IMAGE VOLUME DENOISING USING A FOURIER-WAVELET BASIS

Roland G. Wilson *Nasir M. Rajpoot*

Department of Computer Science
University of Warwick
Coventry CV4 7AL, United Kingdom

e-mail: rgw,nasir@dcs.warwick.ac.uk

ABSTRACT

A novel approach to the removal of noise from three-dimensional image data is described. The image sequence is represented using a non-adaptive wavelet basis, carefully chosen for its ability to compactly represent locally planar surfaces. This is important in many 3-D applications, including video sequences, where they represent moving luminance edges, and medical volume data, where they represent the interfaces between tissue and bone, for example. The new basis, called a Fourier-Wavelet Basis, is an extension of the complex wavelet bases which have been increasing in popularity in recent years. Because transformation to the new basis is efficient computationally, the overall denoising algorithm is highly efficient. The success of the approach is demonstrated using a noisy video sequence and a noisy human knee MR image volume.

1. INTRODUCTION

The use of multiresolution image representations in the restoration of noisy images dates back to the late 1980's [1], in which a statistical estimator based on a simple quadtree image model was used in smoothing images corrupted by additive Gaussian noise. Although the model was simple, it suffered from blocking artefacts associated with Haar wavelets, due to their non-overlapping support. In recent years, considerable interest has been shown in image restoration techniques based on wavelets [2, 3]. While these are efficient computationally, they avoid the worst aliasing artefacts and are quite robust. At its simplest, wavelet 'denoising' is particularly attractive, requiring computation of the transform and its inverse, combined with a suitable threshold function. A variety of choices of wavelet have been advocated and various 'hard' or 'soft' thresholding schemes have been proposed [2]. Nevertheless, there are clearly cases where lack of selectivity in the Fourier domain limits the scope of such techniques: the effectiveness of the technique relies on the ability of the basis to concentrate signal energy in relatively few coefficients. This is particularly important in denoising 3-D image data, where locally planar structures, such as moving edges or interfaces between volumes, convey most information. For example, in image sequences, moving edges sweep out locally planar shapes. Similarly, the surfaces of objects within medical images should be preserved during filtering. Preservation of such features requires a basis in which they are sparsely represented. Furthermore, translation invariance has been shown to be important in this application, as in many others [3]. An obvious choice might, therefore, seem to be the Fourier basis, since a plane surface in a volume corresponds to a line in the Fourier domain. Of course, this misses the key epithet: *local*; all image data show only local planarity, with curvature a significant feature at larger scales.

For this reason, we have developed a modification of the complex wavelet bases proposed in [4, 5], which show translation invariance, good direction-selectivity and yet can be implemented efficiently. The new basis, which we label a *Fourier-Wavelet Basis* (FWB), has a combination of scaling, locality and directional characteristics which are well matched to the locally planar surfaces of interest in applications. Furthermore, it is easy to compute and is invertible. After briefly presenting the new basis, we shall demonstrate its effectiveness in the denoising of an image sequence, based on a modified, adaptive version of a soft thresholding technique [2]. The computational complexity of our algorithm is essentially equivalent to a windowed Fourier transform (WFT) and it does not require any motion estimation.

In the next section, the new basis is briefly described and its invertibility demonstrated. This is followed by an outline of the denoising algorithm. Experimental results are presented and discussed for two image datasets: a noisy video sequence and an MR image volume. The paper is concluded with some remarks on the limitations of the current technique and suggestions for its improvement.

2. FOURIER-WAVELET BASIS

The discrete Fourier-Wavelet transform is a combination of two well known image transforms: the Laplacian pyramid [6] and the windowed Fourier transform. In some ways, it is similar to the octave band Gabor representation proposed in [7], but avoids some of the more unpleasant numerical properties of the Gabor functions. Although the pyramid is overcomplete, by some 33% in 2-D, this becomes negligible in 3-D (14%) unless overlapping windows are used in the WFT. Thus in 1-D, in the continuous domain, a prototypical FWB vector has the form

$$f_{\xi,\omega,a}(x) = w(\frac{x-\xi}{a}) \exp[-\jmath\frac{\omega(x-\xi)}{a}] \qquad (1)$$

where ξ, ω and a are respectively the location, frequency and scale parameters of the function. The function $w(.)$ is a window function, chosen along with the sampling interval to ensure invertibility of the discrete form of the transform. In 3-D, the basis comprises the set of Cartesian products over ξ, ω at each scale a. That the continuous transform defined by (1) is invertible follows directly from the observation that it is simply the multiresolution Fourier transform (MFT) [4].

The discrete form, however, is significantly different from that described in [4], for in the present case, the transform of an image volume f, in vector form, is given at scale m by

$$\hat{\boldsymbol{f}}_{\boldsymbol{m}} = \boldsymbol{\mathcal{F}_n}(\boldsymbol{I} - \boldsymbol{G_{m,m+1}G_{m+1,m}})\boldsymbol{f_m} \tag{2}$$

where $\hat{}$ denotes the FWT at scale m, $\mathcal{F}_n$ is the WFT operator with block size $n \times n \times n$, I is the identity operator, f_m is the Gaussian pyramid representation of f on level m

$$\boldsymbol{f_m} = \prod_{\mathbf{0}}^{\boldsymbol{m-1}} \boldsymbol{G_{l+1,l}f} \tag{3}$$

and $G_{m,m+1}, G_{m+1,m}$ are the raising and lowering operators associated with transitions between levels in the Gaussian pyramid. Invertibility follows directly from equations (2) and (3):

Theorem 1 *The representation defined by equation (2) is invertible.*

Proof

First we note that the WFT operator $\mathcal{F}_n$ has an inverse, which can be denoted $\mathcal{F}_n^{-1}$. Secondly, we know from Burt and Adelson that the Laplacian pyramid is invertible, since, trivially,

$$\boldsymbol{f_m} = \boldsymbol{f_m} - \boldsymbol{G_{m,m+1}f_{m+1}} + \boldsymbol{G_{m,m+1}f_{m+1}} \tag{4}$$

and the proof is completed by induction on m.

Importantly, although both the pyramid and FT operators are Cartesian separable, the closeness of the Burt and Adelson filter to a Gaussian function gives the pyramid virtually isotropic behaviour, which can be exploited well by the high frequency resolution of a Fourier basis. Although the non-orthogonality of the FWB may be a handicap in some applications, this is not a significant disadvantage in denoising.

3. THE DENOISING ALGORITHM

The denoising algorithm works in three steps: forward FWT, adaptive thresholding, and inverse FWT. The coefficients of three-dimensional forward FWT of the image volume are computed using the algorithm outlined in the previous section. Since the presence of additive Gaussian white noise means that almost all the FWT coefficients are affected by it, soft thresholding would reduce the contribution of noise to the restored image volume. Based on the assumption that the coefficients relatively small in magnitude at each resolution (ie, below a certain threshold) are most probably due to the noise variation, coefficients with magnitude above a certain threshold are kept while the remaining ones are discarded. The inverse FWT, therefore, provides an estimation of the original uncorrupted image volume.

The choice of threshold is crucial to the performance of this type of transform domain denoising. Donoho and Johnstone [2] have shown that an adaptive threshold θ given by

$$\theta = \sigma\sqrt{2\log n} \tag{5}$$

is an asymptotically optimal choice for threshold value when denoising a 1-D noisy signal, where n denotes the number of samples in the signal and σ is standard deviation of the additive Gaussian white noise. Our experiments showed that using an adaptive threshold for coefficients at different resolutions gives better denoising results as compared to using same threshold value for transform coefficients at all resolutions. We use threshold value θ_i for coefficients at level i of the pyramid as given by

$$\theta_i = \mathcal{L}(\sigma)\sqrt{2\log n_i} \tag{6}$$

where n_i denotes the number of pixels at level i of the pyramid and $\mathcal{L}(\sigma)$ is a suitably chosen function of σ. The following expression for $\mathcal{L}(\sigma)$ was empirically chosen for our experiments

$$\mathcal{L}(\sigma) = a\log_{10}\sigma + b \tag{7}$$

where $a, b \in \Re$ and $b = 2a$.

4. EXPERIMENTAL RESULTS

The algorithm described above for denoising 3-D image data was tested on the *Miss America* sequence of size 128^3 and a human knee MR image volume of size $256\times256\times64$. The image data was corrupted with additive Gaussian noise, and adaptive thresholding was applied to transform coefficients of the noisy sequence represented in a 3-level Fourier-wavelet domain using a 16^3 window. Experimental results for the FWB denoising and translation invariant wavelet (TIW) denoising [3] of the noisy *Miss America* sequence and noisy human knee image volume for signal-to-noise-ratio (SNR) of 0dB, 5dB, and 10dB are shown in Figures 1 and 2. The value of $a = 0.46$ for equation (7) was chosen empirically using least squares fitting.

As is evident from these results, the FWB denoising performs quite well when compared to the TIW denoising, in terms of both visual quality and SNR gain. In particular, the SNR gain of 7.6dB over TIW denoising for noisy *Miss America* sequence of 0dB SNR shows the promise that FWB representation offers for efficiently representing planar structures. However, SNR gains for knee image volume are not so high, which can be explained by the fact that most structures prevalent in this volume are mainly solid rather than planar. Two types of artifacts can be observed from these results: blocky artifacts due to the use of a 16^3 window, and *fake textures* which sometimes persist within these windows due to suppressing a significant amount of high frequency details.

The computational complexity of our algorithm is $O(n)$ as compared to $O\left(n\log_2(n)\right)$ in case of the TIW denoising, making our algorithm faster by orders of magnitude. It is to

be noted, however, that the TIW denoising with a soft thresholding was applied to the individual 2-D frames (or slices) as opposed to the FWB denoising which benefits from representation of the noisy image volume in a 3-D Fourier-wavelet domain. This means a higher storage requirement for our method.

Experimental results in terms of the SNR gain for five different SNR values of the noisy *Miss America* sequence and the human knee MR image volume are provided in Tables 1 and 2, respectively. Overall, these results compare favourably with those found in the literature [8], in terms of both visual quality and SNR gain.

Noisy Volume	TIW Denoising	FWB Denoising
0	9.5	17.1
5	12.6	18.9
10	15.3	20.9
15	18.1	23.5
20	20.6	25.3

Table 1. SNR (in dB) values for the *Miss America* video sequence

Noisy Volume	TIW Denoising	FWB Denoising
0	10.2	13.1
5	12.3	14.9
10	14.6	16.7
15	17.0	18.8
20	19.4	21.3

Table 2. SNR (in dB) values for the human knee MR image volume

5. CONCLUSIONS

This paper has introduced a new, computationally efficient wavelet representation particularly suited to the surface structures which are important in many 3-D imaging applications. After a brief review of its properties, the new representation was used in a denoising scheme based on soft thresholding and was shown to give high gains in terms of signal-to-noise-ratio. This was achieved without motion compensation. Although the technique appears promising, it remains to be demonstrated just how effective it might be with a more complex, adaptive thresholding method. The effects of oversampling and using tapered windows for the WFT also remain to be investigated.

6. REFERENCES

[1] S. Clippingdale and R. Wilson. Least squares image restoration based on a multiresolution model. In *Proceedings ICASSP-89*, 1989.

[2] D.L. Donoho and I.M. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 31:425–455, 1994.

[3] R.R. Coifman and D.L. Donoho. Translation-invariant denoising. In A. Antoniadis and G. Oppenheim, editors, *Wavelets and Statistics*, pages 125–150. Springer-Verlag, 1995.

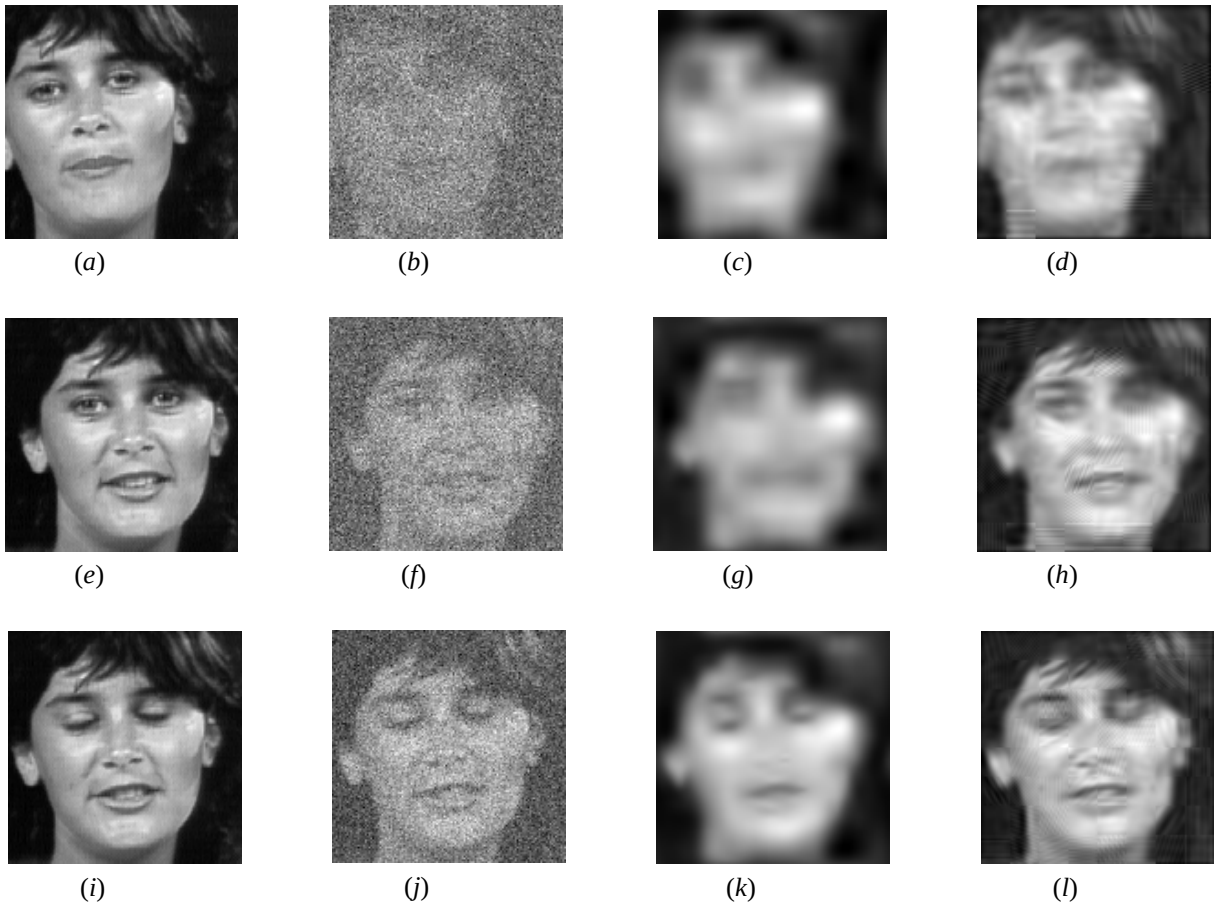
[4] R.G. Wilson, A.D. Calway, and E.R.S. Pearson. A generalized wavelet transform for Fourier analysis: The multiresolution Fourier transform and its applications to image and audio signal analysis. *IEEE Transancation on Information Theory*, 38(2):674–690, March 1992.

[5] N.G. Kingsbury. Image processing with complex wavelets. *Phil. Trans. Royal Soc. A*, 1999.

[6] P.J. Burt and E.H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31:532–540, 1983.

[7] M. Porat and Y.Y. Zeevi. The generalized Gabor scheme of image representation in biological and machine vision. *IEEE Trans. PAMI*, 10:452–468, 1988.

[8] A. Kokaram. *Motion Picture Restoration*. Springer-Verlag, London, 1998.



(*a*) (*b*) (*c*) (*d*)

(*e*) (*f*) (*g*) (*h*)

(*i*) (*j*) (*k*) (*l*)

Fig. 1. Denoising results for the *Miss America* video sequence

Frame# 90: (*a*) Original (*b*) Noisy (SNR=0dB) (*c*) TIW Denoised (SNR=9.5dB) (*d*) FWB Denoised (SNR=17.1dB)

Frame# 70: (*e*) Original (*f*) Noisy (SNR=5dB) (*g*) TIW Denoised (SNR=12.6dB) (*h*) FWB Denoised (SNR=18.9dB)

Frame# 75: (*i*) Original (*j*) Noisy (SNR=10dB) (*k*) TIW Denoised (SNR=15.3dB) (*l*) FWB Denoised (SNR=20.9dB)

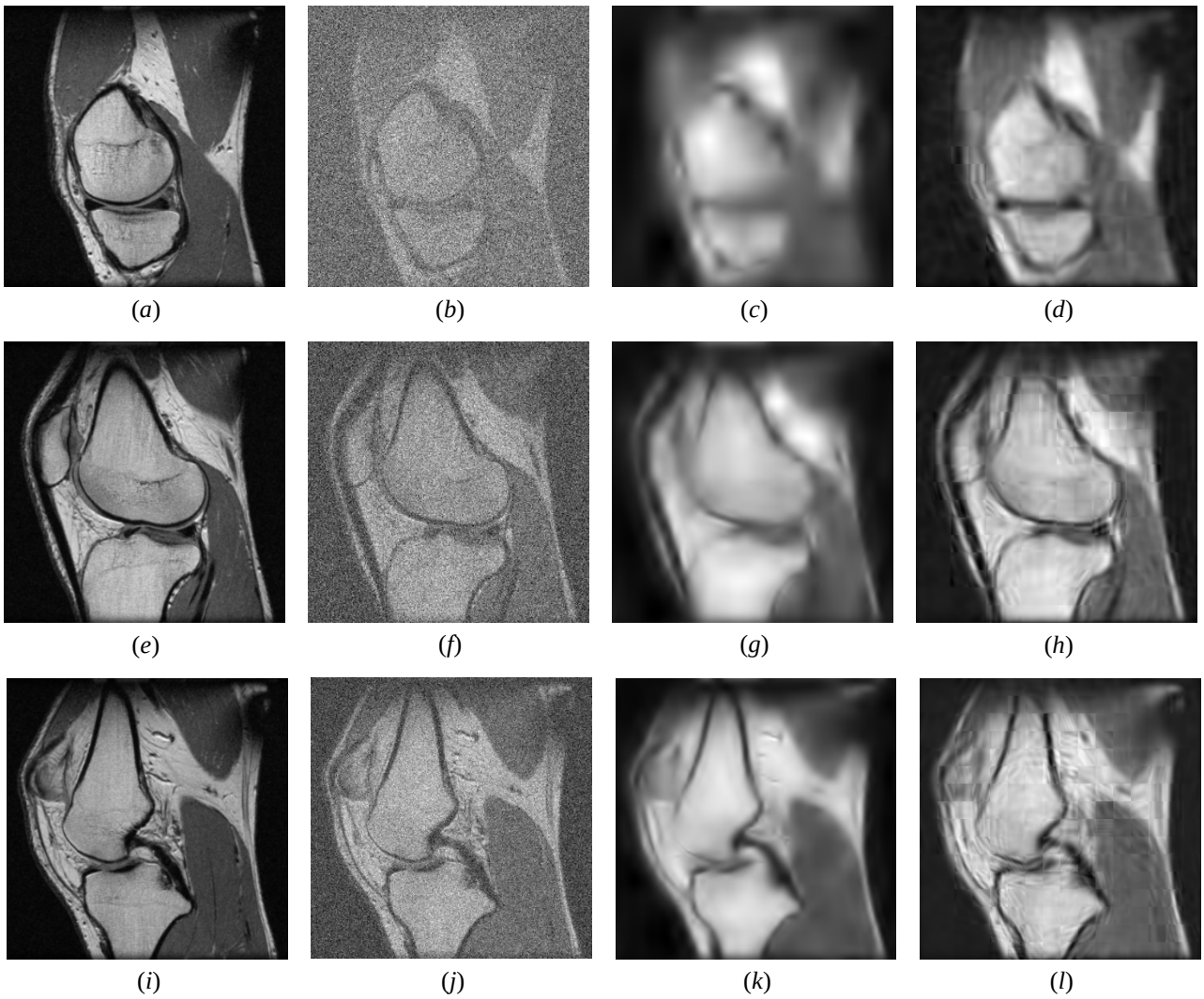


Fig. 2. Denoising results for human knee MR image volume

Slice# 20: (*a*) Original (*b*) Noisy (SNR=0dB) (*c*) TIW Denoised (SNR=9.5dB) (*d*) FWB Denoised (SNR=13.1dB)

Slice# 45: (*e*) Original (*f*) Noisy (SNR=5dB) (*g*) TIW Denoised (SNR=12.3dB) (*h*) FWB Denoised (SNR=14.9dB)

Slice# 30: (*i*) Original (*j*) Noisy (SNR=10dB) (*k*) TIW Denoised (SNR=14.6dB) (*l*) FWB Denoised (SNR=16.7dB)