

The Morphlet Transform: A Multiscale Representation for Diffeomorphisms

Jonathan R. Kaplan, David L. Donoho
Dept. of Mathematics, Statistics
Stanford University
Stanford, CA, 94305-2125, USA
jkaplan@math.stanford.edu

Abstract

We describe a multiscale representation for diffeomorphisms. Our representation allows synthesis – e.g. generate random diffeomorphisms – and analysis – e.g. identify the scales and locations where the diffeomorphism has behavior that would be unpredictable based on its coarse-scale behavior. Our representation has a forward transform with coefficients that are organized dyadically, in a way that is familiar from wavelet analysis, and an inverse transform that is nonlinear, and generates true diffeomorphisms when the underlying object satisfies a certain *sampling condition*.

Although both the forward and inverse transforms are nonlinear, it is possible to operate on the coefficients in the same way that one operates on wavelet coefficients; they can be shrunk towards zero, quantized, and can be randomized; such procedures are useful for denoising, compressing, and stochastic simulation. Observations include: (a) if a template image with edges is morphed by a complex but known transform, compressing the morphism is far more effective than compressing the morphed image. (b) One can create random morphisms with and desired self-similarity exponents by inverse transforming scaled Gaussian noise. (c) Denoising morphisms in a sense smooths the underlying level sets of the object.

1 Introduction

Temporal or spacial deformation is a common underlying source of variability in many signal and image analysis problems. This deformation may be the result of measurement distortions, as in the case of satellite imagery [1] and GC/MS data [12] or the deformation may be the actual phenomenon of study [9, 11]. In the first case, the deformation is seen as a nuisance and must be removed before further analysis. In the second case, the goal is not the removal of the deformation but rather the extraction of the deformation. Once the deformation has been extracted, understanding the phenomenon at hand consists of analyzing the deformation itself. This paper presents a novel representation for the deformation after extraction that takes advantage of smoothness and multiscale organization to both ease the computational burden of analysis and reveal geometric structure.

For the problem of deformation analysis we will limit ourselves to deformations that are diffeomorphisms—special deformations that are invertible, smooth, and have smooth inverses – this specialization allows us to limit the problem to one that is mathematically amenable. Although they are one of the basic building blocks of modern theoretical mathematics, and an everyday object in pure mathematics, diffeomorphisms are a new data type in image and signal analysis. This new data type needs a representation native to its own structures. This representation should allow for fast computation and storage while maintaining as much transparency as possible. In this paper we present one possible approach to addressing this need.

We present a novel nonlinear invertible multiscale transform on the space of diffeomorphisms that can be used to store, manipulate, and analyze the variability in a collection of signals that are all diffeomorphic to a given template. This multiscale transform is known as the *Morphlet Transform*. The use of such a transform is motivated by the success of multiscale methods in data compression [5], noise removal [4], and fast template registration [7]. Many realistic diffeomorphisms that appear in image and signal analysis are sparsely represented in the Morphlet transform space.

There is one chief obstacle to representing diffeomorphisms: the space of diffeomorphisms is not a linear function space. Diffeomorphisms are functions, and so any given diffeomorphism may be expressed in a basis expansion. But perturbations or manipulations of the expansion coefficients may produce, after reconstruction, a function that is no longer a diffeomorphism. This will be true regardless of the basis functions that are used.

There are two typical strategies for representing diffeomorphisms. The first is to restrict to a parametric family of functions like affine maps, polynomials, or thin-plate splines. Although these methods are attractive due to their simplicity, with the exception of affine maps, none of these methods can guarantee that the matching function is in fact a diffeomorphism. The second strategy is to use a deformation vector field such as in [6]. The diffeomorphism is then the unit time flow induced by the vector field. The vector field representation has three drawbacks: first, calculating features of the full diffeomorphism involves solving a first order PDE, which can be computationally intensive; second, the representation is global, so the values of the vector field outside of some region can effect the value of the diffeomorphism in the region; third, because of the global nature of the vector field the multiscale structure of the field may not correspond to the multiscale structure of the diffeomorphism. In contrast, the Morphlet transform offers a local and computationally efficient description for diffeomorphisms that ensures that the matching functions are, in fact, diffeomorphisms.

The Morphlet transform takes uniform samples of the diffeomorphism on the standard tetrahedral lattice and returns a collection of coefficients indexed by scale and location. These coefficients have the same organizational structure as the coefficients produced by the wavelet transform. Like the wavelet transform, the Morphlet coefficients reflect the deviation of the diffeomorphism at one scale and location from the behavior predicted from the next coarser scale values. Unlike the wavelet transform, a given coefficient no longer corresponds to an element of a basis set. Rather, both decomposition and reconstruction are nonlinear.

1.1 Related Work

There is now a long history of the use of multiscale methods in image registration and deformation [8]. On the question of representing diffeomorphism, D. Mumford and E. Sharon have worked on the question of the conformal fingerprint for conformal diffeomorphisms between two simply connected regions of the plane [11]. T.F. Cootes *et al.* [2] work on building a parametric statistical model on the space of diffeomorphisms. A great deal of interest in pattern analysis using diffeomorphic registration has been generated in the field of computational anatomy. See [6, 9].

2 Template Warping and Diffeomorphisms

Before we continue, let us define some terms and make precise the problem we wish to address. For our purposes, a diffeomorphism is a smooth function from \mathbb{R}^n to \mathbb{R}^n which is one to one, onto, and has a smooth inverse. We shall model signals and images as real-valued functions on \mathbb{R}^n . The deformation of a signal will be the action of a diffeomorphism on the signal.

$$I_{\text{deform}}(x) = \phi^*(I)(x) = I \circ \phi^{-1}(x). \quad (1)$$

We assume we have a collection of signals $\{I_n\}_{n=1}^N$ and a fixed template signal I_{template} such that for each signal I_n there exists a diffeomorphism ϕ_n such that:

$$I_n = \phi_n^*(I_{\text{template}}). \quad (2)$$

We will not address the question of how to calculate ϕ_n given I_n and I_{template} . There is a large existing literature devoted to solving the diffeomorphic registration problem [10], and we postpone discussing the relationship between the morphlet transform and registration algorithms for a later publication. We will simply assume that we can calculate ϕ_n satisfying (2) or some appropriate regularization.¹

Once the set of registering diffeomorphisms, $\{\phi_n\}_{n=1}^N$, has been obtained, all of the information contained in the sample $\{I_n\}_{n=1}^N$ is now contained (up to the accuracy of the diffeomorphic assumption) in the set of registering diffeomorphisms and the template I_{template} . If we want to study the variability of $\{I_n\}_{n=1}^N$ we need only study the variability of $\{\phi_n\}_{n=1}^N$. This is advantageous due to the smoothness of the registering diffeomorphisms.

Typically, the measured signals are not smooth. In 1-d, the signals may have jumps and spikes. In 2-d, the images have edges and textures. Representing these features can be difficult, and most of the science of image analysis focuses on ways of dealing with these non-smooth features. In contrast, the registering diffeomorphisms are frequently very smooth functions with a few localized regions of sharp transition. These regions of sharp transition in the registering diffeomorphisms are exactly the regions responsible for the variability in the collection of signals $\{I_n\}_{n=1}^N$.

Bijections that are smooth except for a few local singularities have sparse Morphlet transforms. Thus, even when analyzing images with high spatial resolution, only a small fraction of the morphlet coefficients of the resulting registering diffeomorphisms will be large. Only these large coefficients are important. Morphlet transform preserves diffeomorphisms, and has approximation properties similar to the wavelet transform. Thus, if

¹In the presence of noise, (2) is never satisfied. Rather, registration algorithms typically search for a diffeomorphism that satisfies a regularized least squares problem.

we threshold and discard small coefficients, the reconstructed diffeomorphism will have a sparse representation and will be very close to the original diffeomorphism. We will give examples of such a reconstruction in section 5.2.

3 The Interpolating Wavelet Transforms

The Morphlet transform is a nonlinear variant of the wavelet transform. We do not have the space to give a thorough introduction to the theory of linear wavelet transforms, but because the Morphlet transform explicitly builds off of the interpolating wavelet transform [3] we will briefly describe its construction and properties. For ease of presentation we will only discuss the one dimensional case; the higher dimensional case is similar.

The linear interpolating wavelet transform is defined on the dyadic samples of a continuous real-valued function. Let f be a continuous function on \mathbb{R} . Fix integers J_0 and J_1 which will serve as the coarsest and finest dyadic sampling scales respectively. Sample f at $\frac{k}{2^j}$ for $k \in \mathbb{Z}$. Define β_k^j as:

$$\beta_k^j = f\left(\frac{k}{2^j}\right) \quad (3)$$

for $J_0 \leq j \leq J_1$.

Fix a positive odd integer D . And define the prediction Pred_k^j as:

$$\text{Pred}_k^j = \pi_k^j\left(\frac{k}{2^j} + \frac{1}{2^{j+1}}\right), \quad (4)$$

where π_k^j is a local interpolating polynomial of order D . Specifically,

$$\pi_k^j \text{ interpolates the values } \left(\frac{k+i}{2^j}, \beta_{k+i}^j\right) \text{ for } i = -\frac{D-1}{2}, \dots, \frac{D+1}{2}. \quad (5)$$

We define the linear interpolating wavelet transform as the set of coefficients:

$$\left\{ \beta_k^{J_0}, \alpha_k^{j,\text{linear}} \right\}_{j,k \in \mathbb{Z}} \quad (6)$$

where

$$\alpha_k^{j,\text{linear}} = \beta_{2k+1}^{j+1} - \text{Pred}_k^j. \quad (7)$$

Thus, $\alpha_k^{j,\text{linear}}$ measures the discrepancy between the true value of the function and the value predicted by using the samples at the next coarser scale. The coefficients $\{\beta_k^{J_0}\}$ are called the coarse scale coefficients and $\{\alpha_k^{j,\text{linear}}\}$ are called the detail coefficients. In regions where f is smooth, the detail coefficients decay exponentially in j . The rate of decay measures the degree of smoothness of f . If a function is smooth everywhere except for a few isolated singularities, then the fine scale coefficients of the wavelet transform will be very small away from the singularities.

To invert the transform we employ a pyramidal scheme starting with the coarsest sample $\beta_k^{J_0}$. For each scale j , we predict the values at the next finer scale $j+1$ using (4) and reconstruct the samples at scale $j+1$ using:

$$\beta_{2k+1}^{j+1} = \alpha_k^{j,\text{linear}} + \text{Pred}_k^j. \quad (8)$$

Both the forward and inverse transforms involve $O(N)$ flops, where N is the number of samples of f .

4 The Morphlet Transform

The Morphlet transform acts on the dyadic samples of a continuous diffeomorphism of \mathbb{R}^n . We will show the 2-d version of the transform, as the simplicity of diffeomorphisms in dimension $n = 1$ makes the 1-d transform insufficiently instructive. The high dimensional versions follow the same pattern as the 2-d transform.

4.1 The Sampling Condition

Due to a sampling condition, the Morphlet transform is only actually defined for a special sub-manifold of the space of diffeomorphisms. The idea behind the sampling condition is to ensure, at each scale, that the reconstructed function “looks like a diffeomorphism.” In particular, we demand that the discrete Jacobians of the samples of the diffeomorphism are all positively oriented affine maps at all scales. For any given diffeomorphism there exists a dyadic scale J such that for all scales finer than J the discrete Jacobians satisfy this condition. Thus, a diffeomorphism needs to be sufficiently finely sampled before the Morphlet transform may be applied.

Let ϕ be a diffeomorphism of the plane. As in the linear case define $\beta_{k,l}^j$:

$$\beta_{k,l}^j = \phi\left(\frac{k}{2^j}, \frac{l}{2^j}\right) \quad (9)$$

To clarify the sampling condition, and ease the notation for the definition of the fine scale coefficients, we define three intermediate sets of samples, $\{\beta_{k,l}^{j,0}\}$, $\{\beta_{k,l}^{j,1}\}$, and $\{\beta_{k,l}^{j,2}\}$.

$$\beta_{k,l}^{j,3} = \beta_{k,l}^j \quad (10)$$

$$\beta_{2k,2l}^{j+1,1} = \beta_{2k,2l}^{j+1,2} = \beta_{k,l}^{j,3} \quad (11)$$

$$\beta_{2k+1,2l+1}^{j+1,1} = \beta_{2k+1,2l+1}^{j+1,2} = \frac{1}{2}(\beta_{k+1,l}^{j,3} + \beta_{k,l+1}^{j,3}) \quad (12)$$

$$\beta_{2k+1,2l}^{j+1,1} = \frac{1}{2}(\beta_{k+1,l}^{j,3} + \beta_{k,l}^{j,3}), \beta_{2k+1,2l}^{j+1,2} = \beta_{2k+1,2l}^{j+1,3} \quad (13)$$

We say the diffeomorphism *satisfies the sampling condition* for the scales $j \in [J_0, J_1]$ if the following condition is satisfied:

Discrete Bijectivity Constraint For all $j \in [J_0, J_1]$, $(k, l) \in \mathbb{Z}^2$, $n = 0, 1, 2$,

$$\delta \in \Delta = \left\{ \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \pm \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \pm \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} \right\}: \quad (14)$$

$$\text{sign}(\det[\beta_{k,l}^{j,n} - \beta_{k+\delta_{1,2},l+\delta_{2,2}}^{j,n}, \beta_{k+\delta_{1,1},l+\delta_{2,1}}^{j,n} - \beta_{k+\delta_{1,2},l+\delta_{2,2}}^{j,n}]) = \text{sign}(\det(\delta)). \quad (15)$$

4.2 Defining the Morphlet Coefficients

To define the fine scale coefficients we begin by fixing a coarsest and a finest dyadic scale J_0 and J_1 for which ϕ satisfies the *Discrete Bijectivity Constraint*. As in the interpolating wavelet transform, we also fix an odd integer D which will serve as the order of an underlying linear interpolating wavelet transform as in section 3. In addition, we fix a sequence of exponentially decaying integers λ_j for $j = J_0, J_0 + 1, \dots, J_1$. The order of the transform

D and the rate of decay of λ_j will determine the relationship between the decay of the Morphlet coefficients and the smoothness of the diffeomorphism.

For $J_0 \leq j \leq J_1$ and Δ as in (14), we first define the boundary penalty terms as:

$$\Lambda_i^j(k, l) = \sum_{\delta \in \Delta} \frac{\beta_{k+\delta_{1,1}, l+\delta_{1,2}}^{j,i} - \beta_{k+\delta_{1,2}, l+\delta_{2,2}}^{j,i}}{\det[\beta_{k,l}^j - \beta_{k+\delta_{1,2}, l+\delta_{2,2}}^{j,i}, \beta_{k+\delta_{1,1}, l+\delta_{1,2}}^{j,i} - \beta_{k+\delta_{1,2}, l+\delta_{2,2}}^{j,i}]}. \quad (16)$$

Then we define the fine scale coefficients

$$\alpha_{2k+1, 2l}^j = \alpha_{2k+1, 2l}^{j, \text{linear}} - \lambda_j \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \Lambda_1^{j+1}(2k+1, 2l) \quad (17)$$

$$\alpha_{2k, 2l+1}^j = \alpha_{2k, 2l+1}^{j, \text{linear}} - \lambda_j \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \Lambda_2^{j+1}(2k, 2l+1) \quad (18)$$

$$\alpha_{2k+1, 2l+1}^j = \alpha_{2k+1, 2l+1}^{j, \text{linear}} - \lambda_j \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \Lambda_3^{j+1}(2k+1, 2l+1) \quad (19)$$

The Morphlet transform for ϕ is then:

$$\mathcal{M}(\phi) = \{\beta_{k,l}^{j_0}, \alpha_{k,l}^j\}. \quad (20)$$

Note that the detail coefficients of the Morphlet transform are perturbed versions of the linear interpolating wavelet coefficients. For a smooth diffeomorphism, the difference between the linear coefficients and Morphlet coefficients at fine scales is $O(\lambda_j)$. Thus, the exponential decay of λ_j indicates that the fine scale coefficient of both transforms are very similar. The biggest difference comes in the coarse and medium scales, where the perturbation can be large relative to λ_j . The perturbation is large when, due to a Gibbs' phenomenon, the local polynomial interpolation of the diffeomorphism has a vanishing Jacobian. Under these circumstances, the perturbation can dominate the coefficient.

4.3 Basic Properties of the Morphlet Transform

- Functions reconstructed with the inverse transform will always be discrete diffeomorphisms. On the set of diffeomorphisms that satisfy the *Discrete Bijectivity Constraint*, the Morphlet transform is invertible. Both transforms require $O(N)$ flops where N is the number of samples, though the inverse transform requires more work as it requires the use of Newton's method or another similar nonlinear solver.
- If ϕ is an affine map then $\mathcal{M}_{\text{detail}}(\phi) = 0$. In particular, the detail coefficients of the identity map vanish.
- If $\tau \in \mathbb{R}^2$ then $\mathcal{M}_{\text{coarse}}(\phi + \tau) = \mathcal{M}_{\text{coarse}}(\phi) + \tau$ and $\mathcal{M}_{\text{detail}}(\phi + \tau) = \mathcal{M}_{\text{detail}}(\phi)$. The detail coefficients are invariant under translation.
- If $A \in SO_n$, $\alpha_{k,l}^j(A\phi) = A\alpha_{k,l}^j(\phi)$. The detail coefficients are covariant under linear orthogonal transformation.
- If $\Omega \subset \text{dom}(\phi)$ is an open domain such that $\phi|_{\Omega}$ and $\phi^{-1}|_{\phi(\Omega)}$ are smooth then all Morphlet coefficients with support² in Ω decay geometrically as a function of the

²The support of a Morphlet coefficient is the collection of all points in the domain of the bijection which appear in the penalty term (16) and the linear coefficient (7) for the corresponding formula (17) - (19).

discrete scale index. The rate of decay is determined by the smoothness of the refinement scheme and the smoothness of ϕ .

- If $\lambda_j = O(2^{-(D+1)j})$, the approximation rate of the Morphlet transform is as good as the approximation rate of the associated wavelet transform.

5 Stylized Applications

5.1 Random Diffeomorphisms

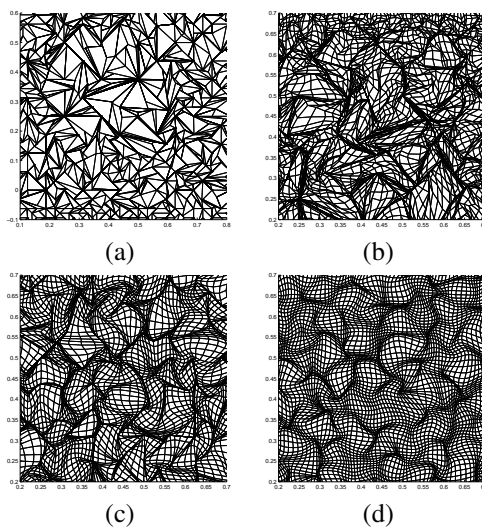


Figure 1: (a) $\omega = 1$, (b) $\omega = 2$, (c) $\omega = 4$, (d) $\omega = 5$.

The inverse Morphlet transform can be used to generate random diffeomorphisms of any preselected smoothness: first, generate a random affine diffeomorphism³; second, subsample the affine map and use the samples as the coarse scale coefficients; third, randomly generate detail coefficients with a preselected decay; finally, apply the inverse transform. In particular, one can choose uncorrelated random Gaussian coefficients with scale-dependent standard deviations $\alpha_{k,l}^j \sim N(0, C2^{-\omega j})$ for some fixed $C, \omega > 0$. The larger ω , the smoother the random diffeomorphism. Figure 1 shows the action of four random diffeomorphisms on a rectangular grid. Each diffeomorphism was generated with coarse scale coefficients set to the identity map and with $C = 1$. The value of ω is varied.

5.2 Compression

In the case where one must compress an image consisting of a known template that has been deformed smoothly, note that if the template has sharp features, such as edges, it can be very difficult to compress a warped instance by standard means. Indeed traditional compression schemes will not work well on figure 2(b). However, as panels (c) and (d) in

³This is just a matter of linear algebra.

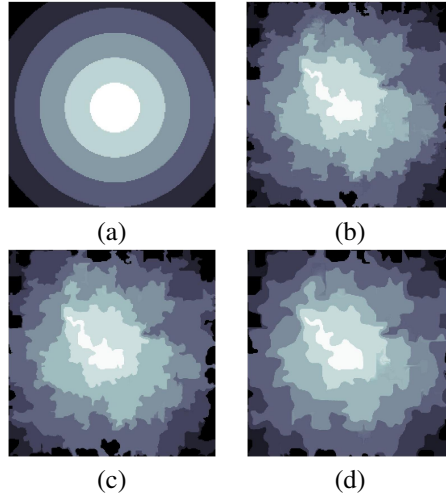


Figure 2: (a) Image I , (b) $\phi^*(I)$, (c) $\phi_{10\%}^*(I)$, Rel. L^2 -error = $9.0e - 4$ (d) $\phi_{2\%}^*(I)$, L^2 -error = $1.9e - 3$.

our example show, if we take the template as separately known to the compressor and the decompressor, and simply compress the morphism, we can reconstruct the warped image yielding a dramatically better visual fidelity than standard compression could offer.

The Morphlet coefficients of a smooth diffeomorphism decay rapidly. Because the approximation rate of the Morphlet transform is at least as good as the approximation rate of the wavelet transform, we may discard a large percentage of the coefficients and still reconstruct the diffeomorphism with high accuracy. This approximation rate can be used to “compress” an image that is a warped version of a known template. To store the image, we store the template and the registering diffeomorphism. If we have a collection of images all of which are diffeomorphic to the template, then the ability to compress the diffeomorphisms translates into low storage for all of the images in the collection.

In figure 2 we show a simple “bulls-eye” template, (a), and the action of a randomly generated diffeomorphism ϕ on the template, (b). Applying the Morphlet transform to ϕ , we threshold all but the largest 10% of the detail coefficients and apply the inverse transform to construct $\phi_{10\%}$. Similarly, we threshold all but the largest 2% of the detail coefficients to construct $\phi_{2\%}$. We apply both $\phi_{10\%}$ and $\phi_{2\%}$ to I and record the relative L^2 -error between $\phi^*(I)$ and $\phi_{10\%}(I)$, $\phi_{2\%}(I)$ respectively. All images are resolved at 512 x 512.

Notice that thresholding the diffeomorphism and applying it to the template effectively smooths the level curves of $\phi(I)$. No linear method can achieve this.

5.3 Interpolation

Given two diffeomorphisms, ϕ_1 and ϕ_0 , that are affine maps at the coarsest scale, the Morphlet transform can be used to interpolate between them in the space of diffeomorphisms. To do so, we interpolate between the two affine maps at the coarsest scale⁴ and then we

⁴Pull back to the Lie algebra and linearly interpolate.

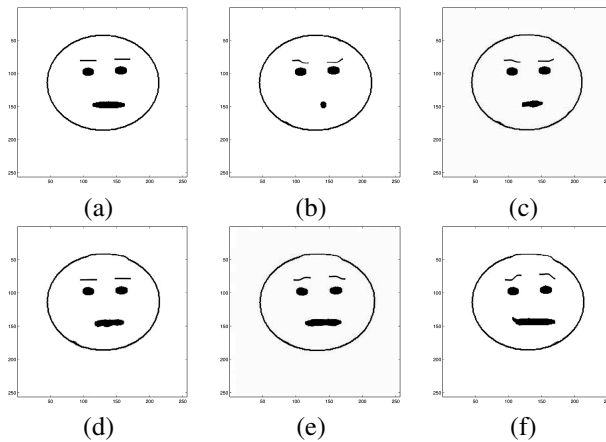


Figure 3: (a) The original template I , (b) $\phi_0^*(I)$, (c) $\phi_{\frac{1}{4}}^*(I)$, (d) $\phi_{\frac{1}{2}}^*(I)$, (e) $\phi_{\frac{3}{4}}^*(I)$, (f) $\phi_1^*(I)$.

simply linearly interpolate the respective detail coefficients of the Morphlet transform:

$$\alpha_{k,l}^j(\phi_t) = t\alpha_{k,l}^j(\phi_1) + (1-t)\alpha_{k,l}^j(\phi_0). \quad (21)$$

Figure 3 shows a simple example of the action of interpolated diffeomorphisms on a template. Figure 3 (a) is the original template, I . Figure 3 (b) and (f) show the action of ϕ_0 and ϕ_1 on the template, respectively, where both diffeomorphisms were synthetically generated by the authors. Both ϕ_0 and ϕ_1 are the identity map at the coarsest scale. For $t = \frac{1}{4}, \frac{1}{2},$ and $\frac{3}{4}$ we interpolate between the two diffeomorphisms as in (21) and apply the maps to the template.

This interpolation provides a cartoon model for articulated motion. The underlying object—here a cartoon face—undergoes diffeomorphic changes. The Morphlet coefficients act like small control knobs with which we can change the expression of the face.

6 Conclusion

The space of Euclidean diffeomorphisms is highly nonlinear. Yet, there is a large submanifold of diffeomorphisms that satisfy the *Discrete Bijectivity Constraint* and for this submanifold the Morphlet transform acts as an embedding into L^∞ (or \mathbb{R}^n in the finitely sampled case). All of the coefficients are calculated using only local information and, similar to wavelets, diffeomorphisms have sparse Morphlet transforms. The Morphlet transform provides a representation where approximation and manipulation are simple arithmetic operations. Future work will further explore how these properties can be used in image and signal analysis.

7 Acknowledgements

Thanks to: Claire Tomlin and other participants in the ONR-MURI “CoMotion” project that partially supported this work; NSF DMS 0072661, NSF DMS 0140698(FRG) and CNS-0085984 for partial support; Peter Schroeder (CalTech), Victoria Stodden, and Inam

Rahman for discussions about software for Multiscale Analysis of Riemannian Symmetric-Space Valued data; IPAM for its support of the first author during the “Multiscale Geometric Analysis” long program; MSRI for its support of both authors during its “Mathematical, Statistical, and Computational Aspects of Vision” long program. The first author has been supported by an NSF Postdoctoral Fellowship.

References

- [1] R. Bernstein. Digital image processing of earth observation sensor data. *IBM Journal of Research and Development*, (20):40–57.
- [2] T. F. Cootes, C. J. Twining, and C. J. Taylor. Diffeomorphic statistical shape models. In *Proceedings of BMVC 2004*, volume 1, pages 447–456, 2004.
- [3] David Donoho. Interpolating wavelet transforms. Technical report, Stanford University, 1992.
- [4] David L. Donoho. De-noising by soft-thresholding. *IEEE Trans. Inform. Theory*, 41(3):613–627, 1995.
- [5] David L. Donoho, Martin Vetterli, R. A. DeVore, and Ingrid Daubechies. Data compression and harmonic analysis. *IEEE Trans. Inform. Theory*, 44(6):2435–2476, 1998. Information theory: 1948–1998.
- [6] S. Joshi and M. I. Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Transactions on Image Processing*, 9(8):1357–1370, August 2000.
- [7] Y. Keller and A. Averbuch. Fast motion estimation using bidirectional gradient methods. *Image Processing, IEEE Transactions on*, 13(8):1042 – 1054, August 2004.
- [8] S. Kovačič and R. Bajcsy. Multiscale/multiresolution representations. In A. W. Toga, editor, *Brain Warping*, pages 45–65. Academic Press, 1999.
- [9] M. I. Miller and U. Grenander. Computational anatomy: An emerging discipline. *Quarterly of Applied Mathematics*, LVI(4):617–694, December 1998.
- [10] Jan Modersitzki. *Numerical methods for image registration*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2004. Oxford Science Publications.
- [11] E. Sharon and D. Mumford. 2d-shape analysis using conformal mapping. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2004*, volume 2, pages 350–357, 2004.
- [12] A. Willse, A. Belcher, G. Preti, J. Wahl, M. Thresher, P. Yang, K. Yamazaki, and G. Beauchamp. Identification of major histocompatibility complex-regulated body odorants by statistical analysis of a comparative gas chromatography/mass spectrometry experiment. *Anal. Chem.*, 77(8):2348 –2361, 2005.