The Multiscale Structure of
Non-Differentiable Image Manifolds

Michael B. Wakin,[r] David L. Donoho,[s] Hyeokho Choi,[r] and Richard G. Baraniuk[r]

[r] Department of Electrical and Computer Engineering, Rice University

[s] Department of Statistics, Stanford University

Email: {wakin, choi, richb}@rice.edu, donoho@stat.stanford.edu

ABSTRACT

In this paper, we study families of images generated by varying a parameter that controls the appearance of the object/scene in each image. Each image is viewed as a point in high-dimensional space; the family of images forms a low-dimensional submanifold that we call an image appearance manifold (IAM). We conduct a detailed study of some representative IAMs generated by translations/rotations of simple objects in the plane and by rotations of objects in 3-D space. Our central, somewhat surprising, finding is that IAMs generated by images with sharp edges are nowhere differentiable. Moreover, IAMs have an inherent multiscale structure in that approximate tangent planes fitted to ϵ-neighborhoods continually twist off into new dimensions as the scale parameter ϵ varies. We explore and explain this phenomenon. An additional, more exotic kind of local non-differentiability happens at some exceptional parameter points where occlusions cause image edges to disappear. These non-differentiabilities help to understand some key phenomena in image processing. They imply that Newton's method will not work in general for image registration, but that a multiscale Newton's method will work. Such a multiscale Newton's method is similar to existing coarse-to-fine differential estimation algorithms for image registration; the manifold perspective offers a well-founded theoretical motivation for the multiscale approach and allows quantitative study of convergence and approximation. The manifold viewpoint is also generalizable to other image understanding problems.

Keywords: Image appearance manifolds, non-differentiable manifolds, angle between subspaces, sampling theorems, multiscale registration, pose estimation.

1. INTRODUCTION

Many problems of interest in image and video processing, pattern recognition, and computer vision involve *families* of related images. If one considers images as points in a high-dimensional ambient space (e.g., million-pixel black-and-white digital images inhabit $\mathbb{R}^{10^6}$), then effective solutions to such image/video processing problems often require exploiting the geometric relationships among the images in the family of interest. In classical signal processing and statistics, for example, the data form subspaces of the ambient space, which leads to linear processing algorithms.

In this paper, we study image families where the images are related by varying a parameter that controls the *appearance* of the objects within each image. Examples include: *translation*, specifying the location of an object in a scene; *orientation*, specifying its pose; or *articulation*, specifying, for a composite object, the relative placement of mobile components. Instead of the more prosaic linear subspaces, such image families form low-dimensional manifolds in the high-dimensional ambient space. We call them *image appearance manifolds* (IAMs).

Much research in high-dimensional data analysis addresses problems of "learning" a manifold from point samples or performing dimensionality reduction for the purposes of classification and parameter extraction [1–3]. The LLE algorithm [2], for example, works by considering locally linear approximations to the manifold, which essentially capture the local geometric structure of the tangent spaces. Such research tends to attack generic data analysis problems, however, and does not address the specific features of IAMs.

In this paper we conduct a detailed study of some representative IAMs corresponding to simple articulations of objects in the image plane and in three-dimensional (3-D) space. Our first, somewhat surprising, finding is that IAMs consisting of images with sharp edges are nowhere differentiable. We consider the approximate tangent spaces formed by fitting hyperplanes to the ϵ-neighborhood of a point in such a manifold and show that, as ϵ varies, the approximate tangent spaces continually "twist off" into new dimensions, never converging to a fixed tangent plane. We believe that this multiscale structure has far-reaching implications for our understanding of current and future image processing tasks.

The family of approximate tangent planes to an IAM, $T(\epsilon, \theta)$, has an index set reminiscent of a continuous wavelet transform, with a scale parameter ϵ and a location parameter θ. We analyze in detail how these multiscale tangent planes twist as we vary the position on the manifold (intra-scale analysis) and as we vary the scale of the manifold (inter-scale analysis). Asking how finely we must sample such a family of planes in scale and location in order to achieve uniform approximation, we find, surprisingly, a behavior similar to classical wavelet analysis of functions, with dyadic scales and integer multiples of the scale being a natural spacing. This has implications for algorithms that need approximate tangent planes (see below).

A more special kind of non-differentiability is caused by *occlusion* phenomena. When an occluding surface exists in a scene, there will generally exist special parameter points at which infinitesimal changes in the parameter can make an edge vanish/appear from behind the occlusion (e.g., a rotating cube in 3-D at the point where a face is appearing/disappearing from view). These articulations correspond to multiscale cusps in the IAM with different "left" and "right" approximate tangent spaces; the local dimensionality of the tangent space changes abruptly at such points.

These non-differentiabilities should force us to rethink certain standard image analysis tasks. Indeed, in many problems in image processing, calculus is a standard tool for deriving algorithms. In the IAM setting, where there are no tangent planes, the approach must be more subtle. In the problem of image registration, for example, one must infer the underlying parameter (location/orientation of an object in the scene) from an unlabeled image. Ordinarily, one might expect to approach this by making an initial guess, projecting on the tangent plane at that point, and reading off the parameter of a linear parametrization of the tangent plane. However, since IAMs are not differentiable and have no tangent planes, image registration really should not be based on standard calculus. This lesson was learned the hard way, without theoretical insights based on the non-differentiable manifold viewpoint, by numerous authors since the early 1980's, who developed "coarse-to-fine differential estimation" methods of image registration [4–9]. The non-differentiable manifold viewpoint explains why multiscale methods are needed in this problem, gives a much more precise way of understanding how certain details need to be handled, and generalizes to a wide range of other image understanding problems.

As a stylized application, we consider the problem of recovering, from a single image, the underlying articulation parameters using this multiscale viewpoint. We develop a multiscale estimation process explicitly using the multiscale approximate tangent planes, and explicitly change the scale of approximation as the accuracy demands increase. This parallels empirically developed methods of image registration, giving them a new quantitative justification and opening the way for problems far beyond simple registration.

Section 2 of this paper elaborates on the manifold viewpoint for articulated image families. Section 3 explores the first type of non-differentiability, caused by the migration of edges. Section 4 analyzes the multiscale tangent twisting behavior in more depth. Section 5 explores the second type of non-differentiability, due to occlusion of edges. Section 6 considers the problem of parameter estimation given an unlabeled image. Section 7 offers a final discussion and conclusions.

2. IMAGE APPEARANCE MANIFOLDS (IAMs)

2.1. The manifold viewpoint

Consider a mathematical model of images as functions $I : \mathbb{R}^2 \mapsto \mathbb{R}$. We will consider images both over the unbounded domain $\mathbb{R}^2$ and over bounded domains, such as the unit square $[0,1] \times [0,1]$. We use $x = (x_0, x_1)$ to denote the coordinates of the image plane.

We are interested in families of images formed by varying a parameter $\theta \in \Theta$ that controls the *articulation* of an object being imaged and thus its *appearance* in each image. For example, θ could be a translation
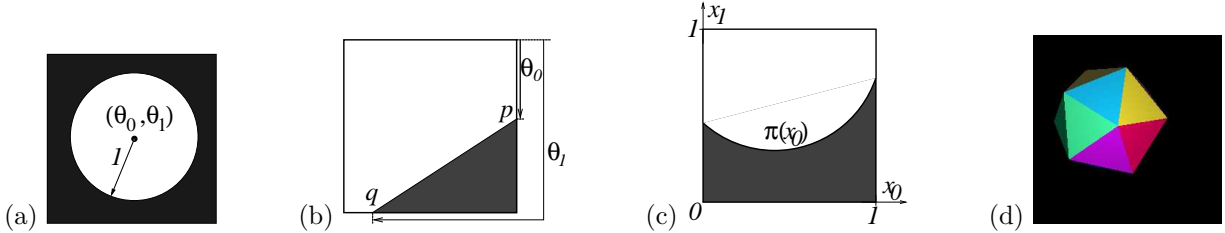
Figure 1. *Simple image articulation models. (a) Parameterization of translating disk image f_θ. (b) Parameterization of a wedgelet. (c) Parameterization of a polynomial horizon. (d) Simulated photograph of a 3-D icosahedron.*

parameter in $\mathbb{R}^3$ specifying the location of the object in a scene; an orientation parameter in SO(3) specifying its pose; or an articulation parameter specifying, for a composite object, the relative placement of mobile components. We let d denote the dimension of θ.

The image formed with parameter θ is a function $f_\theta : \mathbb{R}^2 \mapsto \mathbb{R}$; the corresponding family is the *image appearance manifold* (IAM) $\mathbf{F} = \{f_\theta : \theta \in \Theta\}$. The equation $I = f_\theta$ is our way of saying that the observed image I is a particular member f_θ of the family, with underlying parameter θ. We assume that the relation $\theta \mapsto f_\theta$ is one-to-one.

The set $\mathbf{F}$ is a collection of functions, and we suppose that all of these functions are square-integrable: $\mathbf{F} \subset L^2(\mathbb{R}^2)$. Equipping $\mathbf{F}$ with the L^2 metric, we induce a metric on Θ

$$\mu\left(\theta^{(0)}, \theta^{(1)}\right) = \|f_{\theta^{(0)}} - f_{\theta^{(1)}}\|_{L^2} . \tag{1}$$

Assuming that $\theta \mapsto f_\theta$ is a continuous mapping for the L^2 metric, $M = (\Theta, \mu)$ is a metric space.

2.2. IAM examples

We will use a range of models to illustrate the structural phenomena of IAMs and highlight the basic challenges that can arise in image processing. Similar models are discussed in [10, 11]; the most elaborate such involve combining models to create, for example, articulating cartoon faces.

2.2.1. Articulations in the image plane

The simplest IAMs are formed by articulating cartoon shapes within the image plane. First, consider *translations of an indicator function* in the image plane. Let f_0 be an indicator function in $\mathbb{R}^2$ — a disk, ellipse, square, or rectangle, for example. Let $\Theta = \mathbb{R}^2$ act on the indicator function according to $f_\theta(x) = f_0(x - \theta)$; see Fig. 1(a) for an example with the unit disk. Then it is easy to see that $\mu(\theta^{(0)}, \theta^{(1)}) = m(\|\theta^{(0)} - \theta^{(1)}\|)$ for a monotone increasing function $m \geq 0$, $m(0) = 0$. In fact, if we let B_y denote the indicator function centered at $y \in \mathbb{R}^2$, then

$$m(\rho) = \mathrm{Area}(B_{(0,0)} \triangle B_{(\rho,0)})^{1/2},$$

where $\triangle$ denotes the symmetric difference: $A \triangle B = (A \backslash B) \cup (B \backslash A)$.

In a bounded image domain, a translating indicator function will eventually reach one or both frontiers, where it begins changing shape until it finally disappears completely. We will discuss this *occlusion* phenomenon in more detail in Sec. 5.

Wedgelets offer another bounded domain model. Here we consider the unit square and two points p, q on the boundary of the square. Letting the wedgelet $W_{p,q}$ denote the indicator of the set on the "right hand side" of the line from p to q, we obtain a collection of image elements that are useful for modeling edges in images [12, 13]. To parameterize the wedgelets, let $\theta = (\theta_0, \theta_1)$ denote the position of p and q by measuring the arc length from the northeast corner of the square to each point in a clockwise traverse (see Fig. 1(b)). Thus $\Theta = [0, 4)^2$. Generalizing wedgelets to *polynomial horizons*, let $\pi : [0, 1] \mapsto \mathbb{R}$ be a polynomial of degree $P > 1$, and let $H_\pi(x) = \mathbf{1}_{\{x_1 \leq \pi(x_0),\ x_1 \in [0,1]\}}$ (see Fig. 1(c)). To parameterize these horizons, fix a basis for the set of polynomials, say the standard Legendre basis, $\{L_i\}$ and set $\pi = \sum_{i=0}^{P} \theta_i L_i$. Then $\Theta \subset \mathbb{R}^{P+1}$.

2.2.2. Articulations of 3-D objects

Our model is not limited just to articulations in the image plane. Consider, for example, photography of a 3-D object. In this case, the object may be subject to translations ($\Theta = \mathbb{R}^3$), rotations ($\Theta = \mathrm{SO}(3)$), or a combination of both; the metric on Θ simply involves the difference between two rendered images as in (1). Figure 1(d) shows an example rendering of an icosahedron at an arbitrary position. Additional articulation parameters, such as camera position or lighting conditions [14], can also be considered.

3. NON-DIFFERENTIABILITY FROM EDGE MIGRATION

Each of the image models mentioned in Sec. 2.2 involves sharp edges that move as a function of the parameter θ. This simple effect, relevant in many natural settings where images may feature objects having unknown or moving locations, has a profound consequence on the structure of the resulting IAMs: these manifolds are *nowhere* differentiable. This presents an apparent difficulty for image understanding algorithms that might attempt to exploit the local manifold geometry using calculus.

3.1. The problem

This lack of differentiability can be seen analytically: the metric spaces given in Sec. 2 all have a *non-Lipschitz* relation between the metric distance and the Euclidean distance. As one can check by detailed computations [11], we have

$$\mu\left(\theta^{(0)}, \theta^{(1)}\right) \geq c\left\|\theta^{(0)} - \theta^{(1)}\right\|_2^{1/2} \qquad \text{as } \mu \to 0.$$

The exponent $1/2$ — rather than 1 — implies that the parameterization $\theta \mapsto f_\theta$ is not differentiable. As with a standard function of Hölder regularity $1/2$, we are unable to compute the derivative. For example, to estimate $\frac{\partial f_\theta}{\partial \theta_i}\big|_{\theta=\theta^{(0)}}$, we would let $\theta^{(0)}$ and $\theta^{(1)}$ differ only in component θ_i and would observe that

$$\left\|\frac{f_{\theta^{(1)}} - f_{\theta^{(0)}}}{\theta_i^{(1)} - \theta_i^{(0)}}\right\|_2 \geq c\left\|\theta^{(1)} - \theta^{(0)}\right\|_2^{-1/2} \to \infty \qquad \text{as } \theta^{(1)} \to \theta^{(0)}.$$

This relation is non-differentiable at *every* parameter θ for which local perturbations cause edges to move. Moreover, this failure of differentiability is not something removable by mere reparameterization; no parameterization exists under which there would be a differentiable relationship.

We can also view this geometrically. The metric space $M = (\Theta, \mu)$ is isometric to $\mathbf{F} = (\mathbf{F}, \|\cdot\|_{L^2})$. $\mathbf{F}$ is not a smooth manifold; there simply is no system of charts that can make $\mathbf{F}$ even a C^1 manifold. At base, the lack of differentiability of the manifold $\mathbf{F}$ is due to the lack of *spatial* differentiability of these images [11]. In brief: *images have edges, and if the locations of edges move as the parameters change then the manifold is not smooth.*

3.2. Approximate tangent planes via local PCA

An intrinsic way to think about non-smoothness is to consider approximate tangent planes generated by local principal component analysis (PCA) [15]. Suppose we pick an ϵ-neighborhood $N_\epsilon(\theta^{(0)}; \Theta)$ of some $\theta^{(0)} \in \Theta$; this induces a neighborhood $N_\epsilon(f_{\theta^{(0)}}; \mathbf{F})$ around the point $f_{\theta^{(0)}} \in \mathbf{F}$. We define the ϵ-*tangent plane to* $\mathbf{F}$ *at* $f_{\theta^{(0)}}$ as follows. We place a uniform probability measure on $\theta \in N_\epsilon(\theta^{(0)}; \Theta)$, thereby inducing a measure ν on the neighborhood $N_\epsilon(f_{\theta^{(0)}})$. Viewing this measure as a probability measure on a subset of L^2, we can obtain the first d principal components of that probability measure. These d functions span a d-dimensional affine hyperplane, the approximate tangent plane $T^\epsilon_{f_{\theta^{(0)}}}(\mathbf{F})$; it is an approximate least-squares fit to the manifold over the neighborhood $N_\epsilon(f_{\theta^{(0)}})$.

If the manifold were differentiable, then the approximate tangent planes $T^\epsilon_{f_{\theta^{(0)}}}(\mathbf{F})$ would converge to a fixed d-dimensional space as $\epsilon \to 0$; namely, the plane spanned by the d directional derivatives $\frac{\partial}{\partial \theta_i} f_\theta\big|_{\theta=\theta^{(0)}}$, $i = 0, 1, \ldots, d-1$. However, when these do not exist, the approximate tangent planes do not converge as $\epsilon \to 0$, but continually "twist off" into other dimensions.

Figure 2. *Tangent plane basis vectors of the translating disk IAM estimated: using local PCA at (a) scale $\epsilon = 1/4$ and (b) scale $\epsilon = 1/8$; using image regularization at (c) scale $s = 1/8$ and (d) scale $s = 1/16$.*

As an example, consider the translating disk model, so that the underlying parametrization is 2-D and the tangent planes are 2-D as well. Figure 2(a) shows the approximate tangent plane obtained from this approach at scale $\epsilon = 1/4$. The tangent plane has a basis consisting of two elements, each of which can be considered an image. Figure 2(b) shows the tangent plane basis images at the finer scale $\epsilon = 1/8$. It is visually evident that the tangent plane bases at these two scales are different; in fact the angle between the two subspaces is approximately 30°. Moreover, since the basis elements resemble annuli of shrinking width and growing amplitude, it is apparent for continuous-domain images that as $\epsilon \to 0$, the tangent plane bases cannot converge in L^2.[*]

3.3. Approximate tangent planes via regularization

The lack of IAM differentiability poses an apparent difficulty for image processing: the geometric relationship among images nearby in articulation space seems to be quite complicated. In addition to illuminating this challenge, however, the local PCA experiments in Sec. 3.2 also suggest a way out. Namely, the "twisting off" phenomenon can be understood as the existence of an intrinsic *multiscale structure* to the manifold. Tangent planes, instead of being associated with a location only, as in traditional monoscale analysis, are now associated with a location and a scale.

For a variety of reasons, it is convenient in formalizing this notion to work with a different notion of approximate tangent plane. We first define the *family of regularized manifolds* as follows. Associated with a given IAM, we have a family of *regularization operators* Φ_s that act on functions $f \in \mathbf{F}$ to smooth them; the parameter $s > 0$ is a scale parameter. For example, for the translated disk model, we let Φ_s be the operator of convolution with a Gaussian of standard deviation s: $\Phi_s f = \phi_s * f$, where $\phi_s(x) = \frac{1}{2\pi s^2} \exp\{\frac{-\|x\|^2}{2s^2}\}$. We also define $f_{\theta,s} = \Phi_s f_\theta$. The functions $f_{\theta,s}$ are smooth, and the collection of such functions for θ varying and $s > 0$ makes a manifold $\mathbf{F}_s$. The operator family $(\Phi_s)_{s>0}$ has the property that, as we smooth less, we do less: $\Phi_s f_\theta \to_{L^2} f_\theta$, $s \to 0$. It follows that, at least on compact subsets of $\mathbf{F}$,

$$\mathbf{F}_s \to_{L^2} \mathbf{F}, \quad s \to 0. \tag{2}$$

Because the regularized images contain no sharp edges, it follows that the regularized IAMs are differentiable. We define the *approximate tangent plane* at scale $s > 0$, $T(s, \theta^{(0)}; \mathbf{F})$, to be the exact tangent plane of the approximate manifold $\mathbf{F}_s$; that is $T_{f_{\theta^{(0)},s}}(\mathbf{F}_s)$.

$T(s, \theta^{(0)})$ is the affine span of the functions $\frac{\partial}{\partial \theta_i} f_{\theta,s}\big|_{\theta=\theta^{(0)}}$, $i = 0, 1, \ldots, d-1$. This notion of approximate tangent plane is different from the more intrinsic local PCA approach but is far more amenable to analysis and computation. In practice, the two notions are similar: regularizing an image averages nearby pixel values, whereas local PCA analyzes a set of images related approximately by small shifts in space.

As an example, consider again the translating disk model. Figures 2(c),(d) show the tangent planes obtained from the image regularization process at scales $s = 1/8$ and $s = 1/16$. It is again visually evident that the tangent plane bases at the two scales are different, with behavior analogous to the bases obtained using the local PCA approach in Figures 2(a),(b). In this case, the angle between the two tangent planes is 26.4°.

---

[*]In the case of a *pixelized* image, this phenomenon cannot continue indefinitely. However, the twisting behavior does continue up until the very finest scale, making our analysis relevant for practical algorithms (e.g., see Sec. 6).

3.4. Regularized tangent images

It is instructive to pursue an explicit description for the multiscale tangent images. We begin by deriving the regularized tangents for a restricted class of IAMs, where we have smooth articulations of an indicator set in the plane. This work follows closely certain computations in [11].

Let B denote an indicator set (for example, a disk), and let ∂B denote the boundary of B, which we assume to be C^2. For a point $b \in \partial B$, let $n(b)$ denote the outward-pointing normal vector to ∂B. The set $B = B_\theta$ may change as a function of θ, but we assume the evolution of ∂B_θ to be smooth. Thus we can attach to each boundary point $b \in \partial B_\theta$ a *motion vector* $v_i(b, \theta)$ that indicates the local direction in which the boundary shifts with respect to changes in component θ_i. For example, note that v_i is constant-valued when the articulations simply translate the set B.

From Lemma A.2 in [11], it follows that

$$\frac{\partial}{\partial \theta_i} f_{\theta,s}(x) \bigg|_{\theta = \theta^{(0)}} = \int_{\partial B} \phi_s(x - b) \, \sigma_i(b) \, db,$$

where $\sigma_i(b) := \langle v_i(b, \theta^{(0)}), n(b) \rangle$ measures the amount of shift in the direction normal to the edge. This can be rewritten as the convolution of the regularization kernel ϕ_s with a *Schwartz distribution* $\gamma_i(x)$. This distribution can be understood as a 1-D ridge of delta functions around the boundary ∂B with "height" $\sigma_i(p)$ for $p \in \partial B$ (and height zero elsewhere). Indeed, this distribution also corresponds to the limiting "tangent image" on the unregularized manifold $\mathbf{F}$. We have essentially justified the last link in this chain of equalities

$$\frac{\partial}{\partial \theta_i} f_{\theta,s} = \frac{\partial}{\partial \theta_i} (\phi_s * f_\theta) = \left(\phi_s * \frac{\partial}{\partial \theta_i} f_\theta \right) = \phi_s * \gamma_i. \tag{3}$$

The problem, of course, is that $\gamma_i \notin L^2(\mathbb{R}^2)$, and so we rely on the regularization process. The formula (3) is one on which we may rely for general images f_θ — the regularized tangents can be obtained by convolving a Gaussian with the distributional tangent images.

4. MULTISCALE TWISTING OF IAMs

IAMs of images with sharp edges are non-differentiable, because their tangent planes continually "twist off" into new dimensions. In this section, we examine the multiscale structure of this phenomenon, for the example case of the translating disk IAM. First, we study the twisting phenomenon of the family of smoothed manifolds $\mathbf{F}_s$ as a function of scale s; next we examine twisting at a single scale as a function of position on the manifold. As we will discover, the multiscale characterization of the manifold is not unlike the wavelet analysis of non-differentiable functions.

4.1. Tangent bases for translating disk IAM

We can provide some quantitative values for regularized tangent images in the case of a translated disk. For technical reasons we let the image be the full plane $\mathbb{R}^2$ and also let $\Theta = \mathbb{R}^2$.

We start by identifying the boundary ∂B with the circle $[0, 2\pi)$ (we let $b = 0$ denote the rightmost point of B and traverse ∂B in the counterclockwise direction). For clarity, we write $\vec{b}$ when referring to the boundary point in $\mathbb{R}^2$ and write b when referring to the corresponding angle. For example, we have that $n(\vec{b}) = [\cos(b), \sin(b)]^T$. For translations we have simply that $v_0(\vec{b}) = [1, 0]^T$ and $v_1(\vec{b}) = [0, 1]^T$. This gives $\sigma_0(\vec{b}) = \cos(b)$ and $\sigma_1(\vec{b}) = \sin(b)$.

In order to examine the inter-scale twisting of the tangent planes, we use as a basis for the approximate tangent space $T(s, \theta^{(0)})$ the functions

$$\tau_s^i = \frac{\partial}{\partial \theta_i} f_{\theta,s} \bigg|_{\theta = \theta^{(0)}}.$$

The $L^2(\mathbb{R}^2)$ inner product between these tangent images is given by

$$
\begin{aligned}
\langle \tau_s^i, \tau_s^j \rangle &= \langle \phi_s * \gamma_i, \phi_s * \gamma_j \rangle \\
&= \int_{\mathbb{R}^2} \int_{\partial B} \phi_s(x - \vec{b})\, \sigma_i(\vec{b})\, d\vec{b} \int_{\partial B} \phi_s(x - \vec{\beta})\, \sigma_j(\vec{\beta})\, d\vec{\beta}\, dx \\
&= \int_{\partial B} \int_{\partial B} \sigma_i(\vec{b})\, \sigma_j(\vec{\beta}) \int_{\mathbb{R}^2} \phi_s(x - \vec{b})\, \phi_s(x - \vec{\beta})\, dx\, d\vec{\beta}\, d\vec{b} \\
&= \int_{\partial B} \int_{\partial B} \sigma_i(\vec{b})\, \sigma_j(\vec{\beta})\, \phi_{\sqrt{2}s}(\vec{b} - \vec{\beta})\, d\vec{\beta}\, d\vec{b}.
\end{aligned}
$$

The last step follows because the convolution of two Gaussians yields another Gaussian; a similar derivation appears in Lemma A.3 of [11]. Considering the case where $i \neq j$, we have

$$
\begin{aligned}
\langle \tau_s^0, \tau_s^1 \rangle &= \int_0^{2\pi} \int_0^{2\pi} \cos(b) \sin(\beta)\, \phi_{\sqrt{2}s}(\vec{b} - \vec{\beta})\, d\beta\, db \\
&= \int_{-\pi/2}^{2\pi - \pi/2} \int_{-\pi/2}^{2\pi - \pi/2} \cos(b + \pi/2) \sin(\beta + \pi/2)\, \phi_{\sqrt{2}s}(\vec{b} - \vec{\beta})\, d\beta\, db \\
&= -\int_0^{2\pi} \int_0^{2\pi} \sin(b) \cos(\beta)\, \phi_{\sqrt{2}s}(\vec{b} - \vec{\beta})\, d\beta\, db \\
&= -\langle \tau_s^1, \tau_s^0 \rangle,
\end{aligned}
$$

which implies that $\langle \tau_s^0, \tau_s^1 \rangle = 0$. Thus we have that $\langle \tau_s^i, \tau_s^j \rangle = c_{s,s} \delta_{i,j}$, where, for generality useful below, we set

$$
c_{s_0, s_1} := \int_{\partial B} \int_{\partial B} \cos(b) \cos(\beta)\, \phi_{\sqrt{s_0^2 + s_1^2}}(\vec{b} - \vec{\beta})\, d\vec{\beta}\, d\vec{b}.
$$

Hence, *the $\{\tau_s^i\}$ form an orthogonal basis for the approximate tangent plane $T(s, \theta^{(0)})$ for every $s > 0$.*

Consider now the bases $\{\tau_{s_0}^i\}_{i=0}^1$, $\{\tau_{s_1}^i\}_{i=0}^1$ at two different scales s_0 and s_1. Then by a similar calculation

$$
\langle \tau_{s_0}^i, \tau_{s_1}^j \rangle = c_{s_0, s_1} \delta_{i,j}. \tag{4}
$$

Hence, a basis element at one scale correlates with only one basis element at another scale.

4.2. Inter-scale twist angle

We can give (4) a geometric interpretation based on angles between subspaces. At each scale, define the new basis

$$
\psi_s^i = c_{s,s}^{-1/2}\, \tau_s^i, \quad i = 0, 1,
$$

which is an orthonormal basis for the approximate tangent space $T(s, \theta^{(0)})$. These bases are canonical for measuring the angles between any two tangent spaces. Formally, if we let P_s denote the linear orthogonal projection operator from $L^2(\mathbb{R}^2)$ onto $T(s, \theta^{(0)})$, then the subspace correlation operator $\Gamma_{s_0, s_1} = P_{s_0} P_{s_1}$ has a singular value decomposition using the two bases as left and right singular systems, respectively:

$$
\Gamma_{s_0, s_1} = \sum_{i=0}^1 \psi_{s_0}^i \lambda^i \langle \psi_{s_1}^i, \cdot \rangle;
$$

or, in an informal but obvious notation,

$$
\Gamma_{s_0, s_1} = [\psi_{s_0}^0;\ \psi_{s_0}^1]\, \mathrm{diag}(\lambda^0, \lambda^1)\, [\psi_{s_1}^0;\ \psi_{s_1}^1]^T.
$$

The diagonal entries are given by

$$
\lambda_{s_0, s_1}^i = \frac{c_{s_0, s_1}}{c_{s_1, s_1}^{1/2} c_{s_0, s_0}^{1/2}}.
$$

Now from the theory of angles between subspaces [16,17], we have that the angles between the subspaces $T(s_0, \theta^{(0)})$ and $T(s_1, \theta^{(0)})$ are naturally expressed as $\cos(\text{angle } \#i) = \lambda^i_{s_0,s_1}$, $i = 0, 1$. In this instance, $\lambda^0 = \lambda^1$, and so we write simply

$$\cos(\text{angle}\{T(s_0, \theta^{(0)}), T(s_1, \theta^{(0)})\}) = \frac{c_{s_0,s_1}}{c^{1/2}_{s_0,s_0} c^{1/2}_{s_1,s_1}}.$$

We can perform a simple asymptotic analysis of the c_{s_0,s_1}.

THEOREM 4.1. *In the translating disk model, let the regularization kernel ϕ_s be a Gaussian with standard deviation $s > 0$. Fix $0 < \alpha < 1$ and let $s_1 = \alpha s_0$. Then*

$$\lim_{s_0 \to 0} \cos(\text{angle}\{T(s_0, \theta^{(0)}), T(s_1, \theta^{(0)})\}) = \sqrt{\frac{2\alpha}{\alpha^2 + 1}}. \tag{5}$$

Proof. See Appendix.

This analytical result is fully in line with the results found in Secs. 3.2 and 3.3 by empirically calculating angles between subspaces (for the case $\alpha = 1/2$, the formula predicts an angle of $26.6°$).

4.3. Intra-scale twist angle

We can also examine the twisting phenomenon of the smoothed IAM $\mathbf{F}_s$ at a single scale s as a function of position on the manifold.

A simple experiment reveals the basic effect. We choose $\delta > 0$ and set $\Delta = [\delta; \ 0]^T$. We then compute

$$\text{angle } \{T(s, \theta), T(s, \theta + s\Delta)\}$$

at a variety of scales s. Figure 3 shows the experimental results for 256×256 images; tangents are estimated using a local difference between two synthesized images. This experiment reveals the following effects. First, and not surprisingly, larger changes in θ cause a larger twist in the tangent spaces. Second, and more surprisingly, the twist angle is approximately constant across scale when the change in θ is proportional to the scale. This behavior can also be confirmed analytically following the techniques of Sec. 4.1, though the analysis is a bit more complicated.

This experiment pertains to images over the unbounded domain $\mathbb{R}^2$. In case of a bounded domain, the disk will ultimately experience occlusion at the boundary of the image. In this region of occlusion, we have found that the twisting of the manifold $\mathbf{F}_s$ will depend not only on δ, but also more strongly on s and θ, unlike the experiment above.

4.4. Sampling

Through the process of regularization, we have defined a continuous multiscale characterization of an IAM tangent space. It is interesting, however, to consider the problem of sampling the multiscale tangent space while still preserving its essential structure. For example, we may be interested in answering the following question: "How finely must we sample in scale s at a fixed $\theta^{(0)}$ so that, between adjacent scales, the manifold twists no more than ρ degrees?" Similarly, "How finely must we sample in θ at a fixed scale s so that, between adjacent samples, the manifold twists no more than ρ degrees?" (For example, the success of our parameter estimation algorithm in Sec. 6 will depend on similar questions.)

From Theorem 4.1, it follows that by choosing a sequence

$$s_i = \alpha^i s_0, \quad i = 1, 2, \ldots$$

with an appropriate $\alpha < 1$, we can ensure that the tangent planes at adjacent scales change by no more than a fixed angle. Within a fixed scale, as we have seen in Sec. 4.3, to obtain a constant angle of twist, the amount of shift should be proportional to the smoothing scale s_i. These "sampling rules" for the multiscale tangent space are reminiscent of the sampling of the continuous wavelet transform to obtain the discrete wavelet transform (a case where $\alpha = 1/2$). Just as a non-differentiable function can be characterized with a multiresolution analysis, the translated disk IAM can be characterized by a multiresolution analysis having a similar scale-space structure. This basic behavior is common among a range of IAM models, though the precise details will vary. For use in an algorithm, additional analytic or experimental investigation may be necessary.
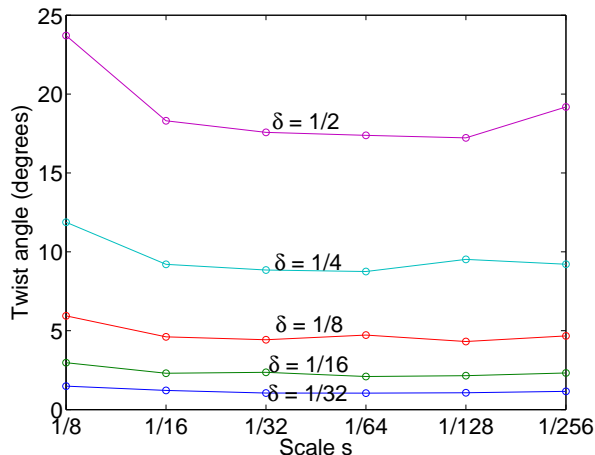
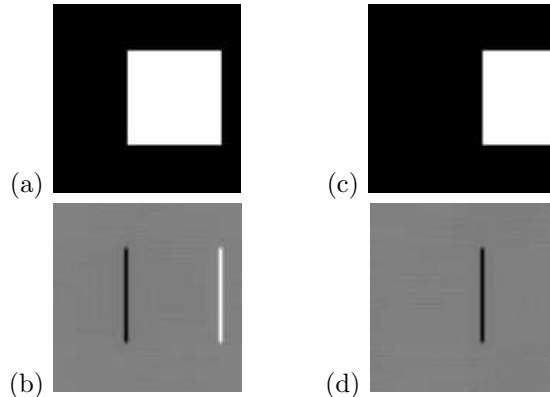Figure 3. *Intra-scale twist angles for translating disk.*

Figure 4. *Changing tangent images for translating square before and after occlusion. Pre-occlusion: (a) image and (b) tangent. Post-occlusion: (c) image and (d) tangent.*

5. NON-DIFFERENTIABILITY FROM EDGE OCCLUSION

The first type of non-differentiability, as discussed in Sections 3 and 4, arises due to the migration of sharp edges. This non-differentiability is global, occurring at every point on the manifold. A second type of non-differentiability, however, can also arise on IAMs. This effect is *local*, occurring at only particular articulations where the tangents (even the regularized tangents) experience a sudden change.

5.1. Articulations in the image plane

To illustrate the basic effect, consider a simple translating-square image model. We assume a bounded image domain of $[-1, 1] \times [-1, 1]$; the occlusion of the square at the image border is the critical effect. The square indicator function has sidelength 1 and is centered at $\theta = (\theta_0, \theta_1)$. We will fix $\theta_1 = 0$ and examine the effects of changing component θ_0.

For the non-occluded regime, where $-1/2 < \theta_0 < 1/2$, it is easy to visualize the tangent images: γ_0 consists of two traveling ridges of delta functions, one with height -1 connecting the points $(\theta_0 - 1/2, \pm 1/2)$, and one with height 1 connecting the points $(\theta_0 + 1/2, \pm 1/2)$. These delta ridges are convolved with ϕ_s to obtain the regularized tangent image (see Fig. 4(a),(b)).

Consider now the occluded regime, for example $1/2 < \theta_0 < 3/2$. In this case, a portion of the square has been eliminated by the image boundary. We can equate the changing image with a *rectangle* sitting against the right side of the image, with width shrinking from the left. In this case γ_0 consists of only *one* traveling delta ridge, having height -1 and connecting the points $(\theta_0 - 1/2, \pm 1/2)$. Again, this ridge is convolved with ϕ_s to obtain the regularized tangent image (see Fig. 4(c),(d)).

This change in the tangent images is *abrupt*, occurring at precisely $\theta = [1/2, \ 0]^T$. Around this point, the manifold has differing "left" and "right" tangent images. It is simple to compute for this case that, as $s \to \infty$, then at the "moment of occlusion", there is an abrupt $45°$ change in tangent direction on each regularized manifold. This effect is not merely an artifact of the regularization process; a local PCA approximation would also be sensitive to the direction in which points are sampled.

This example demonstrates that, aside from the global issues of non-differentiability, IAMs may have localized cusps that persist even after regularization. These cusps indicate that the geometric structure relating nearby images can undergo a sudden change.
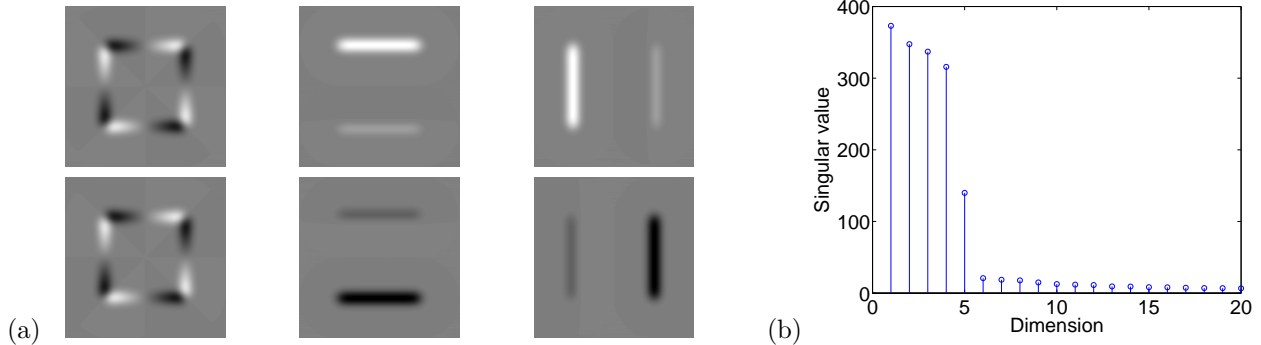
Figure 5. *(a) Tangent images for head-on view of a cube in 3-D space. Left: roll (vectors are similar). Middle: pitch (vectors are different). Right: yaw (vectors are different). (b) PCA on regularized cube images; first 20 singular values are shown.*

5.2. 3-D articulations

Occlusion-based non-differentiability is much more natural in the 3-D case and occurs when an object self-occludes and a new edge appears in view. One example is a 3-D cube viewed face-on and then rotated in some direction. Other examples include polygonal solids, cylinders (when viewed from the end), and so on.

We use two numerical experiments to illustrate this phenomenon. For these experiments, we consider a 3-D cube viewed head-on and examine the tangent space around this point under SO(3) articulations (roll, pitch, yaw) at a fixed scale. For simplicity, we assume an imaging model where the 3-D object has a parallel projection onto the image plane, and we assume that the face of the cube displays a different color/intensity than the sides.

In the first experiment, we compute local tangent approximations on the regularized manifold. We assume θ parameterizes (roll, pitch, yaw) about the face-on appearance $f_{\theta^{(0)}}$. Around $\theta^{(0)}$, we perturb each articulation parameter individually by $+\epsilon$ or $-\epsilon$ and compute the difference relative to the original image (then divide by $\pm\epsilon$ and normalize). The six resulting tangent images are shown in Fig. 5(a). The leftmost two images are almost identical, suggesting that the tangent space is smooth in the roll variable. The next two images differ significantly from one another, as do the last two. Thus with respect to the pitch and yaw parameters, the "left" and "right" tangents apparently differ. Following the arguments in Sec. 5.1, it is easy to understand what causes this discrepancy. For example, when the cube pitches forward, the image shows two moving edges at the bottom, and one at the top. Yet when the cube pitches back, the reverse is true.

In the second experiment, we perform a local PCA approximation to the manifold. We sample points randomly from the 3-D parameter space and run PCA on the resulting regularized images. Figure 5(a) shows a plot of the singular values. This plot suggests that most of the local energy is captured in a 5-D subspace.

These experiments indicate that, at the point where the cube is viewed head-on, we are at a cusp in the IAM with 5 relevant tangent directions — the manifold has a 5-D tangent complex [18] at this point. Clearly, this happens only at a small subset of all possible views of the cube (when only one face is visible). Similar effects (when only two faces are visible) give rise to 4-D tangent complexes. Otherwise, for purely *generic* views of the cube (where three faces are visible), the tangent space has only 3 dimensions, corresponding to the 3 dimensions of Θ. This typical behavior echoes the assumption of "generic view" that is common in models of visual perception [19]: in order to understand a scene, an observer might assume a view to not be accidental (such as seeing a cube face-on).

6. APPLICATION: HIGH-RESOLUTION PARAMETER ESTIMATION

With the multiscale viewpoint as background, we now consider the problem of inferring the articulation parameters from individual images. We will see that while the lack of differentiability prevents the application of conventional techniques, the multiscale perspective offers a way out. This perspective offers new justification for similar multiscale approaches employed in techniques such as image registration.

6.1. The problem

We let $\{f_\theta : \theta \in \Theta\}$ be an articulation family. Suppose we are given an image I that is believed to approximate f_θ for an unknown $\theta \in \Theta$. We aim to recover θ from I, so conceptually we seek a procedure Q that generates parameter estimates $\widehat{\theta} = Q(I)$. A natural approach to the problem is the method of nonlinear least-squares. We pose the objective function

$$D(\theta) = \|f_\theta - I\|_2^2$$

and seek the minimizing θ

$$Q(I) = \mathrm{argmin}_\theta D(\theta), \tag{6}$$

supposing that the minimum is uniquely defined.

6.2. Inspiration — Newton's Method

Standard nonlinear parameter estimation [20] tells us that, if D is differentiable, then we can use calculus to refine an initial guess $\theta^{(0)}$ at the unknown parameter; we take a Newton step from that guess in the direction of the minimum. Moreover, if the initial guess is reasonably good, then the step from $\theta^{(0)}$ to $\theta^{(1)}$ will place us dramatically closer to the minimum, squaring the error. Continuing with $\theta^{(2)}$, $\theta^{(3)}$, etc., we obtain superlinear convergence.

In the case of a *differentiable* manifold, we would let $J = [\partial D/\partial\theta_0 \quad \partial D/\partial\theta_1 \quad \ldots \quad \partial D/\partial\theta_{d-1}]^T$ be the gradient of D, and let H be the $d \times d$ Hessian, $H_{ij} = \frac{\partial^2 D}{\partial\theta_i\partial\theta_j}$. Assuming D is differentiable, Newton's method specifies the following update step:

$$\theta^{(k+1)} \leftarrow \theta^{(k)} + [H(\theta^{(k)})]^{-1} J(\theta^{(k)}).$$

To relate this method to the structure of the manifold, we can express the gradient and Hessian in terms of images, writing

$$D(\theta) = \|f_\theta - I\|_2^2 = \int (f_\theta - I)^2 \ dx = \int f_\theta^2 - 2If_\theta + I^2 \ dx.$$

Differentiating with respect to component θ_i, we obtain

$$
\begin{aligned}
\frac{\partial D}{\partial\theta_i} \ = \ J_i \ &= \ \frac{\partial}{\partial\theta_i}\left(\int f_\theta^2 - 2If_\theta + I^2 \ dx\right) \\
&= \ \int \frac{\partial}{\partial\theta_i}(f_\theta^2) - 2I\frac{\partial}{\partial\theta_i}f_\theta \ dx \\
&= \ \int 2f_\theta\tau_\theta^i - 2I\tau_\theta^i \ dx \\
&= \ 2\langle f_\theta - I, \tau_\theta^i\rangle,
\end{aligned}
$$

where $\tau_\theta^i = \frac{\partial f_\theta}{\partial\theta_i}$ is a tangent image. Continuing, we examine the Hessian,

$$
\begin{aligned}
\frac{\partial^2 D}{\partial\theta_i\partial\theta_j} \ = \ H_{ij} \ &= \ \frac{\partial}{\partial\theta_j}\left(\frac{\partial D}{\partial\theta_i}\right) \\
&= \ \int \frac{\partial}{\partial\theta_j}\left(2f_\theta\tau_\theta^i - 2I\tau_\theta^i\right) \ dx \\
&= \ \int 2\tau_\theta^i\tau_\theta^j + 2f_\theta\tau_\theta^{ij} - 2I\tau_\theta^{ij} \ dx \\
&= \ 2\langle\tau_\theta^i, \tau_\theta^j\rangle + 2\langle f_\theta - I, \tau_\theta^{ij}\rangle, \tag{7}
\end{aligned}
$$

where $\tau_\theta^{ij} = \frac{\partial^2 f_\theta}{\partial\theta_i\partial\theta_j}$ denotes a second-derivative image.

To summarize, an algorithm employing Newton's method to solve the minimization (6) would require an iterative procedure, where each iteration would involve projecting onto tangent images (as well as second derivative images). In our setting, however, the tangent vectors τ_θ^i *do not exist* as functions, making it impossible to directly implement such an algorithm. We turn again to the regularization process in order to remedy this situation.

6.3. Multiscale Newton algorithm

As discussed in Sec. 3, the lack of differentiability can be alleviated by regularizing the images f_θ. Thus, navigation is possible on any of the regularized manifolds $\mathbf{F}_s$ using Newton's method as described above. This fact, in conjunction with the convergence property (2), suggests a multiscale technique for parameter estimation. Note that we focus on dealing with "migration-based" non-differentiability from Sec. 3. In cases where we have occasional occlusion-based non-differentiability as in Sec. 5, it may be necessary to project onto additional tangent images; this adaptation is not difficult, but it does require an awareness of the parameters at which occlusion-based non-differentiability occurs.

The idea is to select a sequence of scales $s_0 > s_1 > \cdots > s_K$, and to start with an initial guess $\theta^{(0)}$. At each scale we take a Newton-like step on the corresponding smoothed manifold. We find it helpful in practice to ignore the second derivative term from equation (7). This is in the typical spirit of making slight changes to Newton's Method; in fact it is similar to the Gauss-Newton method for minimizing D.

To be specific, iteration $k + 1$ of the Multiscale Newton algorithm proceeds as follows:

1. Compute the local tangent vectors on the smoothed manifold $\mathbf{F}_{s_k}$ at the point $f_{\theta^{(k)}, s_k}$:

$$\tau^i_{\theta^{(k)}, s_k} = \left. \frac{\partial}{\partial \theta_i} f_{\theta, s_k} \right|_{\theta = \theta^{(k)}}, \quad i = 0, 1, \ldots, d - 1.$$

2. Project the estimation error $f_{\theta^{(k)}, s_k} - I_{s_k}$ (relative to the *regularized* image $I_{s_k} = \phi_{s_k} * I$) onto the tangent space $T(s_k, \theta^{(k)})$, setting

$$J_i = 2\langle f_{\theta^{(k)}, s_k} - I_{s_k}, \tau^i_{\theta^{(k)}, s_k} \rangle.$$

3. Compute the pairwise inner products between tangent vectors

$$H_{ij} = 2\langle \tau^i_{\theta^{(k)}, s_k}, \tau^j_{\theta^{(k)}, s_k} \rangle.$$

4. Use the projection coefficients to update the estimate

$$\theta^{(k+1)} \leftarrow \theta^{(k)} + H^{-1} J.$$

We note that when the tangent vectors are orthogonal to one another, H is diagonal, and so the update for component $\theta_i^{(k)}$ is simply determined by the inner product of the estimation error vector and the tangent vector $\tau^i_{\theta^{(k)}, s_k}$. Moreover, when the regularized manifold $\mathbf{F}_{s_k}$ is linear in the range of interest, the update in Step 4 immediately achieves the minimizer to D at that scale.

Under certain conditions on the accuracy of the initial guess and the sequence $\{s_k\}$ it can be shown that this algorithm provides estimation accuracy $\|\theta - \theta^{(k)}\| < cs_k^2$. Ideally, we would be able to square the scale between successive iterations, $s_{k+1} = s_k^2$. The exact sequence of steps, and the accuracy required of the initial guess $\theta^{(0)}$, will depend on the specific multiscale structure of the IAM under consideration. We omit the convergence analysis in this paper, instead providing several examples to demonstrate the basic effectiveness of the algorithm.

6.4. Examples

6.4.1. Translating disk

As a basic exercise of the proposed algorithm, we attempt to estimate the articulation parameters for a translated disk. The process is illustrated in Fig. 6. The observed image I is shown on the far left; the top-left image in the grid is the initial guess $f_{\theta^{(0)}}$. For this experiment, we create 256×256 images with "subpixel" accuracy (each pixel is assigned a value based on the proportion of its support that overlaps the disk). Regularized tangent images are estimated using a local difference of synthesized (and regularized) images.

Table 1. *Estimation errors of multiscale Newton iterations, translating disk, no noise.*

s	θ_0 error	θ_1 error	image MSE
Initial	-1.53e-01	1.92e-01	9.75e-02
1/2	-2.98e-02	5.59e-02	3.05e-02
1/4	-4.50e-04	1.39e-03	1.95e-04
1/16	-1.08e-06	8.62e-07	8.29e-10
1/256	1.53e-08	1.55e-07	1.01e-10

Table 2. *Estimation errors of multiscale Newton iterations, translating disk, with noise. MSE between noisy image and true disk = 3.996.*

s	θ_0 error	θ_1 error	image MSE
Initial	-1.53e-01	1.93e-01	4.092
1/2	-3.46e-02	7.40e-02	4.033
1/4	-1.45e-02	-2.61e-03	4.003
1/16	-1.55e-03	-1.77e-03	3.997
1/256	-5.22e-04	1.10e-03	3.996

Table 3. *Estimation errors after multiscale Newton iterations, ellipse.*

s	θ_0 error	θ_1 error	θ_2 error	θ_3 error	θ_4 error	image MSE
Initial	-5.75e-02	3.95e-02	-8.16e+00	7.72e-02	-3.56e-02	8.47e-02
1/2	5.82e-02	1.48e-02	7.91e-01	-3.66e-02	-8.74e-03	3.62e-02
1/4	-4.86e-03	-1.56e-03	-4.14e+00	3.19e-02	-1.28e-02	1.91e-02
1/16	4.25e-04	1.99e-04	-7.95e-01	-2.64e-03	-1.05e-03	1.42e-03
1/256	-3.61e-05	2.71e-05	-3.38e-03	-1.49e-04	-3.86e-05	2.72e-06

Table 4. *Estimation errors after multiscale Newton iterations, 3-D icosahedron. MSE between noisy image and true original = 2.98.*

s	θ_0 error	θ_1 error	θ_2 error	θ_3 error	θ_4 error	θ_5 error	image MSE
Initial	-50	-23	20	1.00e-01	-1.00e-01	5.00e-01	3.13
1/2	-8.81e+01	1.53e+00	6.07e+01	-2.60e-02	5.00e-02	-3.28e-01	3.14
1/4	-5.29e+01	4.70e+00	2.44e+01	3.44e-03	2.42e-02	4.24e-02	3.10
1/8	-1.15e+01	1.12e+00	-9.44e-01	-4.34e-03	3.19e-03	1.26e-01	3.03
1/16	8.93e-01	3.00e-01	-1.69e+00	-1.38e-03	2.21e-03	3.40e-02	2.98
1/256	5.28e-01	2.57e-01	-6.68e-01	6.91e-04	2.44e-03	2.12e-02	2.98

We run the multiscale estimation algorithm using the sequence of stepsizes $s = 1/2, 1/4, 1/16, 1/256$. Figure 6 shows the basic computations of each iteration. Note the geometric significance of the smoothed difference images $I_{s_k} - f_{\theta^{(k)}, s_k}$; at each scale this image is projected onto the tangent plane basis vectors. Table 1 gives the estimation errors at each iteration, both for the articulation parameters θ and the mean square error (MSE) of the estimated image. Using this sequence of scales, we observe rapid convergence to the correct articulation parameters with accuracy far better than the width of a pixel, $1/256 \approx 3.91$e-03.

We now run a similar experiment for the case where the observation $I = f_\theta + n$, where n consists of additive white Gaussian noise of variance 4. Using the same sequence of smoothing filter sizes, the results are shown in Fig. 7 and in Table 2. Note that the estimated articulation parameters are approximately the best possible, since the resulting MSE is approximately equal to the noise energy.

6.4.2. Articulated ellipse

We run a similar experiment for an ellipse image. In this case, the parameter space Θ is 5-D, with two directions of translation, one rotation parameter, and two parameters for the axis lengths of the ellipse. Figure 8 and Table 3 show the estimation results. It is particularly interesting to examine the geometric structure among the tangent vectors and how it reflects the different effects of changing the articulation parameters. Again, the algorithm successfully estimates the articulation parameters with high accuracy.

6.4.3. 3-D articulations

We now consider a different imaging modality, where we have articulations of a 3-D object. In this case, the parameter space Θ is 6-D; the articulations of the object are three rotational coordinates, two shift coordinates parallel to the image plane, and one shift toward/away from the camera. (We now use a pinhole imaging model, so motions toward the camera make the object appear larger.)

For this example, we consider synthesized photographs of an icosahedron. Our image model includes a directional light source (with location and intensity parameters assumed known). We consider color images,
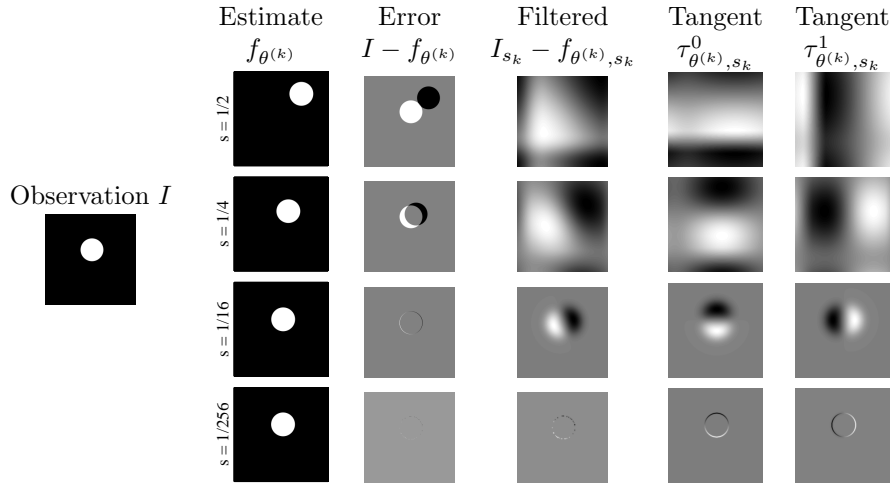
Figure 6. *Multiscale estimation of translation parameters for observed disk image. Each row corresponds to the smoothing and tangent basis vectors for one iteration.*
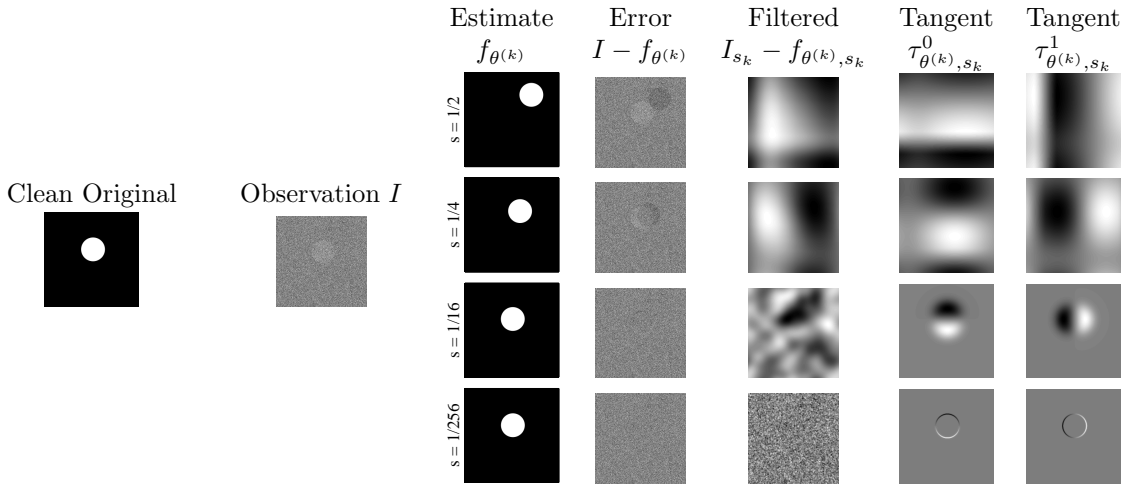

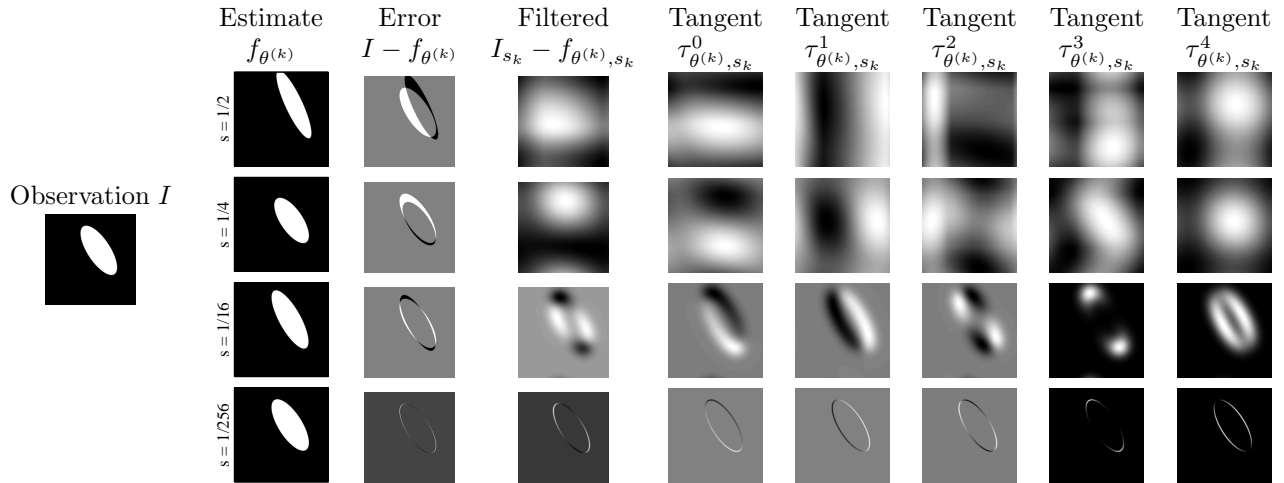**Figure 7**. *Multiscale estimation of translation parameters for observed disk image with noise.*

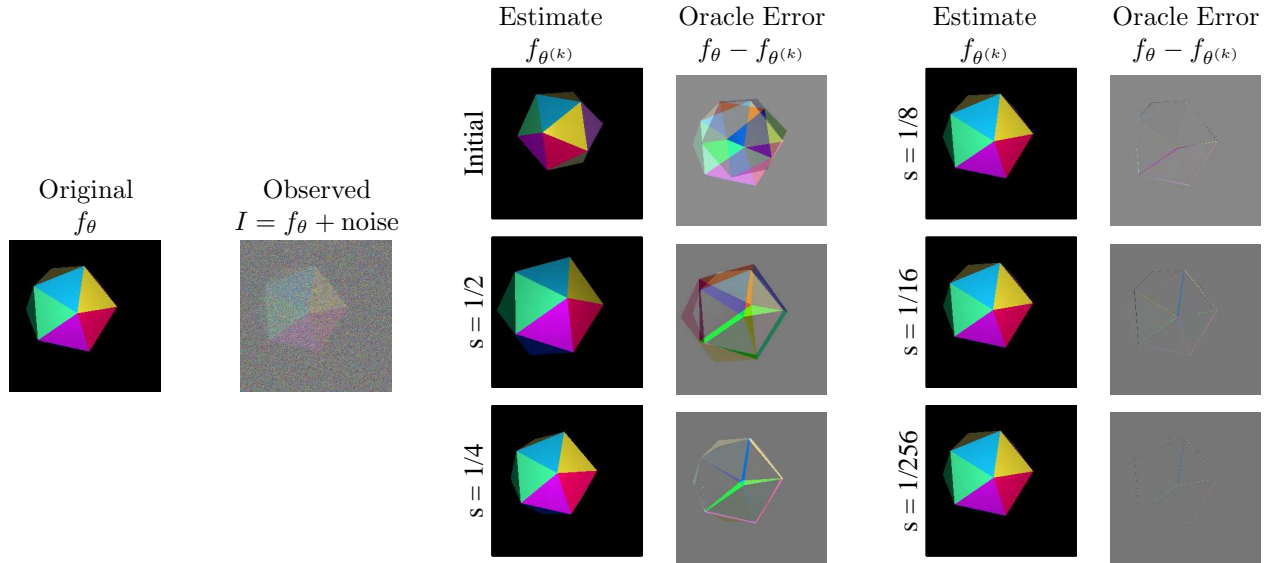**Figure 8**. *Multiscale estimation of articulation parameters for ellipse.*

Figure 9. *Multiscale estimation of articulation parameters for 3-D icosahedron.*

treating each image as an element of $R^{256 \times 256 \times 3}$. Figure 9 and Table 4 show the successful estimation of the articulation parameters for a noisy image. For this example, we must use a slightly less ambitious sequence of smoothing filters. (In this case, while we successfully ignore the occlusion-based effects of the appearance/disappearance of faces, we find that these should not be ignored in general.)

6.5. Related work

Our multiscale framework for estimation with IAMs shares common features with a number of practical image registration algorithms; space considerations permit discussion of only a few here. Irani and Peleg [7] have developed a popular multiscale algorithm for registering an image $I(x)$ with a translated and rotated version for the purposes of super-resolution. They employ a multiscale pyramid to speed up the algorithm and to improve accuracy, but a clear connection is not made with the non-differentiability of the corresponding IAM. While Irani and Peleg compute the tangent basis images with respect to the x_0 and x_1 axes of the image, Keller and Averbach [21] compute them with respect to changes in each of the registration parameters. They also use a multiscale pyramid and conduct a thorough convergence analysis. Belhumeur [14] develops a tangent-based algorithm that estimates not only the pose of a 3-D object, but also its illumination parameters.

Where we differ from these approaches is in deriving the multiscale approach from the structure of the underlying manifold and in explaining the properties of the algorithm (e.g., how quickly the scale can be decreased) in terms of the twisting of the tangent space. More importantly, our approach is general and in principle extends far beyond the registration setting to many other image understanding problems, of learning parameters from example images.

7. DISCUSSION AND CONCLUSIONS

In this paper, we have studied the structural features of IAMs and observed two types of non-differentiability that can arise on IAMs corresponding to seemingly innocuous image models. These effects are highly relevant for certain image processing algorithms that might attempt to use calculus to exploit the structural relationships among "similar" images. Fortunately, although tangent spaces do not exist, a multiscale collection of approximate tangent spaces does exist and leads to the possibility of valid multiscale calculus algorithms. In areas such as image registration, similar multiscale solutions have been found empirically and are considered important tools. We suspect that there are many other settings in which multiscale approximate tangent spaces can be of use.

Regarding our Multiscale Newton algorithm, in addition to the convergence analysis mentioned in Sec. 6.3, a number of issues remain open. For instance, with noisy images the multiscale tangent projections will reach

a point of diminishing returns where finer scales will not benefit; we must develop a stopping criterion for such cases. Additional issues revolve around efficient implementation. We believe that a sampling of the tangent planes needed for the projections can be precomputed and stored using the multiscale representation of [22]. Moreover, since many of the computations are local (as evidenced by the support of the tangent basis images in Fig. 2), we expect that the image projection computations can be implemented in the wavelet domain. This would also lead to a fast method for obtaining the initial guess $\theta^{(0)}$ with the required accuracy.

Appendix — Proof of Theorem 4.1

We start with

$$
\begin{aligned}
c_{s_0,s_1} &= \int_{\partial B} \int_{\partial B} \cos(b)\,\cos(\beta)\,\phi_{\sqrt{s_0^2+s_1^2}}(\vec{b}-\vec{\beta})\,d\vec{\beta}\,d\vec{b} \\
&= \int_0^{2\pi} \cos(b) \int_0^{2\pi} \frac{\cos(\beta)}{2\pi(s_0^2+s_1^2)} \exp\left\{ \frac{-2(1-\cos(b-\beta))}{2(s_0^2+s_1^2)} \right\}\,d\beta\,db.
\end{aligned}
\tag{8}
$$

We now proceed less formally, though the detailed epsilontics can easily be supplied. Note that

$$
2 - 2\cos(x) = x^2 + h(x),
$$

where $h(x)$ is smooth and $O(x^4)$ as $x \to 0$. Hence

$$
\exp\{-n^2(2-2\cos(x))\} = \exp\{-n^2x^2\}\exp\{-n^2h(x)\}.
$$

Now note that $\phi_{1/n}(x) = \frac{n}{\sqrt{\pi}}\exp\{-n^2x^2\}$ is the probability density of a Gaussian distribution concentrated on $(-1/n, +1/n)$, and that $\exp\{-n^2h(x)\}$ is bounded and obeys

$$
\left\| \exp\{-n^2h(x)\} - 1 \right\|_{L^\infty[-1/n^{2/3},+1/n^{2/3}]} = O(n^{-2}).
$$

Hence

$$
f_n(x) \equiv \frac{n}{\sqrt{\pi}} \exp\{-n^2(2-2\cos(x))\}
$$

is essentially a Gaussian probability density and obeys

$$
\cos(b) = \lim_{n\to\infty} \int_{\partial B} \cos(\beta)\,\frac{n}{\sqrt{\pi}} f_n(\beta - b)\,d\beta.
$$

Strengthening this statement with remainder terms shows that we can approximate the integrand (8) by

$$
\int_0^{2\pi} \cos(b)\,\frac{\cos(b)}{\sqrt{2\pi(s_0^2+s_1^2)}}\,db.
$$

This in turn means we can write

$$
c_{s_0,s_1} = (1+o(1))\frac{\sqrt{\pi}}{\sqrt{2(s_0^2+s_1^2)}}
$$

where the asymptotic refers to the situation where $s_1 = \alpha s_0$, and $s_0 \to 0$. Under this same asymptotic,

$$
\frac{c_{s_0,\alpha s_0}}{c_{s_0,s_0}^{1/2}\,c_{\alpha s_0,\alpha s_0}^{1/2}} \sim \frac{\frac{\sqrt{\pi}}{s_0\sqrt{2(1+\alpha^2)}}}{\left(\frac{\sqrt{\pi}}{2s_0}\right)^{1/2}\left(\frac{\sqrt{\pi}}{2\alpha s_0}\right)^{1/2}} \sim \frac{\sqrt{2\alpha}}{\sqrt{1+\alpha^2}}.
$$

□

REFERENCES

1. J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science* **290**, pp. 2319–2323, December 2000.

2. S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science* **290**, pp. 2323–2326, December 2000.

3. D. L. Donoho and C. E. Grimes, "Hessian Eigenmaps: Locally linear embedding techniques for high-dimensional data," *Proc. Natl. Acad. Sci. USA* **100**, pp. 5591–5596, May 2003.

4. B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. 7th Int'l Joint Conf. on Artificial Intelligence*, pp. 674–679, (Vancouver), 1981.

5. L. Quam, "Hierarchical warp stereo," in *Proc. DARPA Image Understanding Workshop*, pp. 149–155, September 1984.

6. W. Enkelmann, "Investigations of multigrid algorithms for the estimation of optical flow fields in image sequences," *Comp. Vision, Graphics, and Image Processing* **43**, pp. 150–177, 1988.

7. M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical Models and Image Processing* **53**, pp. 231–239, May 1991.

8. E. P. Simoncelli, "Coarse-to-fine estimation of visual motion," in *Proc 8th Workshop on Image and Multidimensional Signal Processing*, pp. 128–129, IEEE Sig Proc Society, (Cannes, France), September 1993.

9. E. P. Simoncelli, "Bayesian multi-scale differential optical flow," in *Handbook of Computer Vision and Applications*, B. Jähne, H. Haussecker, and P. Geissler, eds., **2**, ch. 14, pp. 397–422, Academic Press, San Diego, April 1999.

10. C. Grimes, *New methods in nonlinear dimensionality reduction.* PhD thesis, Department of Statistics, Stanford University, 2003.

11. D. L. Donoho and C. Grimes, "Image manifolds which are isometric to euclidean space," *J. Math. Imaging and Computer Vision* **23**, July 2005.

12. D. L. Donoho, "Wedgelets: Nearly-minimax estimation of edges," *Annals of Stat.* **27**, pp. 859–897, 1999.

13. M. B. Wakin, J. K. Romberg, H. Choi, and R. G. Baraniuk, "Wavelet-domain approximation and compression of piecewise smooth images," *IEEE Trans. Image Processing*, 2005. To appear.

14. P. N. Belhumeur and G. D. Hager, "Tracking in 3D: Image variability decomposition for recovering object pose and illumination," *Pattern Analysis and Applications* **2**, pp. 82–91, 1999.

15. M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognitive Neuroscience* **3**(1), 1991.

16. C. Davis and W. Kahan, "The rotation of eigenvectors by a perturbation, III," *SIAM J. Numer. Anal.* **7**(1), pp. 1–46, 1970.

17. G. W. Stewart and J. guang Sun, *Matrix perturbation theory*, Academic Press, Boston, 1990.

18. G. Carlsson, A. Zomorodian, A. Collins, and L. Guibas, "Persistence barcodes for shapes," *Int. J. of Shape Modeling.* To appear.

19. W. T. Freeman, "Exploiting the generic viewpoint assumption," *Int. J. Computer Vision* **20**(3), 1996.

20. D. M. Bates and D. G. Watts, *Nonlinear Regression Analysis and Its Applications*, John Wiley and Sons, New York, 1988.

21. Y. Keller and A. Averbach, "Fast motion estimation using bidirectional gradient methods," *IEEE Trans. Image Processing* **13**, pp. 1042–1054, August 2004.

22. I. U. Rahman, I. Drori, V. C. Stodden, D. L. Donoho, and P. Schroeder, "Multiscale representations for manifold-valued data," 2005. Preprint.