SURFACE INTERPOLATION USING WAVELETS

Alex P. Pentland

Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract. Extremely efficient surface interpolation can be obtained by use of a wavelet transform. This can be accomplished using biologically-plausible filters, requires only $O(n)$ computer operations, and often only a single iteration is required.

1 Introduction

Surface interpolation is a common problem in both human and computer vision. Perhaps the most well-known interpolation theory is regularization [7, 9]. However this theory has the drawback that the interpolation network requires hundreds or even thousands of iterations to produce a smoothly interpolated surface. Thus in computer vision applications surface interpolation is often the single most expensive processing step. In biological vision, timing data from neurophysiology makes it unlikely that many iterations of cell firing are involved in the interpolation process, so that interpolation theories have been forced to assume some sort of analog processing. Unfortunately, there is little experimental evidence supporting such processing outside of the retina. In this paper I will show how efficient solutions to these problems can be obtained by using orthogonal wavelet filters or receptive fields.

1.1 Background

In computer vision the surface interpolation problem typically involves constructing a smooth surface, sometimes allowing a small number of discontinuities, given a sparse set of noisy range or orientation measurements. Mathematically, the problem may be defined as finding a function $\mathcal{U}$ within a linear space $\mathcal{H}$ that minimizes an energy functional $\mathcal{E}(\mathcal{U})$,

$$\mathcal{E}(\mathcal{U}) = \inf_{\mathcal{V}\in\mathcal{H}} \mathcal{E}(\mathcal{V}) = \inf_{\mathcal{V}\in\mathcal{H}} (\mathcal{K}(\mathcal{V}) + \mathcal{R}(\mathcal{V})) \tag{1}$$

where $\mathcal{K}(\mathcal{V})$ is an energy functional that is typically proportional to the curvature of the surface, and $\mathcal{R}(\mathcal{V})$ is an energy functional that is proportional to the residual difference between $\mathcal{V}$ and the sensor measurements. When the solution exists, the variational derivative $\delta_{\mathcal{U}}$ of the energy functional vanishes,

$$\delta_{\mathcal{U}}\mathcal{E}(\mathcal{U}) = \delta_{\mathcal{U}}\mathcal{K}(\mathcal{U}) + \delta_{\mathcal{U}}\mathcal{R}(\mathcal{U}) = 0 \tag{2}$$

The linear operators $\delta_{\mathcal{U}}\mathcal{E}$, $\delta_{\mathcal{U}}\mathcal{K}$, and $\delta_{\mathcal{U}}\mathcal{R}$ are infinite dimensional and normally dense. To solve Equation 2, therefore, it must first be projected onto a discretization S of $\Re$ containing n nodes. The resulting matrix equation is written $\lambda\mathbf{KU} + \mathbf{R} = 0$ where λ is a scalar constant, $\mathbf{U}$, $\mathbf{R}$ are n x 1 vectors and $\mathbf{K}$ an n x n matrix; these are the discretization of $\mathcal{U}$, $\delta_{\mathcal{U}}\mathcal{R}(\mathcal{U})$, and $\delta_{\mathcal{U}}\mathcal{K}(\mathcal{U})$, respectively. To make explicit the dependence of $\mathbf{R}$ on $\mathbf{U}$, I will write the regularization equation as follows:

$$\lambda\mathbf{KU} + \mathbf{SU} - \mathbf{D} = 0 \tag{3}$$

i.e., $\mathbf{R} = \mathbf{SU} - \mathbf{D}$, where $\mathbf{D}$ is a n x 1 vector whose entries are the measured coordinates d_i where sensor measurements exist and zero elsewhere, and $\mathbf{S}$ is a diagonal "selection matrix" with ones for nodes with sensor measurements and zeros elsewhere.

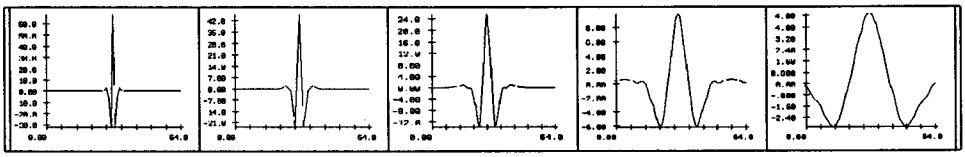

Fig. 1. Wavelet filter family "closest" to Wilson-Gelb filters (arbitrarily scaled for display).

1.2 Choice of Basis

When $\mathcal{K}$ is chosen to be the stress within a bending thin plate (as is standard), then $\mathbf{K}$ is the *stiffness matrix* familiar from physical simulation. Unfortunately, several thousand iterations are often required to the interpolated surface. Although sophisticated multiresolution techniques can improve performance, the best reported algorithms still require several hundred iterations.

The cost of surface interpolation is proportional to both the bandwidth and condition number of $\mathbf{K}$. Both of these quantities can be greatly reduced by choosing the correct *basis* (a set of n orthogonal vectors) and associated coordinate system in which to solve the problem. In neural systems, transformation to a new basis or coordinate system can be accomplished by passing a data vector through a set of receptive fields; the shapes of the receptive fields are the new basis vectors, and the resulting neural activities are the coordinates of the data vector in the coordinate system defined by these basis vectors. If the receptive fields are orthonormal, then we can convert back to the original coordinate system by adding up the same receptive fields in amounts proportional to the associated neurons activity.

For the class of physically-motivated smoothness functionals, the ideal basis would be both spatially and spectrally localized, and (important for computer applications) very fast to compute. The desire for spectral localization stems from the fact that, in the absence of boundary conditions, discontinuities, etc., these sort of physical equilibrium problems can usually be solved in closed form in the frequency domain. In similar fashion, a spectrally-localized basis will tend to produce a banded stiffness matrix $\mathbf{K}$. The requirement for spatial localization stems from the need to account for local variations in $\mathbf{K}$'s band structure due to, for instance, boundary conditions, discontinuities, or other inhomogeneities.

1.3 Orthogonal Wavelet Bases

A class of bases that provide the desired properties are generated by functions known as *orthogonal wavelets* [5, 2, 8]. Orthogonal wavelet functions and receptive fields are different from the wavelets previously used in biological and computational modeling because *all* of the functions or receptive fields within a family, rather than only the functions or receptive fields of one size, are orthogonal to one another. A family of

orthogonal wavelets $h_{a,b}$ is constructed from a single function h by dilation of a and translation of b

$$h_{a,b} = |a|^{-1/2} h\left(\frac{x-b}{a}\right), \qquad a \neq 0 \tag{4}$$

Typically $a = 2^j$ and $b = 1, ..., n = 2^j$ for $j = 1, 2, 3....$ The critical properties of wavelet families that make them well suited to this application are that:

- For appropriate choice of h they can provide an orthonormal basis of $\mathbf{L}^2(\Re)$, i.e., *all* members of the family are orthogonal to one another.
- They can be simultaneously localized in both space and frequency.
- Digital transformations using wavelet bases can be *recursively* computed, and so require only $O(n)$ operations.

Such families of wavelets may be used to define a set of multiscale orthonormal basis vectors. I will call such a basis Φ_w, where the columns of the n x n matrix Φ_w are the basis vectors. Because Φ_w forms an orthonormal basis, $\Phi_w^T \Phi_w = \Phi_w \Phi_w^T = \mathbf{I}$. That is, like the Fourier transform, the wavelet transform is self-inverting. Figure 1 shows a subset of Φ_w; from left to right are the basis vectors corresponding to $a = 1, 2, 4, 8, 16$ and $b = n/2$. All of the examples presented in this paper will all be based on the wavelet basis illustrated in this figure.

The basis vector shapes shown in Figure 1 may be regarded as the neural receptive fields that transform an input signal into, or out of, the wavelet coordinate system. I developed this particular set of wavelets to match as closely as possible the human psychophysical receptive field model of Wilson and Gelb [10]; there is only a 7.5% MSE difference between this set of wavelet receptive fields and the Wilson-Gelb model [1] [6]. *This set of wavelets, therefore, provides a good model of human spatial frequency sensitivity, and of human sensitivity to changes in spatial frequency.*

2 Surface Interpolation using Wavelet Bases

It has been proven that by using wavelet bases linear operators such as $\delta_U \mathcal{K}$ can be represented extremely compactly [1]. This suggests that Φ_w is an effective preconditioning transform, and thus may be used to obtain very fast approximate solutions. The simplest method is to transform a previously-defined $\mathbf{K}$ to the wavelet basis,

$$\tilde{\mathbf{K}} = \Phi_w^T \mathbf{K} \Phi_w \tag{5}$$

then to discard off-diagonal elements,

$$\Omega_w^2 = \text{diag}\left(\Phi_w^T \mathbf{K} \Phi_w\right) \tag{6}$$

and then to solve. Note that for each choice of $\mathbf{K}$ the diagonal matrix Ω_w^2 is calculated only once and then stored; further, its calculation requires only $O(n)$ operations. In numerical experiments I have found that for a typical $\mathbf{K}$ the summed magnitude of the off-diagonals of $\tilde{\mathbf{K}}$ is approximately 5% of the diagonal's magnitude, so that we expect to incur only small errors by discarding off-diagonals.

[1] This set of wavelets were developed by applying the gradient-descent QMF design procedure of Simoncelli and Adelson [8] using the Wilson-Gelb filters as the initial "guess" at an orthogonal basis. Wavelet receptive fields from only five octaves are shown, although the Wilson-Gelb model has six channels. Wilson, in a personal communication, has advised us that the Wilson-Gelb "b" and "c" channels are sufficiently similar that it is reasonable to group them into a single channel.

Case 1. The simplest case of surface interpolation is when sensor measurements exist for every node so that the sampling matrix $\mathbf{S} = \mathbf{I}$. Substituting $\Phi_w \tilde{\mathbf{U}} = \mathbf{U}$ and premultiplying by Φ_w^T converts Equation 3 to

$$\lambda \Phi_w^T \mathbf{K} \Phi_w \tilde{\mathbf{U}} + \Phi_w^T \Phi_w \tilde{\mathbf{U}} = \Phi_w^T \mathbf{D} \tag{7}$$

By employing Equation 6, we then obtain $(\lambda \Omega_w^2 + \mathbf{I})\tilde{\mathbf{U}} = \Phi_w^T \mathbf{D}$, so that the approximate interpolation solution $\mathbf{U}$ is

$$\mathbf{U} = \Phi_w (\lambda \Omega_w^2 + \mathbf{I})^{-1} \Phi_w^T \mathbf{D} \tag{8}$$

Note that this computation is accomplished by simply transforming $\mathbf{D}$ to the wavelet basis, scaling the convolution filters (receptive fields) appropriately at each level of recursion, and then transforming back to the original coordinate system. To obtain an approximate regularized solution for an $\sqrt{n}$ x $\sqrt{n}$ image using a wavelet of width w therefore requires approximately $8wn + n$ add and multiply operations.

Case 2. In the more usual case where not all nodes have sensor measurements, the interpolation solution may require iteration. In this case the sampling matrix $\mathbf{S}$ is diagonal with ones for nodes that have sensor measurements, and zeros elsewhere. Again substituting $\Phi_w \tilde{\mathbf{U}} = \mathbf{U}$ and premultiplying by Φ_w^T converts Equation 3 to

$$\lambda \Phi_w^T \mathbf{K} \Phi_w \tilde{\mathbf{U}} + \Phi_w^T \mathbf{S} \Phi_w \tilde{\mathbf{U}} = \Phi_w^T \mathbf{D} \tag{9}$$

The matrix $\Phi_w^T \mathbf{S} \Phi_w$ is diagonally dominant so that the interpolation solution $\mathbf{U}$ may be obtained by iterating

$$\mathbf{U}^{t+1} = \Phi_w (\lambda \Omega_w^2 + \tilde{\mathbf{S}})^{-1} \Phi_w^T \mathbf{D}^t + \mathbf{U}^t \tag{10}$$

where $\tilde{\mathbf{S}} = \mathrm{diag}(\Phi_w^T \mathbf{S} \Phi_w)$ and $\mathbf{D}^t = \mathbf{D} - (\mathbf{K} + \mathbf{S})\mathbf{U}^t$ is the residual at iteration t. I have found that normally no more than three to five iterations of Equation 10 are required to obtain an accurate estimate of the interpolated surface; often a single iteration will suffice.

Note that for this procedure to be successful, the largest gaps in the data sampling must be significantly smaller than the largest filters in the wavelet transform. Further, when λ is small and the data sampling is sparse and irregular, it can happen that the off-diagonal terms of $\Phi_w^T \mathbf{S} \Phi_w$ introduce significant error. When using small λ I have found that it is best to perform one initial iteration with a large λ, and then reduce λ to the desired value in further iterations.

Discontinuities. The matrix $\mathbf{K}$ describes the connectivity between adjacent points on a continuous surface; thus whenever a discontinuity occurs $\mathbf{K}$ must be altered. Following Terzopoulos [9], we can accomplish this by disabling receptive fields that cross discontinuities. In a computer implementation, the simplest method is to locally halt the recursive construction the wavelet transform whenever one of the resulting bases would cross a discontinuity.

An Example. Figure 2(a) shows the height measurements input to a 64 x 64 node interpolation problem (zero-valued nodes have no data); the vertical axis is height. These data were generated using a sparse (10%) random sampling of the function $z = 100[\sin(kx) + \sin(ky)]$. Figure 2(b) shows the resulting interpolated surface. In this example Equation 10 converged to within 1% of its true equilibrium state with a single iteration. Execution time was approximately 1 second on a Sun 4/330.

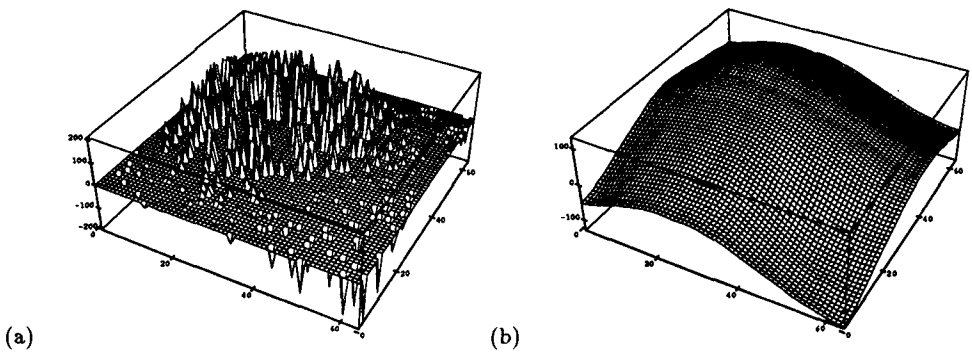

Fig. 2. A surface interpolation problem; solution after one iteration (1 second on a Sun 4/330).

2.1 Summary

I have described a method for surface interpolation that uses orthogonal wavelets to obtain good interpolations with only a very few iterations. The method has a simple biological implementation, and its performance was illustrated with wavelets that accurately model human spatial frequency sensitivity.

References

1. Albert, B., Beylkin, G., Coifman, R., Rokhlin, V. (1990) *Wavelets for the Fast Solution of Second-Kind Integral Equations. Yale Research Report DCS.RR-837*, December 1990.
2. Daubechies, I. (1988) *Orthonormal Bases of Compactly Supported Wavelets. Communications on Pure and Applied Mathematics*, XLI:909-996, 1988.
3. Kohonen, T., (1982) Self-organized formation of topologically correct feature maps, *Biol. Cyber.*, 43, pp. 59-69.
4. Linsker, R. (1986) *From basic network principles to neural architecture, Proc. Nat. Acad. Sci, U.S.A.*, 83, pp. 7508-7512, 8390-8394, 8779-8783.
5. Mallat, S. G., (1989) *A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. PAMI*, 11(7):674–693, 1989
6. Pentland, A., (1991) *Cue integration and surface completion, Invest. Opthal. and Visual Science* 32(4):1197, March 1991.
7. Poggio, T., Torre, V., and Koch, C., (1985) Computational vision and regularization theory, *Nature*, 317:314–319, Sept. 26, 1985.
8. Simoncelli, E., and Adelson, E., (1990) *Non-Separable Extensions of Quadrature Mirror Filters to Multiple Dimensions, Proceedings of the IEEE*, 78(4):652–664, April 1990
9. Terzopoulos, D., (1988) The Computation of visible surface representations, *IEEE Trans. PAMI*, 10(4):417–439, 1988.
10. Wilson, H., and Gelb, G., (1984) *Modified line-element theory for spatial-frequency and width discrimination, J. Opt. Soc. Am. A* 1(1):124-131, Jan. 1984.

This article was processed using the LaTeX macro package with ECCV92 style