514-529, 1991.
4. D. Barash, "Fundamental Relationship Between Bilateral Filtering, Adaptive Smoothing, and the Nonlinear Diffusion Equation," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **24**(6), pp. 844-847, 2002.
5. M. Umasuthan and A. Wallace, "Outlier Removal and Discontinuity Preserving Smoothing of Range Data," in *Proc. IEEE Conf. on Vision, Image and Signal Processing*, **143**(3), pp. 191-200, 1996.
6. M. Desbrun, M. Meyer, P. Schröder, and A. Barr, "Anisotropic Feature-Preserving Denoising of Height Fields and Bivariate Data," in *Proc. Graphics Interface (GI00)*, pp. 588-597, (Montréal, Canada), May 2000.
7. P. Boulanger, O. Jokinen, and A. Beraldin, "Intrinsic filtering of range images using a physically based noise model," in *Proc. Vision Interface (VI02)*, (Calgary, Canada), May 2002.
8. D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika* **81**, pp. 425-455, 1994.
9. D. Donoho, "De-noising by soft thresholding," *IEEE Trans. on Information Theory* **41**, pp. 613-627, 1995.
10. D. Donoho and I. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. of the American Statistical Assoc.* **90**(432), pp. 1200-1224, 1995.
11. R. Coifman and D. Donoho, "Translation-invariant de-noising," in *Wavelets and Statistics*, A. Antoniadis and G. Oppenheim, eds., *Lecture Notes in Statistics*, vol. 103, pp. 125-150, Springer-Verlag, Berlin, 1995.
12. E. Simoncelli, "Bayesian denoising of visual images in the wavelet domain," in *Bayesian Inference in Wavelet Based Models*, P. Müller and B. Vidakovic, eds., *Lecture Notes in Statistics*, vol. 141, ch. 18, pp. 291-308, Springer-Verlag, New York, 1999.
13. S. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. on Image Processing* **9**(9), pp. 1522-1531, 2000.
14. X. Li and M. Orchard, "Spatially adaptive denoising under overcomplete expansion," in *Proc. Int'l. Conf. on Image Processing (ICIP)*, **3**, pp. 300-303, (Vancouver, Canada), Sept. 2000.
15. H. Zhang, A. Nosratinia, and R. Wells, "Modeling the autocorrelation of wavelet coefficients for image denoising," in *Proc. Int'l. Conf. on Image Processing (ICIP)*, **3**, pp. 304-307, (Vancouver, Canada), Sept. 2000.
16. E. Simoncelli, W. Freeman, E. Adelson, and D. Heeger, "Shiftable multi-scale transforms," *IEEE Trans. on Information Theory* **38**(2), pp. 587-607, 1992.
17. J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. on Image Processing* **12**(11), pp. 1338-1351, 2003.
18. N. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Appl. Comput. Harmonic Anal.*, pp. 234-253, 2001.

19. L. Sendur and I. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. on Signal Processing* **50**(11), pp. 2744–2756, 2002.
20. L. Sendur and I. Selesnick, "Bivariate shrinkage with local variance estimation," *IEEE Signal Processing Letters* **9**(12), pp. 438–441, 2002.
21. E. Simoncelli and E. Adelson, "Noise removal via Bayesian wavelet coring," in *Proc. Int'l. Conf. on Image Processing (ICIP)*, **I**, pp. 379–382, (Lausanne, Switzerland), Sept. 1996.
22. S. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. on Image Processing* **9**(9), pp. 1532–1546, 2000.
23. M. Mihçak, I. Kozintsev, and K. Ramchandran, "Low-complexity image modeling based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Letters* **6**(2), pp. 300–303, 1999.
24. A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, "A joint inter- and intrascale statistical model for Bayesian wavelet based image denoising," *IEEE Trans. on Image Processing* **11**(5), pp. 545–557, 2002.
25. J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Acoustics, Speech and Signal Processing* **41**(12), pp. 3445–3462, 1993.
26. S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 2nd ed., 1999.
27. M. Crouse, R. Nowak, and R. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. on Signal Processing* **46**(4), pp. 886–902, 1998.
28. J. Romberg, H. Choi, and R. Baraniuk, "Bayesian Tree-Structured Image Modeling using Wavelet Domain Hidden Markov Models," *IEEE Trans. on Image Processing* **10**(7), pp. 1056–1068, 2001.
29. G. Fan and X.-G. Xia, "Image denoising using local contextual hidden Markov model in the wavelet domain," *IEEE Signal Processing Letters* **8**(5), pp. 125–128, 2001.
30. M. Papadakis, G. Gogoshin, I. Kakadiaris, D. Kouri, and D. Hoffman, "Non-separable radial frame multiresolution analysis in multidimensions," *Numerical Functional Analysis and Optimization* **24**, pp. 907–928, 2003.
31. I. Selesnick, "Smooth wavelet frames with zero moments," *Appl. Comput. Harmonic Anal.* **10**, pp. 163–181, 2001.
32. C. Chui, W. He, and J. Stöckler, "Compactly supported tight and sibling frames with maximum vanishing moments," *Appl. Comput. Harmonic Anal.* **13**, pp. 224–262, 2002.
33. I. Daubechies, B. Han, A. Ron, and Z. Shen, "Framelets: MRA-based constructions of wavelet frames," *Appl. Comput. Harmonic Anal.* **14**, pp. 1–46, 2003.
34. Q. Jiang, "Parameterizations of masks for tight affine frames with two symmetric/antisymmetric generators," *Adv. Comput. Math.* **18**, pp. 247–268, 2003.
35. A. Petukhov, "Symmetric framelets," *Constr. Approx.* **19**, pp. 309–328, 2003.
36. B. Han and Q. Mo, "Tight wavelet frames generated by three symmetric b-spline functions with high vanishing moments," *Proc. Amer. Math. Soc.* **132**, pp. 77–86, 2004.
37. A. Ron and Z. Shen, "Affine systems in $L^2(\mathbf{R}^d)$: The analysis of the analysis operator," *J. of Functional Analysis* **148**, pp. 408–447, 1997.
38. J. Benedetto and O. Treiber, "Wavelet frames: Multiresolution analysis and extension principles," in *Wavelet Transforms and Time-Frequency Analysis*, L. Debnath, ed., ch. 1, pp. 3–36, Birkhäuser, Boston, 2001.
39. P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the Face Recognition Grand Challenge," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 947–954, (San Diego, CA), June 2005.
40. Y. Ohtake, A. G. Belyaev, and H.-P. Seidel, "Mesh smoothing by adaptive and anisotropic gaussian filter applied to mesh normals," in *Vision, Modeling, and Visualization (VMV)*, pp. 203–210, (Erlangen, Germany), 2002.