On Growth and Formlets: Sparse Multi-Scale Coding of Planar Shape

James H. Elder[a,*], Timothy D. Oleskiw[b], Alex Yakubovich[a], Gabriel Peyré[c]

[a]*Centre for Vision Research, York University, Toronto, Canada*
[b]*Department of Applied Mathematics, University of Washington, Seattle, WA, United States*
[c]*CEREMADE, Université Paris-Dauphine, Paris, France*

Abstract

We propose a sparse representation of 2D planar shape through the composition of warping functions, termed formlets, localized in scale and space. Each formlet subjects the 2D space in which the shape is embedded to a localized isotropic radial deformation. By constraining these localized warping transformations to be diffeomorphisms, the topology of shape is preserved, and the set of simple closed curves is closed under any sequence of these warpings. A generative model based on a composition of formlets applied to an embryonic shape, e.g., an ellipse, has the advantage of synthesizing only those shapes that could correspond to the boundaries of physical objects. To compute the set of formlets that represent a given boundary, we demonstrate a greedy coarse-to-fine formlet pursuit algorithm that serves as a non-commutative generalization of matching pursuit for sparse approximations. We evaluate our method by pursuing partially occluded shapes, comparing performance against a contour-based sparse shape coding framework.

Keywords: planar shape, deformation, sparse coding sep contour completion

1. Introduction

Shape information is important for a broad range of computer vision problems. For some detection and recognition tasks, discriminative models that use non-invertible shape codes (e.g., [1]) can be effective. However, many other tasks call for a more complete generative model of shape. Examples include: (1) shape segmentation, recognition, and tracking in cluttered scenes, where

---

*Corresponding author.
Email addresses: `jelder@yorku.ca` (James H. Elder), `oleskiw@uw.edu` (Timothy D. Oleskiw), `yakuboa@yorku.ca` (Alex Yakubovich), `gabriel.peyre@ceremade.dauphine.fr` (Gabriel Peyré)
URL: `www.yorku.ca` (James H. Elder),
`http://depts.washington.edu/amath/people/Timothy.Oleskiw` (Timothy D. Oleskiw),
`elderlab.yorku.ca/~alex` (Alex Yakubovich), `http://www.ceremade.dauphine.fr/~peyre` (Gabriel Peyré)

shapes must be distinguished not just from each other, but from 'phantom' shapes formed by conjunctions of features from multiple objects [2]; (2) modeling of shape articulation, growth, and deformation; and (3) modeling of shape similarity.

Our paper concerns the generative modeling of natural 2D shapes in the plane, represented by their 1D boundary. We restrict our attention to simply-connected shapes whose boundaries are smooth, simple, and closed curves. We seek a generative shape model that satisfies a set of properties that seem to us essential:

1. *Completeness.* The model can produce all shapes.
2. *Closure.* The set of valid shapes is closed under the generative model. In other words, the model generates only valid shapes.
3. *Composition.* Complex shapes are generated by combining simpler components.
4. *Sparsity.* Good approximations of shape can be generated with relatively few components.
5. *Progression.* Approximations can be improved by incorporating more components.
6. *Locality.* Components are localized in space.
7. *Scaling.* Components are tuned to specific scales and are self-similar over scale.
8. *Region & Contour.* Components can capture both region and contour properties in a natural way.

The need for completeness is self-evident if the system is to be general. Closure is critical if we hope to capture the statistics of natural shape in a set of hidden generative variables. Without closure, heuristics must be used to avoid the generation of invalid shapes, e.g., bounding contours with self-intersections. Aside from the resulting inefficiency, this creates a discrepancy between the statistical structure encoded by the model, and samples the model produces. In other words, the model cannot fully capture the statistics of natural boundaries.

Composition (here we use the word in a general sense) is important if we are to handle the richness and complexity of natural shapes while maintaining conceptual simplicity. Given the high dimensionality of natural shapes, sparsity is necessary in order to store shape models [3]. Sparsity also implies that essential shape features have been made explicit [4]. Progression allows the complexity of the model to be matched to the difficulty of the task, facilitating real-time operation and coarse-to-fine optimization.

Locality is a natural goal, since a first-order property of natural images is local coherence. Nearby points on the surface of an object tend to have similar reflectance, attitude, and illumination. Locality also allows for greater robustness to occlusion, since components are more likely to be either entirely visible or removed altogether rather than distorted. Scaling allows invariance over object size, and allows shape features of different sizes to be captured separately.

Finally, it has long been recognized that planar shape description requires attention to both region and contour properties [3]. Some shape properties, e.g., curvature, are naturally described by the bounding contour. Others, e.g., necks, are best described as region properties, since they involve points that are proximal in the image but distant along the contour. A good generative model will allow both to be encoded in a natural way.

We begin by reviewing prior models, with an eye to each of these essential properties.

2. Prior Work

Early models that used chain coding or splines to encode shapes were not generative and failed to succinctly capture global properties of shape. Fourier descriptor, moment, and PCA bases have the potential to be generative, but since all components are global, they are not robust to occlusion or local deformation [5, 3, 6]. For these reasons, most modern approaches attempt to capture structure at intermediate scales, or over a range of scales. Most of these models can be crudely partitioned into two classes: contour-based and symmetry-based.

2.1. Contour-Based Models

Attneave [4] pointed to the concentration of information in the curvature of the bounding contour, and suggested the potential for sparse descriptions based on points of extremal curvature magnitude. Hoffman & Richards [7] linked curvature to the part structure of shapes, proposing that parts are perceptually segmented at negative minima of curvature. Mokhtarian & colleagues emphasized the encoding of curvature inflections across scale space for the purpose of shape recognition [8].

While none of these early models are generative, Dubinskiy & Zhu [9] have more recently proposed a contour-based shape representation that is both generative and sparse. The theory is based upon the representation of a shape by a summation of component *shapelets*. A shapelet is a primitive curve defined by Gabor-like coordinate functions that map arclength to the image, which can be represented by the complex plane.

Specifically, a shapelet $\gamma(t; \sigma, \mu)$ is a mapping of arc length $t \in [0, 1]$ to the image, represented by the complex plane. Each shape let is parameterized by an arc length position parameter μ and a scale parameter σ, and has the specific form:

$$\gamma(t; \sigma, \mu) = \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)\left(\cos\left(\frac{2\pi}{\sigma}(t-\mu)\right) + i\sin\left(\frac{2\pi}{\sigma}(t-\mu)\right)\right). \quad (1)$$

Figure 1 shows the coordinate functions and trace of an example shapelet. Note that the planar curves generated by $\gamma(t; \sigma, \mu)$ are identical on $t \in \mathbb{R}$ up to a linear reparameterization, i.e., they are *self-similar*. However, these functions are only approximately self-similar on any finite domain over which a curve will

be defined. Also, note that γ does not in general generate a simple closed curve. In fact, as $\sigma \to 0$, the number of sinusoidal periods on the interval $t \in [0, 1]$ explodes, generating an infinite number of self-intersections.
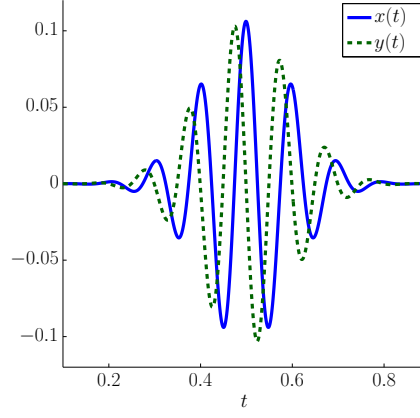
(a) Component functions
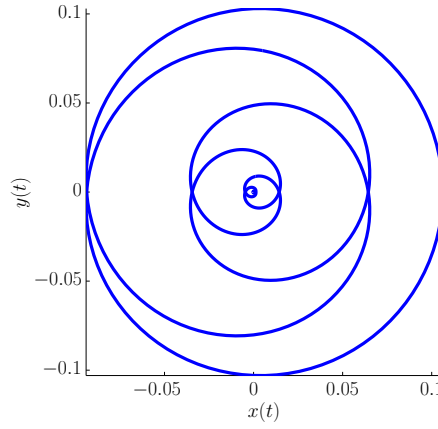
(b) Generated image trace

Figure 1: An example shapelet.

Shifting and scaling shapelets over arclength produces a basis set sufficient to generate arbitrarily complex shapes. In particular, a K-shapelet curve $\Gamma^K(t)$ can be defined as:

$$\Gamma^K(t) = z_0 + \sum_{k=1}^{K} A_k \gamma(t; \sigma_k, \mu_k), \tag{2}$$

4

where the 2×2 matrix A_k applies an affine transformation to each shapelet in image space prior to linear combination.

Dubinskiy & Zhu's shapelet model has many positive features. Components are localized, albeit only in arclength, and scale is made explicit in a natural way. However, like all contour-based methods, the shapelet theory does not explicitly capture regional properties of shape. Perhaps most crucially, the model does not respect the topology of object boundaries: sampling from the model will in general yield non-simple, i.e., self-intersecting, curves (Figure 2). This violates the closure criterion identified in Section 1.
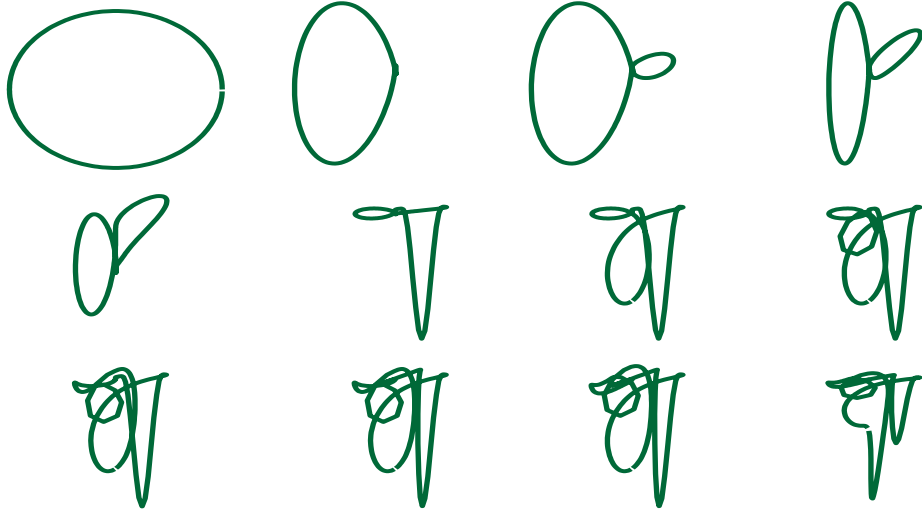

Figure 2: Sampling from the shapelet model generally yields non-simple curves.

2.2. Symmetry-Based Models

Blum and colleagues [10, 11] introduced the symmetry axis representation of shape in which a planar shape is represented by a 1D skeleton function and associated 1D radius function. The symmetry axis representation led to related representations [12] which found application in medical imaging and other domains.

Subsequent work incorporated notions of scale and time with symmetry axis descriptions. Leyton [13] related symmetry axis descriptions to causal deformation processes acting upon prototype shapes. In this view, symmetry axes, terminating at curvature extrema on the boundary, are understood as records of these deformation processes. Subsequent work on curve evolution methods and shock-graph representations [14, 15] has provided a more complete theory of region-based shape representation that has been broadly applied.

Despite the many appealing features of symmetry axis and shock-graph representations, these methods, in general, are not sparse. In fact, the description of

each shape typically requires more storage, and little emphasis has been placed on making symmetry axis representations generative [3]. Recent work of Trinh and Kimia exploring generative and sparse models based upon shock graphs comes some way in overcoming these limitations [16]. However, the constraints required to enforce the closure property, i.e., topological constraints, are fairly complex, and the full potential of the theory has yet to be explored.

A related approach to shape representation (e.g., [17, 18] employs finite element modelling techniques to code the bounding contour in terms of the free vibration modes of the shape, which are said to correspond to the object's generalized axes of symmetry. The main difficulty in developing this approach into a generative model is that points on the boundary are coupled only locally in the intrinsic coordinates of the shape boundary, thus nothing constrains the topology of generated shapes.

2.3. Hybrid Approaches

Recognizing the merits and limitations of both contour-based and symmetry-based approaches, Zhu [19] developed an MRF model for natural 2D shape, employing a neighbourhood structure that can directly encode both contour-based and region-based Gestalt principles. The theory is promising in many respects. It is generative, providing an explicit probabilistic model, and it captures both region and contour properties. It is not sparse, however, and because the underlying graph is lifted from the image plane, there is nothing in the model that encodes the topological constraint that the boundary be simple, i.e., non-intersecting. Instead, when sampling from the model, a 'firewall' is employed to prevent intersections. Again, this is inefficient, and it also creates a disconnect between the generative variables encoding the model and the sampling distribution.

2.4. Coordinate Transformations

A different class of model that could also be called region-based involves the application of coordinate transformations of the planar space in which a shape is embedded. This idea can be traced back at least to D'Arcy Thompson, who considered specific classes of global coordinate transformations of the plane to model the relationship between the shapes of different animal species [20]. In the field of computer vision, Jain et al. [21] were among the first to extend this idea to more general deformations with a complete Fourier deformation basis that they used to match observed shapes to stored prototypes. However, this Fourier basis fails to satisfy the locality property, and as a potential generative model it does not satisfy the closure property: random combinations of Fourier deformation components will not in general preserve the topology of the prototype curve.

More recently, Sharon & Mumford [22] have explored conformal mappings as global coordinate transformations between planar shapes. However, although the Riemann mapping theorem guarantees that any simple closed curve can

be conformally mapped to the unit circle, conformal mappings do not in general preserve the topology of embedded contours. Hence, despite the computational constraints imposed by the Cauchy-Riemann equations, we again have the problem that the set of valid bounding contours is not closed under these transformations, making generative modeling difficult.

2.5. Localized Diffeomorphisms: Formlets

In considering prior generative shape models, the goal that seems most elusive is that of closure: ensuring that the model generates only valid shapes. Our approach originates with the observation that, while general smooth coordinate transformations of the plane will not preserve the topology of an embedded curve, it is straightforward to design a specific family of diffeomorphic transformations that will. It then follows immediately by induction that a generative model based upon arbitrary sequences of diffeomorphisms will satisfy the closure property.

In this paper we specifically consider a family of diffeomorphisms we call *formlets*. A formlet is a simple, isotropic, radial deformation of planar space that is localized within a specified circular region of a selected point in the plane. The family comprises formlets over all locations and spatial scales. While the gain of the deformation is also a free parameter, it is constrained to satisfy a simple criterion that guarantees that the formlet is a diffeomorphism. Since topological changes in an embedded figure can only occur if the deformation mapping is either discontinuous or non-injective, these diffeomorphic deformations are guaranteed to preserve the topology of embedded figures. Thus the model satisfies the closure property.

By construction, formlets satisfy the desired locality and scaling properties. It is straightforward to show that the model also satisfies the composition, completeness, and progression properties in that an arbitrary shape can be approximated to increasing precision by composing an appropriate sequence of localized formlets. Since each formlet may be centered either near the contour, near a symmetry axis, or at any other location in the plane, the model has the potential to capture both region and contour properties directly.

Our formlet model is closely related to recent work by Grenander et al. [23], modeling changes to anatomical parts over time. Their representation, called Growth by Random Iterated Diffeomorphisms (GRID), models growth as a sequence of local and radial deformations. They demonstrate their model by tracking growth in the rat brain, as revealed in sequential planar sections of MRI data.

In the present paper we explore the possibility that these ideas could be extended to model not just differential growth between sequential shapes, but to serve as the basis for a generative model over the entire space of smooth shapes, based upon a universal embryonic shape in the plane such as an ellipse.

Elements of the present paper were first reported at CVPR [24]. The main contributions of this conference paper were:

1. We illustrated the completeness and closure properties of the formlet model through random generation of sample shapes.

7

2. To solve the inverse problem of modeling given shapes, we developed and applied a generalization of matching pursuit, which selects the sequence of formlets that minimizes approximation error. We demonstrated that this *formlet pursuit* algorithm allows for progressive approximation of shape, while preserving topological properties.

3. We assessed the robustness of the formlet model to occlusion by evaluating it on the problem of contour completion. We found that the model compares favourably with the contour-based shapelet model [9] on this important problem.

In the present paper we elaborate substantially on these contributions, including full derivations and complete implementation details. But we also build on this work with several important new contributions:

1. We introduce a method for handling analytically computed optimal gain values that exceed the diffeomorphism bounds.

2. We develop and evaluate an improved parameter optimization method called dictionary descent, and show that it increases accuracy by 11% and decreases run time by 42%, relative to standard dictionary pursuit.

3. We provide derivations for the Jacobian required for this new dictionary descent method.

4. We develop, evaluate and compare several alternative mathematical formulations of the formlet function.

5. We report statistics of formlet model parameters for our database of animal shapes, demonstrating coarse-to-fine scaling properties and an interesting anisotropy in the location distribution.

3. Formlet Coding

3.1. Formlet Bases

We represent the image in the complex plane $\mathbb{C}$, and define a formlet $f : \mathbb{C} \to \mathbb{C}$ to be a diffeomorphism of the complex plane localized in scale and space. Such a deformation can be realized by centering f about the point $\zeta \in \mathbb{C}$ and allowing f to deform the plane within a $(\sigma \in \mathbb{R}^+)$-region of ζ. Our Gabor-inspired deformation is defined as

$$f(z; \zeta, \sigma, \alpha) = \zeta + \frac{z - \zeta}{|z - \zeta|} \rho(|z - \zeta|; \sigma, \alpha), \quad \text{where}$$

$$\rho(r; \sigma, \alpha) = r + \alpha \sin\left(\frac{2\pi r}{\sigma}\right) \exp\left(\frac{-r^2}{\sigma^2}\right). \tag{3}$$

Thus each formlet $f : \mathbb{C} \to \mathbb{C}$ is a localized isotropic and radial deformation of the plane at location ζ and scale σ. The magnitude of the deformation is controlled by the gain parameter $\alpha \in \mathbb{R}$. Figure 3 demonstrates formlet deformations of the plane with positive and negative gain.
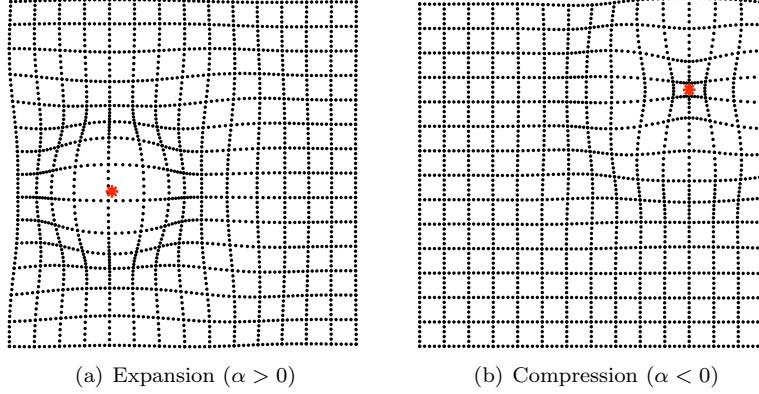
(a) Expansion ($\alpha > 0$) (b) Compression ($\alpha < 0$)

Figure 3: Example formlet deformations. The location of the formlet ζ is indicated by the asterisk.

3.2. Diffeomorphism Constraint

Without any constraints on the parameters, these deformations, though continuous, can fold the plane on itself, changing the topology of an embedded contour. In order to preserve topology, we must constrain the gain parameter to guarantee that each deformation is a diffeomorphism. As the formlets defined in Equation 3 are both isotropic and angle preserving, it is sufficient to require that the radial deformation ρ be a diffeomorphism of $\mathbb{R}^+$, i.e., that $\rho(r; \sigma, \alpha)$ be strictly increasing in r:

$$\frac{\partial}{\partial r} \rho(r; \sigma, \alpha) > \quad 0$$

$$\Rightarrow \quad \alpha \frac{\partial}{\partial r} \sin\left(\frac{2\pi r}{\sigma}\right) \exp\left(\frac{-r^2}{\sigma^2}\right) > \quad -1 \tag{4}$$

$$\Rightarrow \quad \frac{2\alpha}{\sigma} \exp\left(-\frac{r^2}{\sigma^2}\right) \left(\pi \cos\left(\frac{2\pi r}{\sigma}\right) - \frac{r}{\sigma} \sin\left(\frac{2\pi r}{\sigma}\right)\right) > \quad -1$$

For $\alpha < 0$, it is easy to see that the minimal slope of ρ is attained as $r \to 0^+$. Evaluating Equation 4 at $r = 0$ thus yields the lower-bound on the gain α:

$$\alpha > -\frac{\sigma}{2\pi}. \tag{5}$$

For positive α, the location of the minimum in $\rho'(r)$ does not have a closed form solution, but can be computed numerically:

$$\alpha \lessapprox 0.1956\sigma. \tag{6}$$

Thus the diffeormorphism constraint is:

$$\alpha \in \sigma \left(-\frac{1}{2\pi}, 0.1956\right). \tag{7}$$

9

By enforcing this constraint, we guarantee that the formlet $f(z, \zeta, \sigma, \alpha)$ is a diffeomorphism of the plane. Hence, such a formlet acting on a curve embedded in the plane will be a homeomorphism. In particular, let Γ be the continuous mapping

$$\Gamma : [0, 1] \to \mathbb{C}. \tag{8}$$

Recall that Γ is simple if the mapping is injective, and closed by permitting the equality $\Gamma(0) = \Gamma(1)$. Since a formlet f satisfying Equation 7 is bicontinuous, if Γ is simple and closed, the deformed curve

$$\Gamma^f(t) = f(\Gamma(t)) \tag{9}$$

will also be simple and closed.

Figures 4(a) and (b) show the radial deformation function $\rho(r; \sigma, \alpha)$ as a function of r for a range of gain α and scale σ values respectively. Figures 4(c) and (d) show the corresponding trace of the formlet deformation of an ellipse in the plane.

(a) ρ with gain variation

(b) ρ with scale variation

(c) f with gain variation
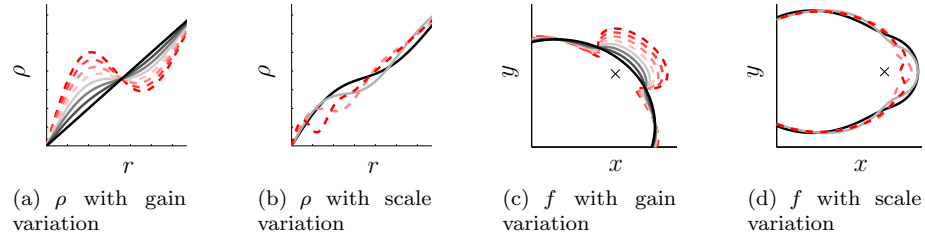
(d) f with scale variation

Figure 4: Formlet transformations as a function of scale and gain. Dashed lines denote invalid formlet parameters outside the diffeomorphism bounds of Equation 7.

3.3. Formlet Composition

The power of formlets is that they can be composed to produce complex shapes while preserving topology. We define the *forward formlet composition* problem as follows. Given an embryonic shape $\Gamma^0(t)$ and a sequence of K formlets $\{f_1 \ldots f_K\}$ drawn from a formlet dictionary $\mathcal{D}$, determine the resulting deformed shape $\Gamma^K(t)$. The problem is well-posed because the set of simple closed curves is closed under formlet deformation: multiple formlets can be composed to generate complex shape transformations. Thus,

$$\Gamma^K(t) = (f_K \circ f_{K-1} \circ \cdots \circ f_1)(\Gamma^0(t)). \tag{10}$$

Figure 5 shows an example of forward composition from a circular embryonic shape, where the formlet parameters ζ, σ, and α have been randomly selected. Note that a rich set of complex shapes is generated without leaving the space of valid shapes (simple, closed contours).

A more difficult but interesting problem is inverse formlet composition: given an *observed* shape $\Gamma^{obs}(t)$, determine a sequence of K formlets $\{f_1 \ldots f_K\}$, drawn
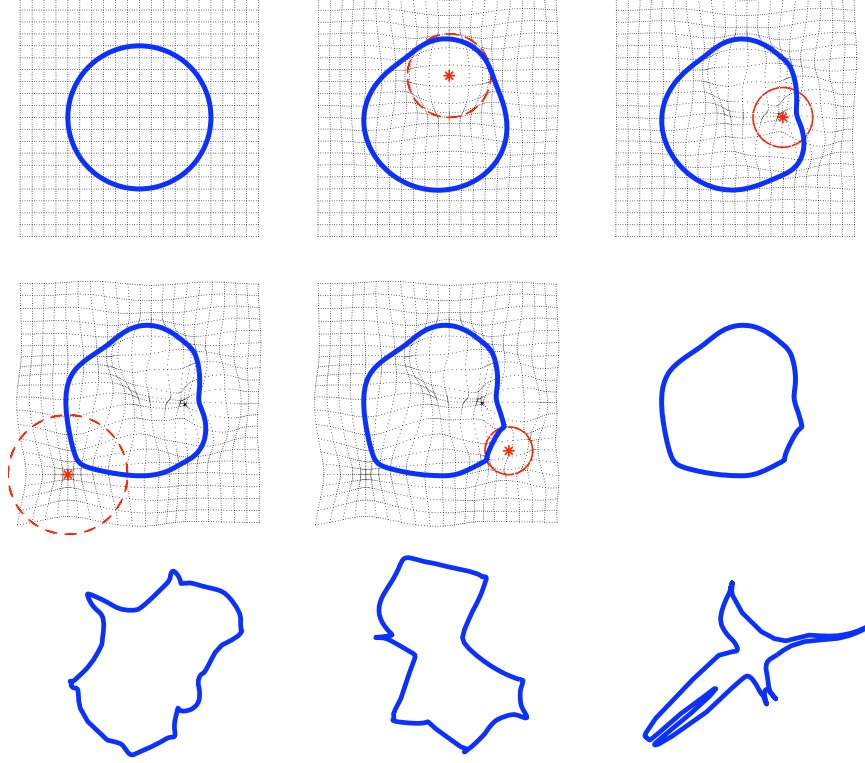
Figure 5: Shapes generated by random formlet composition over the unit circle. The first two rows show the result of applying 5 successive random formlets. The asterisk and circle indicate formlet location ζ and scale σ, respectively. The bottom row shows some example shapes produced from the composition of many random formlets.

from a formlet dictionary $\mathcal{D}$, that best approximate $\Gamma^{obs}(t)$ by minimizing some reconstruction error ξ. Here we measure error as the L^2 norm of the residual:[1]

$$
\begin{aligned}
\xi(\Gamma^{obs}, \Gamma^K) &= \left|\left|\Gamma^{obs}(t) - \Gamma^K(t)\right|\right|_2^2 \\
&= \int_0^1 \left(\Gamma^{obs}(t) - \Gamma^K(t)\right) \overline{\left(\Gamma^{obs}(t) - \Gamma^K(t)\right)} dt.
\end{aligned}
\tag{11}
$$

---

[1]For notational simplicity, we treat contours as continuous functions of arc length t. In practice, we represent contours as 128-point vectors. All integrals map to summations in a straightforward manner.

4. Formlet Pursuit

4.1. Dictionary Method

As a first attempt to estimate the optimal formlet sequence $\{f_1 \ldots f_K\}$, we propose a version of matching pursuit for sparse approximation [25], replacing the linear summation of elements by a non-commutative composition of formlet components. Algorithm 1 shows the flow of the formlet pursuit algorithm.

Algorithm 1: Formlet Pursuit of Γ^{obs}.

Initialization: define $\Gamma^0 = A\Gamma_0 + z_0$ to be a best matching ellipse approximating Γ^{obs}

for $k = 1, \ldots, K$ **do**

　Optimal Formlet: compute maximal error reducing transformation

$$f_k = \operatorname*{argmin}_{f \in \mathcal{D}} \xi(\Gamma^{obs}, f(\Gamma^{k-1}))$$

　Update Approximation: apply optimal formlet

$$\Gamma^k = f_k(\Gamma^{k-1})$$

Initialization. Given an observed target shape Γ^{obs}, we initialize the model as a 128-point polygon sampling the unit circle, and form a 1:1 correspondence between the model and target points that remains fixed throughout pursuit. We next apply an affine transformation to the model to generate an embryonic elliptical shape Γ^0 minimizing the L^2 error $\xi\left(\Gamma^{obs}, f(\Gamma^0)\right)$.

Formlet Selection. At iteration k of the formlet pursuit algorithm, we select the formlet $f_k(z; \zeta_k, \sigma_k, \alpha_k)$ that, when applied to the current model Γ^{k-1}, maximally reduces the approximation error:

$$f_k = \operatorname*{argmin}_{f \in \mathcal{D}} \xi(\Gamma^{obs}, f(\Gamma^{k-1})). \tag{12}$$

This is a difficult non-convex optimization problem, and experimentation with gradient descent methods has shown that the formlet parameter space can have many local minima. One saving grace is that the formlet transformation is linear with respect to the gain α, allowing α to be recovered analytically. Specifically, consider an alternative but equivalent representation of the formlet described by Equation 3:

$$f(z; \zeta, \sigma, \alpha) = z + \alpha \cdot g(z - \zeta, \sigma) \quad \text{where}$$
$$g(z_\zeta; \sigma) = \frac{z_\zeta}{|z_\zeta|} \sin\left(\frac{2\pi|z_\zeta|}{\sigma}\right) \exp\left(-\frac{|z_\zeta|^2}{\sigma^2}\right). \tag{13}$$

In Appendix Appendix A we show that, if we fix both the formlet location ζ and scale σ, the optimal unconstrained gain α^* for formlet f_k is given by

12

$$\alpha^* = \frac{\left\langle \Gamma^{obs} - \Gamma^{k-1}, g\left(\Gamma^{k-1} - \zeta; \sigma\right)\right\rangle dt}{\left\|g\left(\Gamma^{k-1} - \zeta; \sigma\right)\right\|_2^2}. \tag{14}$$

where $\langle \cdot, \cdot \rangle$ is the inner product on functions $f : [0,1] \to \mathbb{C}$ given in equation A.3.

One complication is that Equation 14 may yield a gain value α^* that does not satisfy the diffeomorphism constraint given by Equation 7. However, from Equations 3 and 11 it can be seen that the error is a quadratic function of the gain α. Thus the optimal constrained gain α_c^* for given ζ and σ parameters is simply the optimal unconstrained gain α^* expressed by Equation 14, thresholded by the diffeormorphism constraints:

$$\alpha_c^* = \begin{cases} \alpha_l & \text{for } \alpha^* < \alpha_l \\ \alpha^* & \text{for } \alpha_l \le \alpha^* \le \alpha_u \\ \alpha_u & \text{for } \alpha^* > \alpha_u, \end{cases} \tag{15}$$

where

$$\alpha_l = -(2\pi)^{-1}\sigma$$

and

$$\alpha_u \approx 0.1956\sigma.$$

Thus search for the optimal formlet can proceed by sampling from a dictionary over location ζ and scale σ parameters, computing the optimal constrained gain α_c^* in each case, and then selecting the resulting formlet that yields minimum error.

Figure 6 shows an example of formlet pursuit with this dictionary on an example animal shape.
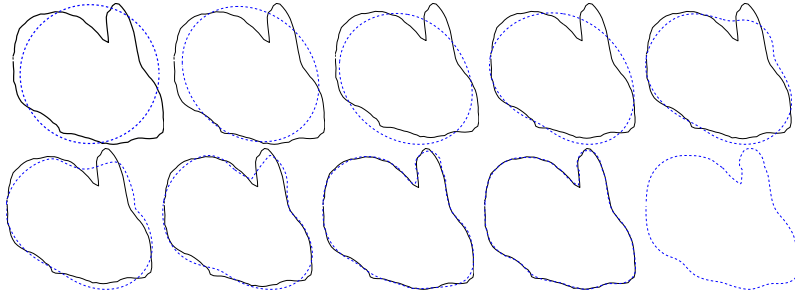

Figure 6: Formlet pursuit of an example animal shape. We first show the initial unit circle, followed by the least-squares ellipse embryo $\Gamma^0(t)$, and the models Γ^k, where $k = 1, 2, 3, 4, 8, 16, 32$. The last curve shows the model Γ^{32} without the target curve Γ^{obs}.

4.2. Dictionary Descent Method

While the formlet pursuit method has the advantage of simplicity, it is far from optimal, as it ignores most smoothness properties that the error function might enjoy, aside from the quadratic dependence upon the gain α. As a consequence one must face the tradeoff between accuracy, which requires that the parameter space be sampled finely, and speed, which limits the capacity of the dictionary.

We can potentially improve upon the standard dictionary method by employing a smaller dictionary, and initiating a local gradient descent search from the m most promising formlets to determine the formlet parameters that locally minimize the error function.

Figure 7 compares pursuit for the standard dictionary and dictionary descent methods on a particular example animal shape: the higher accuracy of the dictionary descent method is evident. Table 1 shows the performance of the two methods on the entire shape dataset. The dictionary descent method improves accuracy by roughly 11%, and runs about 42% faster than standard pursuit. We use the dictionary descent method in our evaluation below. An implementation is available at www.elderlab.yorku.ca/formlets.
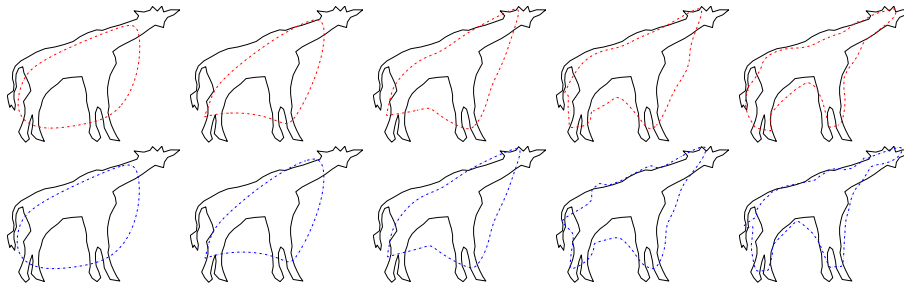

Figure 7: Pursuit of an example animal shape with standard dictionary search (top row) and dictionary descent (bottom row) for K=1,2,4,8,16.

Table 1: Comparison of Dictionary and Dictionary Descent methods on entire animal dataset.

Optimization Method	L^2 Error	Run Time (min)
Dictionary	0.00535	1.9
Dictionary Descent	0.00476	1.1

5. Implementation Details

5.1. Shape Dataset

To explore the inverse problem of constructing formlet representations of planar shapes, we employ a database consisting of 391 blue-screened images of

animal models from the Hemera Photo-Object database. The boundary of each object was sampled at 128 points at regular arc-length intervals. Each resulting polygon was then shifted to have 0 mean and scaled to have unit L^2 norm in both vertical and horizontal directions:

$$\int_0^1 \mathbf{Re}(\Gamma^{obs}(t))^2 dt = \int_0^1 \mathbf{Im}(\Gamma^{obs}(t))^2 dt = 1. \tag{16}$$

This scaling generally alters the aspect ratio of the shape: we invert this distortion when displaying our results. The full dataset of object shapes used in this paper is available at www.elderlab.yorku.ca/formlets.

5.2. Dictionary Method: Discretization

To evaluate this formlet pursuit algorithm, we constructed a dictionary consisting of a regular sampling of the position parameter ζ on a 64×64 grid roughly 4 times the extent of the average shape, and the scale parameter σ at 16 regularly-spaced values over $(0, 0.8]$.

5.2.1. Tuning the Dictionary Descent Method

Since our objective function is the L^2-norm of the residual error between the observed curve and the approximation, we employed the MATLAB function $lsqnonlin()$, which is optimized for non-linear least squares problems, and compute the Jacobian of the objective function analytically (Appendix Appendix B). We tuned the parameters of our Dictionary Descent method in stepwise fashion. First, we determined appropriate values for the tolerance parameters $xTol$ and $fTol$ of $lsqnonlin()$, which determine the stopping criteria for the parameters and error function, respectively. We employed a sparse dictionary, sampling the position parameter ζ on a 16×16 grid, and the scale parameter σ at 4 regularly-spaced values over $(0, 0.8]$. We initiated descent at the $m = 100$ lowest error solutions. Using a small subset of our animal dataset containing only four animal shapes, we performed a grid search in log space over the $xTol$ and $fTol$ parameters in the range 10^{-1} to 10^{-9}, computing the average running time and L^2 error for a 32-formlet approximation. All experiments were conducted on a Power Mac G5 with a 2.66 Ghz quad-core Intel Xeon processor, running MATLAB R2009b.

The results are shown in Table 2. Error was found to be minimized for parameter values of $xTol = 10^{-3}, fTol = 10^{-6}$: we used these values for all further experiments.

Second, we optimized the density of the dictionary and number m of dictionary formlets selected for descent, using the descent parameters optimized above, the same 4 training shapes, and 32-formlet approximation. The running time and accuracy results are shown in Tables 3 and 4 respectively. Sampling ζ on a 51×51 grid, the scale parameter σ at 13 values, and launching $m = 25$ descents from the most promising formlets, we found that for these four training images we could improve the accuracy over the standard dictionary method by a factor of more than two, while saving roughly 30% in computation time.

15

Table 2: Average L^2 error ($\times 100$) for a 32-formlet model, as a function of the gradient descent termination criteria.

fTol/xTol	1E-01	1E-02	1E-03	1E-04	1E-05	1E-06	1E-07	1E-08	1E-09
1E-01	6.86	6.31	6.54	6.54	6.54	6.54	6.54	6.54	6.54
1E-02	5.68	5.26	5.11	5.02	5.02	5.02	5.02	5.02	5.02
1E-03	5.91	4.66	4.04	4.16	4.16	4.16	4.16	4.16	4.16
1E-04	5.80	3.65	4.02	3.72	3.72	3.72	3.72	3.72	3.72
1E-05	5.75	3.69	3.66	3.73	3.90	3.90	3.90	3.90	3.90
1E-06	5.75	3.69	3.52	3.73	3.74	3.74	3.74	3.74	3.74
1E-07	5.75	3.69	3.54	3.71	3.66	3.66	3.66	3.66	3.66
1E-08	5.75	3.69	3.54	3.67	3.66	3.66	3.66	3.66	3.66
1E-09	5.75	3.69	3.54	3.67	3.66	3.66	3.66	3.66	3.66

Table 3: Average running time per shape (min) for a 32-formlet model, as a function of dictionary size n and number of descents m.

m/n	$64^2 \times 16$	$58^2 \times 14$	$51^2 \times 13$	$45^2 \times 11$	$38^2 \times 10$	$32^2 \times 8$	$26^2 \times 6$	$19^2 \times 5$
0	1.28	0.93	0.68	0.45	0.29	0.16	0.09	0.04
1	1.38	1.01	0.74	0.50	0.33	0.21	0.11	0.07
5	1.44	1.06	0.78	0.55	0.38	0.24	0.18	0.13
10	1.49	1.10	0.84	0.60	0.44	0.31	0.25	0.20
15	1.56	1.17	0.90	0.67	0.50	0.37	0.32	0.28
20	1.60	1.23	0.95	0.73	0.56	0.42	0.41	0.35
25	1.67	1.28	1.01	0.78	0.62	0.48	0.47	0.42
30	1.72	1.34	1.07	0.84	0.68	0.55	0.53	0.51

Table 4: Average residual ($\times 1000$) for a 32-formlet model, as a function of dictionary size n and number of descents m.

m/n	$64^2 \times 16$	$58^2 \times 14$	$51^2 \times 13$	$45^2 \times 11$	$38^2 \times 10$	$32^2 \times 8$	$26^2 \times 6$	$19^2 \times 5$
0	8.0	6.5	8.1	9.5	9.3	14.0	23.5	28.1
1	3.7	4.7	4.8	5.8	5.0	5.9	7.4	14.0
5	3.6	3.8	4.7	5.9	4.2	6.4	7.4	7.1
10	4.1	3.7	3.8	45	4.5	6.1	6.9	6.4
15	3.6	3.9	4.1	4.0	4.1	5.7	6.7	8.0
20	3.4	3.8	3.8	4.5	4.1	4.8	6.6	8.0
25	3.7	3.8	3.7	4.5	4.2	4.8	6.5	7.4
30	3.4	3.9	3.7	4.4	4.2	4.4	5.7	7.5

Interestingly, we found that tightening tolerance parameters, increasing the dictionary density, or increasing the number of deployments of the optimizer did not always decrease the error. However, at a given iteration, error did decrease monotonically as a function of each of these parameters, as expected. Thus the non-monotonic variation in error with these parameters appears to reflect the non-optimality of the greedy pursuit algorithm. In other words, selecting the formlet that minimizes the residual at stage i will not necessarily lead to the smallest error at stage $k > i$.

6. Evaluation

To evaluate and compare shape models, we address the problem of *contour completion*, using our animal shape dataset. In natural scenes, object boundaries are often fragmented by occlusion and loss of contrast: contour completion is the process of filling in the missing parts. Completion can also be an important component of perceptual organization algorithms: given one or more partial contour hypotheses, completion can be used to estimate the locations of missing parts. These estimate can then guide search for corroborating evidence.

We compare our formlet model with the shapelet model described in Section 2.1 [9]. For each shape in the dataset, we simulate the occlusion of a 10% or 30% continuous section of the contour, and allow the two methods to pursue only the remaining visible portion.

The rate of convergence of both formlet and shapelet methods depends upon how the parameters are sampled. For formlet pursuit, we use the dictionary descent method described in Section 4.2. For the shapelet method, we used the standard dictionary method of Dubinskiy et al. [9], optimizing performance by sampling as finely as possible given time constraints. The shapelet representation assumes an arc-length representation of the curves on $t \in [0,1]$, and each shapelet component has an arc-length position μ and scale σ. We sampled the position parameter μ at 128 regularly-spaced values over $[0,1]$, and the scale parameter σ at 128 regularly-spaced values over $(0,1]$. The affine parameters were computed analytically [9].

The formlet and shapelet pursuit algorithms were initialized with the same embryonic ellipses, and were governed by a minimization of the L^2 error (Equation 11) over the visible points of the curves only. While pursuit is based on a fixed 1:1 correspondence between points on the target and model curves, we measure performance using the L^2 Hausdorff distance to avoid potential dependence of the evaluation upon the parameterization of the curves. Specifically, we define the error between the target shape and the model as the average minimum distance of a point on one of the shapes to the other shape:

$$\xi^H(\Gamma^{obs}, \Gamma^k) = \sqrt{\int_0^1 \frac{1}{2} \left(\min_{t' \in [0,1)} |\Gamma^{obs}(t) - \Gamma^k(t')|^2 + \min_{t' \in [0,1)} |\Gamma^{obs}(t') - \Gamma^k(t)|^2 \right) dt}.$$
(17)

We measured the residual error between the model and target for both the visible and occluded portions of the shapes. Performance on the occluded portion, where the model is under-constrained by the data, reveals how well the structure of the model captures properties of natural shapes.

Implementations for both the formlet and shapelet models are available at www.elderlab.yorku.ca/formlets.

6.1. Results

Figure 8 shows some example qualitative results for this experiment. While shapelet pursuit introduces topological errors in both visible and occluded regions, formlet pursuit remains topologically valid, as predicted.
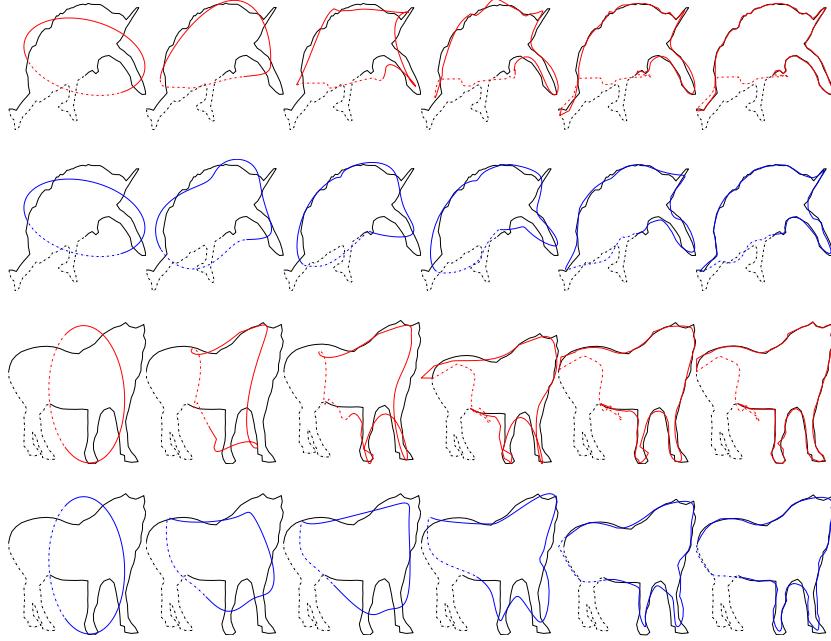

Figure 8: Examples of 30% occlusion pursuit with shapelets (red) and formlets (blue) for $k = 0, 2, 4, 8, 16, 32$. Solid lines indicate visible contour, dashed lines indicate occluded contour.

Figure 9 shows quantitative results for this experiment. While the shapelet and formlet models achieve comparable error on the visible portions of the boundaries, the error is substantially lower for the formlet representation on the occluded portions. This suggests that the structure of the formlet model better captures regularities in the shapes of natural objects. We believe that the two principal reasons for this are a) respecting the topology of the shape prunes off many inferior completion solutions and b) by working in the image space, rather than arc length, the formlet model is better able to capture important regional properties of shape.
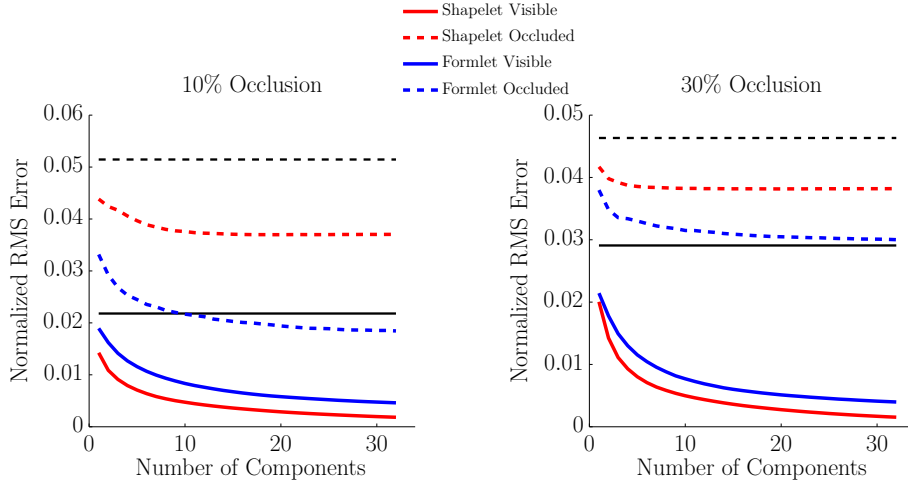
18

Figure 9: Results of occlusion pursuit evaluation. Black denotes error for $\Gamma^0(t)$, the affine-fit ellipse.

7. Discussion

7.1. Formlet Parameter Distributions

The focus of this paper is to establish the appropriate structural properties for a generative model of planar shape. To ultimately apply this representation to problems such as object detection and recognition, statistical models over this representation must be developed. One small step is to consider the distribution of formlet parameters selected in pursuit of the shapes in our animal dataset.

Figure 10 shows how the means of the formlet parameters vary as pursuit unfolds. We observe that scales decrease over time (a), reflecting a coarse-to-fine approximation. Gains also decrease over time (b), although when normalized by scale (c), this decline is moderated substantially. Finally, formlet locations are biased to the centre of the shape and are roughly isotropic (d), with a slight bias to the lower field, presumably reflecting the additional details required to represent the legs of the animals.

7.2. Alternative Formlet Bases

In this paper we have chosen a particular Gabor-like formlet representation (Equation 3) that confers several key properties:

1. The family of formlets forms a self-similar scale space.
2. Each formlet acts within a σ-ball around a specific location ζ, converging to the identity as $|z - \zeta| \to \infty$.
3. The mapping is smooth everywhere except at ζ, where it is C^0.
4. Deformation is isotropic and radial around ζ.

There are of course other formulations that would also satisfy these properties. Here we consider two specific alternatives and compare them with the Gabor formulation.
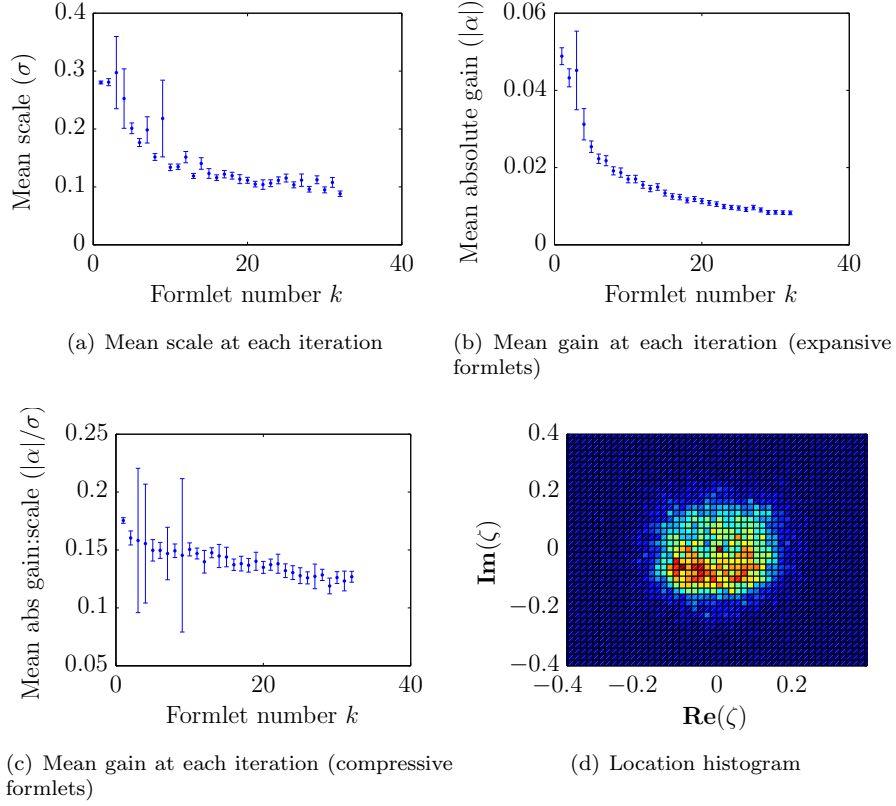
19

(a) Mean scale at each iteration

(b) Mean gain at each iteration (expansive formlets)

(c) Mean gain at each iteration (compressive formlets)

(d) Location histogram

Figure 10: Marginal distributions of formlet parameters. Error bars indicate standard error of the mean.

7.3. Gaussian Derivative Formlets

We simplify the original Gabor formulation of Equation 3 by replacing the sinusoidal factor with a first-order Taylor series approximation, yielding:

$$f(z; \zeta, \sigma, \alpha) = \zeta + \frac{z - \zeta}{|z - \zeta|} \rho(|z - \zeta|), \quad \text{where} \tag{18}$$

$$\rho(r) = r + \alpha \frac{2\pi r}{\sigma} \exp\left(\frac{-r^2}{\sigma^2}\right) \tag{19}$$

Note that the deformation term of the radial deformation function $\rho(r)$ is proportional to the first Gaussian derivative in r.

f is a diffeomorphism iff $\rho'(r) > 0$ everywhere:

$$\rho'(r) = 1 + \exp\left(\frac{-r^2}{\sigma^2}\right) \frac{2\pi\alpha}{\sigma} \left[1 - \frac{2r^2}{\sigma^2}\right] > 0. \tag{20}$$

20

For $\alpha < 0$, the minimum is attained when $r = 0$:

$$\Rightarrow \alpha > -\frac{1}{2\pi}\sigma \tag{21}$$

For $\alpha > 0$, by solving $\rho''(r) = 0$ it can be shown that the minimum is attained when $r = \sqrt{3/2}\sigma$. Substituting into Equation 20 then yields

$$\alpha < \frac{\exp(3/2)}{4\pi}\sigma \tag{22}$$

Thus f is a diffeomorphism iff $\alpha \in \dfrac{\sigma}{2\pi}\left(-1, \dfrac{1}{2}\exp(3/2)\right)$.

7.4. Spline Formlets

Both the Gabor and Gaussian formlets have infinite support, which increases computation time and limits the degree to which formlets can be computed in parallel. To achieve strictly compact support we impose the constraint that $\rho(r;\sigma) = r \iff f(z) = z$ whenever $r > \sigma$. To guarantee smoothness, we require $\rho(\sigma;\sigma) = \sigma$ and $\rho'(\sigma;\sigma) = 1$ and to achieve continuity at ζ we require $\rho(0) = 0$. The simplest spline meeting all these conditions is:

$$\rho(r;\sigma) = \begin{cases} r + \alpha \dfrac{r}{\sigma^2}(r - \sigma)^2 & \text{for } r \leq \sigma \\ r & \text{for } r > \sigma \end{cases} \tag{23}$$

We derive the diffeomorphism constraints as before:

$$\rho'(r) = 1 + \frac{\alpha}{\sigma^2}\left[(r - \sigma)^2 + r \cdot 2(r - \sigma)\right] > 0 \tag{24}$$

$$\Rightarrow \frac{\alpha}{\sigma^2}\left[3r^2 - 4r\sigma + \sigma^2\right] > -1 \tag{25}$$

For $\alpha < 0$, the minimum is attained when $r = 0$, yielding $\alpha > -1$.

For $\alpha > 0$, by solving $\rho''(r) = 0$ it can be shown that the minimum is attained when $r = 2\sigma/3$. Substituting into Equation 24 then yields $\alpha < 3$. Thus f is a diffeomorphism iff $\alpha \in (-1, 3)$

7.5. Comparison of Formlet Bases

Figures 11 - 12(b) show the radial deformation functions, examples of pursuit and rate of convergence for these three different formulations. Empirically, we find that the Gabor formulation achieves a better rate of convergence on the animal dataset than the competing formulations, although at this stage we do not have a clear theoretical explanation for this result.
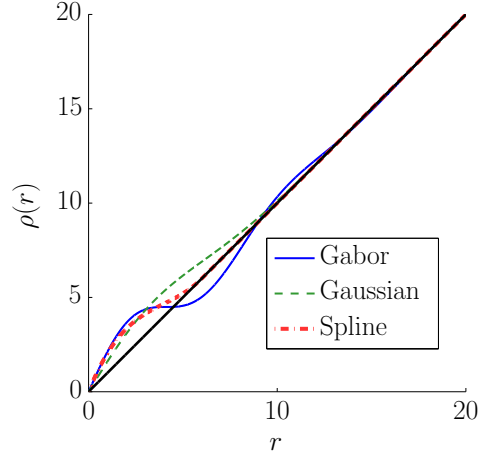
Figure 11: Radial deformation function for three formlet bases.

8. Conclusion

We have developed a novel generative model of planar shape that satisfies a number of essential properties. In this model, complex shapes are seen as the evolution of a simple embryonic shape by successive application of simple diffeomorphic transformations of the plane called *formlets*. The system is both complete and closed, since arbitrary shapes can be modeled, and generated shapes are guaranteed to be topologically valid. This means that the model has the potential to support accurate probabilistic modeling. We have demonstrated a novel dictionary descent formlet pursuit algorithm that selects formlets to efficiently approximate given target shapes. Evaluation of the formlet pursuit model on the problem of shape completion revealed that the model is better able to approximate parts of shapes missing due to occlusion than a competing contour-based method. Our animal object dataset, experimental results, example movies and implementations for both the formlet and shapelet models are available at www.elderlab.yorku.ca/formlets.

Future Work. We hope to extend the present work in a number of ways. First, we would like to generalize our definition of formlets to allow for anisotropic deformation that could efficiently model elongated parts such as animal limbs. Second, we would like to develop probabilistic models over the formlet representation. Finally, we are interested in using the formlet pursuit algorithm for contour grouping, using detected fragments to generate predictions for where other fragments of the same object boundary might be found.

22
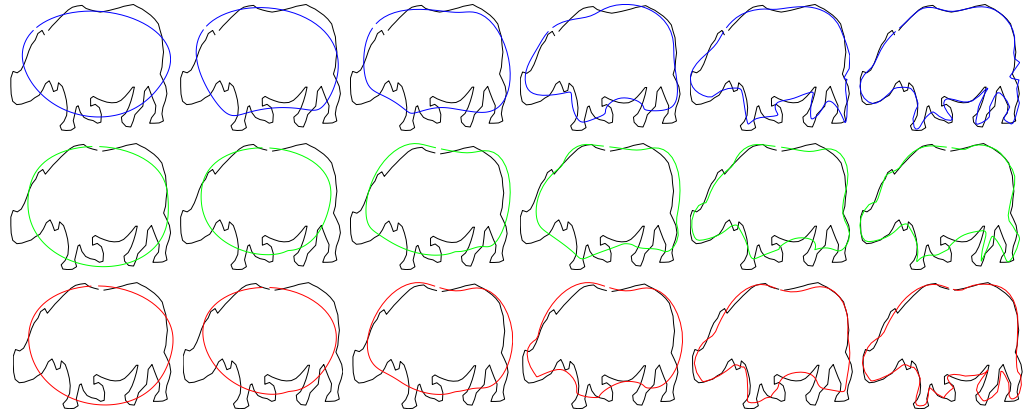
(a) Pursuit of an example shape with Gabor (blue), Gaussian (green) and spline (red) bases for K=1,2,4,8,16.
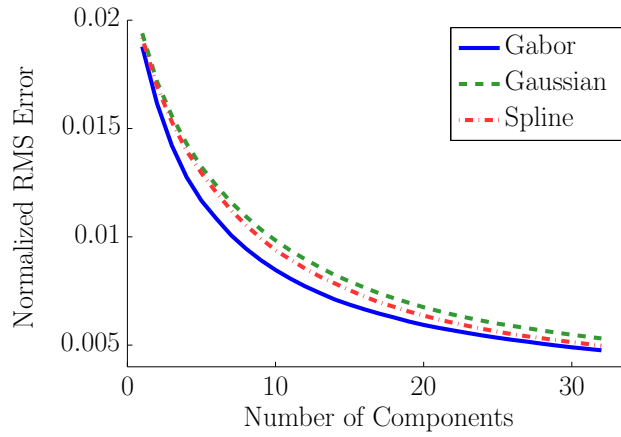
(b) Mean L^2 Hausdorff error for formlet pursuit over animal dataset with three different formlet bases.

Figure 12: Comparison of the three different formlet bases.

References

[1] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, Pattern Analysis and Machine Intelligence, IEEE Trans. 24 (2002) 509–522.

[2] P. Cavanagh, What's up in top-down processing, in: A. Gorea (Ed.), Representations of Vision, Cambridge University Press, Cambridge, UK, 1991 edition, 1991, pp. 295–304.

[3] D. Mumford, Mathematical theories of shape: do they model perception?, in: B. C. Vemuri (Ed.), Geometric Methods in Computer Vision, volume 1570, SPIE, 1991, pp. 2–10.

[4] F. Attneave, Some informational aspects of visual perception, Psychol. Rev. 61 (1954) 183–193.

[5] T. F. Cootes, C. J. Taylor, D. H. Cooper, J. Graham, Active shape models - their training and application, Comput. Vis. Image Underst. 61 (1995) 38–59.

[6] T. Pavlidis, Structural pattern recognition, volume 1, Springer-Verlag, Berlin, illustrated edition, 1977.

[7] D. D. Hoffman, W. A. Richards, Parts of recognition, Cognition 18 (1984) 65–96.

[8] F. Mokhtarian, A. Mackworth, Scale-based description and recognition of planar curves and two-dimensional shapes, Pattern Analysis and Machine Intelligence, IEEE Trans. 8 (1986) 34–43.

[9] A. Dubinskiy, S. Zhu, A multiscale generative model for animate shapes and parts, in: Proc. 9th IEEE ICCV, volume 1, pp. 249–256.

[10] H. Blum, Biological shape and visual science (part i), J. Theoretical Biology 38 (1973) 205–287.

[11] H. Blum, R. N. Nagel, Shape description using weighted symmetric axis features, Pattern Recognition 10 (1978) 167 – 180.

[12] M. Brady, H. Asada, Smoothed local symmetries and their implementation, Int. J. Robotics Res. 3 (1984) 36–61.

[13] M. Leyton, A process-grammar for shape, Artificial Intelligence 34 (1988) 213 – 247.

[14] B. B. Kimia, A. R. Tannenbaum, S. W. Zucker, Shapes shocks and deformations i: the components of two dimensional shape and the reaction diffusion space, Int. J. Comput. Vision 15 (1995) 189–224.

[15] S. Osher, J. A. Sethian, Fronts propagating with curvature-dependent speed, J. Comput. Phys. 79 (1988) 12–49.

[16] N. Trinh, B. Kimia, A symmetry-based generative model for shape, in: Proc. 11th IEEE ICCV, pp. 1–8.

[17] A. Pentland, S. Sclaroff, Closed-form solutions for physically based shape modeling and recognition, Pattern Analysis and Machine Intelligence, IEEE Transactions on 13 (1991) 715–729.

[18] S. Scarloff, A. Pentland, Modal matching for correspondence and recognition., Pattern Analysis and Machine Intelligence, IEEE Trans. 17 (1995) 545–561.

[19] S.-C. Zhu, Embedding gestalt laws in markov random fields, Pattern Analysis and Machine Intelligence, IEEE Trans. 21 (1999) 1170–1187.

[20] D. W. Thompson, On growth and form, Cambridge University Press, Cambridge, UK, abridged ed./edited edition, 1961.

[21] A. Jain, Y. Zhong, S. Lakshmanan, Object matching using deformable templates, Pattern Analysis and Machine Intelligence, IEEE Trans. 18 (1996) 267–278.

[22] E. Sharon, D. Mumford, 2d-shape analysis using conformal mapping, Computer Vision and Pattern Recognition, IEEE Comp. Soc. Conf. 2 (2004) 350–357.

[23] U. Grenander, A. Srivastava, S. Saini, A pattern-theoretic characterization of biological growth, Medical Imaging, IEEE Trans. 26 (2007) 648–659.

[24] T. Oleskiw, J. Elder, G. Peyré, On growth and formlets, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2010).

[25] S. Mallat, Z. Zhang, Matching pursuits with time frequency dictionaries, Signal Processing, IEEE Trans. 41 (1993) 3397–3415.

Appendix A. Computation of Optimal Gain

Since the formlet deformation of Equation 3 is linear in the gain α, given fixed location ζ and scale σ parameters, the gain that minimizes the L^2 deviation from the target shape can be computed analytically. Specifically, suppose that the observed curve Γ^{obs} is currently approximated by Γ^{k-1}. For given formlet location and scale parameters ζ and σ, we define the optimal unconstrained gain α^* for formlet f_k as:

$$\alpha^* = \underset{\alpha \in \mathbb{R}}{\operatorname{argmin}} \, \xi(\Gamma^{obs}, f(\Gamma^{k-1}; \zeta, \sigma, \alpha)) \tag{A.1}$$

where, for curves a and b, $\xi(\Gamma^a, \Gamma^b)$ denotes the L^2 error metric

$$\int_0^1 \mathbf{Re}\left(\Gamma^a(t) - \Gamma^b(t)\right)^2 + \mathbf{Im}\left(\Gamma^a(t) - \Gamma^b(t)\right)^2 dt. \qquad (A.2)$$

induced by the inner product:

$$\left\langle \Gamma^a, \Gamma^b \right\rangle = \int_0^1 \mathbf{Re}\,\Gamma^a(t)\,\mathbf{Re}\,\Gamma^b(t) + \mathbf{Im}\,\Gamma^a(t)\,\mathbf{Im}\,\Gamma^b(t)\,dt \qquad (A.3)$$

Using Equation 13, we differentiate ξ with respect to α and set to zero:

$$\frac{\partial}{\partial \alpha}\|\Gamma^{obs} - f(\Gamma^k - 1)\|^2 = \frac{\partial}{\partial \alpha}\|\Gamma^{res} - \alpha g\|^2 \qquad (A.4)$$

$$= \frac{\partial}{\partial \alpha}\left(\|\Gamma^{res}\|^2 - 2\alpha\left\langle\Gamma^{res}, g\right\rangle + \alpha^2\|g\|^2\right) \qquad (A.5)$$

$$= 2(\left\langle\Gamma^{res}, g\right\rangle - \alpha\|g\|^2) = 0 \qquad (A.6)$$

$$\Rightarrow \alpha = \frac{\left\langle\Gamma^{res}, g\right\rangle}{\|g\|^2}$$

where we used the shorthand $g = g\left(\Gamma^{k-1} - \zeta; \sigma\right), \Gamma^{res} = \Gamma^{obs} - \Gamma^{k-1}$.

As a result, given fixed ζ and σ, the optimal unconstrained gain α^* that maximally reduces the L^2 error between the observed curve Γ^{obs} and current approximation Γ^{k-1} is given by

$$\alpha^* = \frac{\left\langle\Gamma^{obs} - \Gamma^{k-1}, g\left(\Gamma^{k-1} - \zeta; \sigma\right)\right\rangle dt}{\left\|g\left(\Gamma^{k-1} - \zeta; \sigma\right)\right\|_2^2}. \qquad (A.7)$$

Note that in general Equation A.7 may produce an optimal gain outside the diffeomorphism bounds of Equation 7. However, the optimal gain that satisfies the constraint is simply the unconstrained gain α^* thresholded by the diffeomorphism constraints, as described in Section 4.1.

Appendix B. Jacobian Computation for Nonlinear Least Squares Minimization

The dictionary descent optimization method described in Section 4.2 employs the MATLAB gradient descent method *lsqnonlin* to determine the location parameter ζ and scale parameter σ. *lsqnonlin* uses the Jacobian of the error function in the unknown parameters to iterate toward the local minimum. The method performs best if an analytic form of the Jacobian can be provided. Note that since the optimal gain α_c^* is determined analytically (Equation 15), this value must be used in all computations of the Jacobian in order to determine locally optimally values for the other parameters.

26

Combining Equations 11 and 13, and using $r = |\Gamma^{k-1} - \zeta|$, the error function can be written as

$$
\begin{aligned}
\xi(\Gamma^{obs}, \Gamma^K) &= \|(\Gamma^{obs} - f(\Gamma^{k-1})\|_2 \\
&= \left\| \Gamma^{obs} - \Gamma^{k-1} - \alpha_c^* \frac{\Gamma^{k-1} - \zeta}{r} \sin\left(\frac{2\pi r}{\sigma}\right) \exp\left(\frac{-r^2}{\sigma^2}\right) \right\|_2.
\end{aligned}
$$

Now defining

$$
\begin{aligned}
\Gamma^{res} &= \Gamma^{obs} - \Gamma^{k-1} \\
\text{and} \\
G &= G(r; \sigma) = \frac{1}{r} \sin\left(\frac{2\pi r}{\sigma}\right) \exp\left(\frac{-r^2}{\sigma^2}\right),
\end{aligned}
$$

and using x and y subscripts to denote real and imaginary components, we can rewrite this expression as

$$
\begin{aligned}
\xi(\Gamma^{obs}, \Gamma^K) &= \int_0^1 \left[(\Gamma_x^{res}(t)) - \alpha_c^*(\Gamma_x^{k-1}(t) - \zeta_x)G \right]^2 dt \\
&+ \int_0^1 \left[(\Gamma_y^{res}(t)) - \alpha_c^*(\Gamma_y^{k-1}(t) - \zeta_y)G \right]^2 dt \\
&\equiv \int_0^1 \left(\xi_x(t)^2 + \xi_y(t)^2 \right) dt.
\end{aligned}
$$

Since the error is a function of the optimal gain α_c^* and α_c^* is a function of the location parameter ζ and the scale parameter σ, we will need the partial derivative of α_c^* with respect to these two parameters. From Equation 15, we have

$$
\alpha_c^* = \begin{cases} \alpha_l & \text{for } \alpha^* < \alpha_l \\ \alpha^* & \text{for } \alpha_l \leq \alpha^* \leq \alpha_u \\ \alpha_u & \text{for } \alpha^* > \alpha_u, \end{cases}
$$

where

$$
\begin{aligned}
\alpha_l &= -(2\pi)^{-1}\sigma \\
\text{and} \\
\alpha_u &\approx 0.1956\sigma,
\end{aligned}
$$

and α^* is given by Equation 14. Thus we have

$$
\frac{\partial \alpha_c^*}{\partial \sigma} = \begin{cases} -(2\pi)^{-1} & \text{for } \alpha^* < \alpha_l \\ \dfrac{\partial \alpha^*}{\partial \sigma} & \text{for } \alpha_l \leq \alpha_* \leq \alpha_u \\ 0.1956 & \text{for } \alpha^* > \alpha_u \end{cases}
$$

$$\frac{\partial \alpha_c^*}{\partial \zeta_x} = \begin{cases} 0 & \text{for } \alpha^* < \alpha_l \\ \dfrac{\partial \alpha^*}{\partial \zeta_x} & \text{for } \alpha_l \leq \alpha^* \leq \alpha_u \\ 0 & \text{for } \alpha^* > \alpha_u \end{cases}$$

$$\frac{\partial \alpha_c^*}{\partial \zeta_y} = \begin{cases} 0 & \text{for } \alpha^* < \alpha_l \\ \dfrac{\partial \alpha^*}{\partial \zeta_y} & \text{for } \alpha_l \leq \alpha^* \leq \alpha_u \\ 0 & \text{for } \alpha^* > \alpha_u \end{cases}$$

Thus to determine the partial derivatives of the constrained gain α_c^*, we must compute the partial derivatives of the unconstrained gain α^*, which is defined by Equation 14:

$$\alpha^* = \frac{\left\langle \Gamma^{obs} - \Gamma^{k-1}, g\left(\Gamma^{k-1} - \zeta; \sigma\right)\right\rangle dt}{\left|\left| g\left(\Gamma^{k-1} - \zeta; \sigma\right)\right|\right|_2^2}.$$

where we have used

$$g(\Gamma^{k-1} - \zeta; \sigma) = (\Gamma^{k-1} - \zeta)\frac{1}{r}\sin\left(\frac{2\pi r}{\sigma}\right)\exp\left(\frac{-r^2}{\sigma^2}\right) = (\Gamma^{k-1} - \zeta)G(r). \quad \text{(B.1)}$$

Computing the partial derivatives with respect to the scale σ parameter and location parameters ζ_x and ζ_y, we obtain:

$$\frac{\partial \alpha_*}{\partial \sigma} = \frac{\left(\frac{\partial}{\partial \sigma}\left\langle \Gamma^{res}, g\right\rangle\right)\|g\|^2 - \left\langle \Gamma^{res}, g\right\rangle \frac{\partial}{\partial \sigma}\|g\|^2}{\|g\|^4}, \text{ where:}$$

$$\frac{\partial}{\partial \sigma}\left\langle \Gamma^{res}, g\right\rangle = \frac{\partial}{\partial \sigma}\int_0^1 \Gamma_x^{res}(t)(\Gamma_x^{k-1}(t) - \zeta_x)G + \Gamma_y^{res}(t)(\Gamma_y^{k-1}(t) - \zeta_y)G dt$$

$$= \int_0^1 \left\langle \Gamma^{res}(t), \Gamma^{k-1}(t) - \zeta\right\rangle \frac{\partial G}{\partial \sigma}dt$$

$$\frac{\partial}{\partial \sigma}\|g\| = \frac{\partial}{\partial \sigma}\int_0^1 [(\Gamma_x^{k-1}(t) - \zeta_x)G]^2 + [(\Gamma_y^{k-1}(t) - \zeta_y)G]^2 dt$$

$$= \int_0^1 2G\frac{\partial G}{\partial \sigma}\|\Gamma^{k-1}(t) - \zeta\|^2$$

$$\frac{\partial \alpha_*}{\partial \zeta_x} = \frac{\left(\frac{\partial}{\partial \zeta_x}\left\langle \Gamma^{res}, g\right\rangle\right)\|g\|^2 - \left\langle \Gamma^{res}, g\right\rangle \frac{\partial}{\partial \zeta_x}\|g\|^2}{\|g\|^4}, \text{ where:}$$

28

$$\frac{\partial}{\partial \zeta_x} \langle \Gamma^{res}, g \rangle = \frac{\partial}{\partial \zeta_x} \int_0^1 \Gamma_x^{res}(t)(\Gamma_x^{k-1}(t) - \zeta_x)G + \Gamma_y^{res}(t)(\Gamma_y^{k-1}(t) - \zeta_y)G dt$$

$$= \int_0^1 \Gamma_x^{res}(t) \left[-G + (\Gamma_x^{k-1}(t) - \zeta_x)\frac{\partial G}{\partial \zeta_x} \right] + \Gamma_y^{res}(t) \left[(\Gamma_y^{k-1}(t) - \zeta_y)\frac{\partial G}{\partial \zeta_x} \right] dt$$

$$\frac{\partial}{\partial \zeta_x} \|g\|^2 = \frac{\partial}{\partial \zeta_x} \int_0^1 [(\Gamma_x^{k-1}(t) - \zeta_x)G]^2 + [(\Gamma_y^{k-1}(t) - \zeta_y)G]^2 dt$$

$$= \int_0^1 2\left[(\Gamma_x^{k-1}(t) - \zeta_x)G \right] \left[-G + (\Gamma_x^{k-1}(t) - \zeta_x)\frac{\partial G}{\partial \zeta_x} \right] + 2G\frac{\partial G}{\partial \zeta_x}(\Gamma_y^{k-1}(t) - \zeta_y)^2 dt$$

$$\frac{\partial \alpha_*}{\partial \zeta_y} = \frac{\left(\frac{\partial}{\partial \zeta_y} \langle \Gamma^{res}, g \rangle \right) \|g\|^2 - \langle \Gamma^{res}, g \rangle \frac{\partial}{\partial \zeta_y} \|g\|^2}{\|g\|^4}, \text{ where:}$$

$$\frac{\partial}{\partial \zeta_y} \langle \Gamma^{res}, g \rangle = \frac{\partial}{\partial \zeta_y} \int_0^1 \Gamma_x^{res}(t)(\Gamma_x^{k-1}(t) - \zeta_x)G + \Gamma_y^{res}(t)(\Gamma_y^{k-1}(t) - \zeta_y)G dt$$

$$= \int_0^1 \Gamma_x^{res}(t) \left[(\Gamma_x^{k-1}(t) - \zeta_x)\frac{\partial G}{\partial \zeta_y} \right] + \Gamma_y^{res}(t) \left[-G + (\Gamma_y^{k-1}(t) - \zeta_y)\frac{\partial G}{\partial \zeta_y} \right] dt$$

$$\frac{\partial}{\partial \zeta_y} \|g\|^2 = \frac{\partial}{\partial \zeta_y} \int_0^1 [(\Gamma_x^{k-1}(t) - \zeta_x)G]^2 + [(\Gamma_y^{k-1}(t) - \zeta_y)G]^2 dt$$

$$= \int_0^1 2G\frac{\partial G}{\partial \zeta_y}(\Gamma_x^{k-1}(t) - \zeta_x)^2 + 2(\Gamma_y^{k-1}(t) - \zeta_y)G \left[-G + (\Gamma_y^{k-1}(t) - \zeta_y)\frac{\partial G}{\partial \zeta_y} \right] dt$$

We are now ready to compute the Jacobian matrix. From Equation B.1 we have that:

$$\xi_x(t_i) = \Gamma_x^{res}(t_i) - \alpha \left(\Gamma_x^{k-1}(t_i) - \zeta_x \right) G$$
$$\xi_y(t_i) = \Gamma_y^{res}(t_i) - \alpha \left(\Gamma_y^{k-1}(t_i) - \zeta_y \right) G$$

Thus,

$$\frac{\partial \xi_x(t_i)}{\partial \sigma} = -(\Gamma_x^{k-1}(t_i) - \zeta_x) \left[\frac{\partial \alpha}{\partial \sigma}G + \alpha\frac{\partial G}{\partial \sigma} \right] \tag{B.2}$$

$$\frac{\partial \xi_y(t_i)}{\partial \sigma} = -(\Gamma_y^{k-1}(t_i) - \zeta_y) \left[\frac{\partial \alpha}{\partial \sigma}G + \alpha\frac{\partial G}{\partial \sigma} \right] \tag{B.3}$$

$$\frac{\partial \xi_x(t_i)}{\partial \zeta_x} = -\frac{\partial \alpha}{\partial \zeta_x}(\Gamma_x^{k-1}(t_i) - \zeta_x)G + \alpha G - \alpha(\Gamma_x^{k-1}(t_i) - \zeta_x)\frac{\partial G}{\partial \zeta_x} \tag{B.4}$$

$$\frac{\partial \xi_y(t_i)}{\partial \zeta_x} = -(\Gamma_y^{k-1}(t_i) - \zeta_y) \left[\frac{\partial \alpha}{\partial \zeta_x}G + \alpha\frac{\partial G}{\partial \zeta_x} \right] \tag{B.5}$$

$$\frac{\partial \xi_x(t_i)}{\partial \zeta_y} = -(\Gamma_x^{k-1}(t_i) - \zeta_x) \left[\frac{\partial \alpha}{\partial \zeta_y}G + \alpha\frac{\partial G}{\partial \zeta_y} \right] \tag{B.6}$$

$$\frac{\partial \xi_y(t_i)}{\partial \zeta_y} = -\frac{\partial \alpha}{\partial \zeta_y}(\Gamma_y^{k-1}(t_i) - \zeta_y)G + \alpha G - \alpha(\Gamma_y^{k-1}(t_i) - \zeta_y)\frac{\partial G}{\partial \zeta_y} \tag{B.7}$$

where for the Gabor basis, we have:

$$\frac{\partial G}{\partial \sigma} = \exp\left(-\frac{r^2}{\sigma^2}\right)\left[\frac{-2\pi}{\sigma^2}\cos\left(\frac{2\pi r}{\sigma}\right) + \frac{2r}{\sigma^3}\sin\left(\frac{2\pi r}{\sigma}\right)\right]$$

$$\frac{\partial G}{\partial \zeta_x} = \exp\left(-\frac{r^2}{\sigma^2}\right)\frac{(\Gamma_x^{k-1}(t_i) - \zeta_x)}{r}\left[-\frac{2\pi}{\sigma r}\cos\left(\frac{2\pi r}{\sigma}\right) + \frac{2}{\sigma^2}\sin\left(\frac{2\pi r}{\sigma}\right) + \frac{1}{r^2}\sin\left(\frac{2\pi r}{\sigma}\right)\right]$$

$$\frac{\partial G}{\partial \zeta_y} = \exp\left(-\frac{r^2}{\sigma^2}\right)\frac{(\Gamma_y^{k-1}(t_i) - \zeta_y)}{r}\left[-\frac{2\pi}{\sigma r}\cos\left(\frac{2\pi r}{\sigma}\right) + \frac{2}{\sigma^2}\sin\left(\frac{2\pi r}{\sigma}\right) + \frac{1}{r^2}\sin\left(\frac{2\pi r}{\sigma}\right)\right]$$

It is straightforward to show that Equations B.2 - B.7 also apply to the Gaussian and Spline bases (Section 7.2), with suitable definitions of $G(r; \sigma)$:

Gaussian Basis:

$$G(r; \sigma) = \frac{2\pi}{\sigma}\exp\left(\frac{-r^2}{\sigma^2}\right)$$

$$\frac{\partial G}{\partial \sigma} = 2\pi\exp\left(\frac{-r^2}{\sigma^2}\right)\left[\frac{-1}{\sigma^2} + \frac{2r^2}{\sigma^4}\right]$$

$$\frac{\partial G}{\partial \zeta_x} = \frac{4\pi}{\sigma^3}\exp\left(\frac{-r^2}{\sigma^2}\right)(\Gamma_x^{k-1} - \zeta_x)$$

$$\frac{\partial G}{\partial \zeta_y} = \frac{4\pi}{\sigma^3}\exp\left(\frac{-r^2}{\sigma^2}\right)(\Gamma_y^{k-1} - \zeta_y)$$

Spline Basis:

$$G(r; \sigma) = \frac{(r-\sigma)^2}{\sigma^2}$$

$$\frac{\partial G}{\partial \sigma} = \frac{-2}{\sigma^3}(r-\sigma)^2 - \frac{2}{\sigma^2}(r-\sigma)$$

$$\frac{\partial G}{\partial \zeta_x} = \frac{-2}{r\sigma^2}(r-\sigma)(\Gamma_x^{k-1} - \zeta_x)$$

$$\frac{\partial G}{\partial \zeta_y} = \frac{-2}{r\sigma^2}(r-\sigma)(\Gamma_y^{k-1} - \zeta_y)$$