Building better formlet codes for planar shape

Alex Yakubovich
Centre for Vision Research
York University
Toronto, Canada
yakuboa@yorku.ca

James H. Elder
Centre for Vision Research
York University
Toronto, Canada
jelder@yorku.ca

***Abstract*—The GRID/formlet representation of planar shape has a number of nice properties [4], [10], [3], but there are also limitations: it is slow to converge for shapes with elongated parts, and it can be sensitive to parameterization as well as grossly ill-conditioned. Here we describe a number of innovations on the GRID/formlet model that address these problems: 1) By generalizing the formlet basis to include oriented deformations we achieve faster convergence for elongated parts. 2) By introducing a modest regularizing term that penalizes the total energy of each deformation we limit redundancy in formlet parameters and improve identifiability of the model. 3) By applying a recent contour remapping method [9] we eliminate problems due to drift of the model parameterization during matching pursuit. These innovations are shown to both speed convergence and to improve performance on a shape completion task.**

***Keywords*-GRID; Formlets; Planar shape; Generative models; Contour completion; Deformation; Diffeomorphisms**

I. Introduction

Discriminative models for shape (e.g. [1]) can be sufficient for specific recognition tasks, but there are many applications where a full generative model could be useful, to generate complete shape hypotheses from localized shape features in cluttered scenes, for example. Surprisingly, there has been relatively little recent work on generative models for planar shape. One notable exception of the shapelet theory of Dubinskiy & Zhu [2], which can be thought of as a 'waveletization' of the Fourier descriptor model of shape [11]. While the shapelet theory enjoys many attractive wavelet-like properties (e.g., linearity, sparsity, locality, self-similarity), it also suffers from a crucial limitation as a generative model: samples drawn from the model are very often topologically incorrect. For example, although one may wish to model simple closed curves, samples often contain self-intersections.

A different approach to modeling shape is based on coordinate transformations of the image space. This idea goes back at least to D'Arcy Thompson, who modelled variation in animal shapes using global coordinate transformations [14]. Coordinate transformation have also been used in computer vision (e.g. [13], [6]), but in general these methods do not preserve the topology of embedded contours.

Grenander et al. address this limitation in their GRID (Growth By Random Iterated Diffeomorphisms) model [5], [4]. GRID expresses changes in time-series medical images as a composition of small, local deformations of the plane. By constraining these deformations to be diffeomorphic, the topology of the image structures are preserved [5], [4], [12]. This idea has been extended to the complete representation of a planar shape as a sequence of localized diffeomorphisms called *formlets*, applied to a simple embryonic shape (e.g., a circle or an ellipse)[10], [3]. The theory can be used to approximate an arbitrary target shape using a variation of matching pursuit [7], [2].

The GRID/formlet theory shares many of the desirable properties enjoyed by shapelets. Both are capable of modeling arbitrarily complex shapes, through the action of multiple relatively simple components. As they are both essentially wavelet theories, they both have good localization and scaling properties, and are capable of producing sparse descriptions in a progressive manner. However, formlets have a key advantage: since they are diffeomorphic mappings of the image space, the topology of the evolving shape will not deviate from the topology of the original embryonic shape. Moreover, by working in the image space, formlets capture regional properties of shape. Formlets have been found to outperform shapelets on the important task of shape completion [10], [3], suggesting that they may be a better model for inferring complete shapes from contours fragmented by occlusions and shadows in complex scenes.

While the GRID/formlet theory offers many advantages, it has limitations: 1) the isotropic nature of formlets is not well suited to modelling shapes with elongated features (e.g. limbs of a giraffe), 2) the correspondence between a shape and its representation remains fixed throughout the pursuit, which slows convergence and leads to unintuitive representations, and 3) formlet estimation is ill-conditioned in the sense that very different formlet parameters can yield very similar deformations. By addressing these limitations we show that convergence and performance on a shape completion task is substantially improved.

978-1-4799-4337-1/14 $31.00 © 2014 IEEE
DOI 10.1109/CRV.2014.19

A. Dataset and evaluation methodology

We evaluate our methods on the same dataset of 391 animal outlines used in the original formlet papers [10], [3]. Each of these outlines has been sampled at 128 points at regular arc-length intervals. In addition to qualitative comparison, we quantitatively evaluate performance using the two methods employed in the original formlet papers [10], [3]:

1) **Rate of convergence.** We measure how quickly the mean $\mathcal{L}^2$ error drops as a function of the number of formlets used in the representation.
2) **Shape completion.** We simulate the occlusion of a 10% or 30% continuous portion of the contour, using the remaining data to learn the representation. We then measure the $\mathcal{L}^2$ error on the occluded portion. Lower error suggests that the model embodies more knowledge of the properties of natural shape.

We measure error using an $\mathcal{L}^2$ Hausdorff metric, in order to be independent of the parameterization of the target and model curves. This metric defines the error between the target shape and its representation as the symmetrical average minimum distance of a point on one shape to the other:

$$\xi^H(\Gamma^{obs},\Gamma^k) = \frac{1}{2}\Big(\overline{\min_{t'\in[0,1)} d\left(\Gamma^{obs}(t),\Gamma^k(t')\right)}\ldots + \overline{\min_{t'\in[0,1)} d\left(\Gamma^{obs}(t'),\Gamma^k(t)\right)}\Big) \quad (1)$$

where $d(\mathbf{x},\mathbf{y})$ denotes the Euclidean distance between vectors $\mathbf{x}$ and $\mathbf{y}$ and $\overline{\Gamma} = \frac{1}{n}\sum_{i=1}^{n}\Gamma(t_i)$ denotes the mean over all points on the shape.

Mean rate of convergence will be analyzed as we introduce each innovation on the GRID/formlet model. Shape completion results are presented in Section VI.

II. Formlets

The isotropic formlet [10], [3] is a localized, Gabor-like, radial deformation $f:\mathbb{R}^2\to\mathbb{R}^2$: [1]

$$f(\mathbf{x};\mathbf{x_0},\sigma,\alpha) = \mathbf{x_0} + \frac{\mathbf{x}-\mathbf{x_0}}{\|\mathbf{x}-\mathbf{x_0}\|}\rho(\|\mathbf{x}-\mathbf{x_0}\|) \quad (2)$$

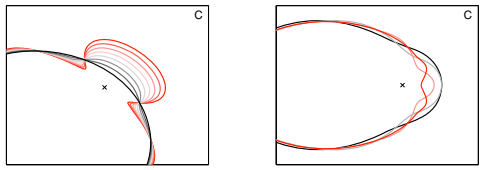
$$\rho(r;\alpha,\sigma) = r + \alpha\sin\left(\frac{2\pi r}{\sigma}\right)\exp\left(\frac{-r^2}{\sigma^2}\right) \quad (3)$$

This transformation can be shown to be a diffeomorphism of the plane as long as the gain parameter satisfies the linear constraint

$$-\frac{\sigma}{2\pi} \le \alpha \lessapprox .1956\sigma. \quad (4)$$

[1]This Gabor formulation is similar but not identical to one version of the diffeomorphisms used in GRID, where the radial deformation function can be expressed as $\rho(r;\alpha,\sigma) = r + \alpha r\exp\left(\frac{-r^2}{\sigma^2}\right)$. The Gabor formulation has been found to yield faster convergence for natural shapes [3].

Each formlet can be thought of as a packet of three parameters: location $\mathbf{x_0}\in\mathbb{R}^2$, scale $\sigma\in\mathbb{R}^+$ and gain $\alpha\in\mathbb{R}$. The scale controls the extent of the image region affected by the formlet, and the gain controls the magnitude and sign of the perturbation (compression or expansion). Figure 1 shows how the gain and scale parameters affect the nature of the deformation.



(a) f with gain variation (b) f with scale variation

Figure 1: Several formlets applied to an ellipse. The formlets shown in red do not satisfy the diffeomorphism constraints.

Starting with an embryonic shape Γ^0, formlets can be composed to represent an arbitrarily complex shape:

$$\Gamma^K(t) = (f_K\circ f_{K-1}\circ\ldots\circ f_1)(\Gamma^0(t))$$

To learn a formlet sequence $\{f_1\ldots f_K\}$ that represents an observed shape Γ^{obs}, a variation of matching pursuit [7] is employed, which iteratively selects formlets in greedy fashion to minimize reconstruction error:

$$f_k = \arg\min_f E(\Gamma^{obs}, f(\Gamma^k - 1)), \quad (5)$$

where error is measured by the $\mathcal{L}^2$ norm of the residual:

$$E(\Gamma^{obs},\Gamma^K) = \|\Gamma^{obs}-\Gamma^K\|_2^2 = \int(\Gamma_x^{k-1}(t)) - \Gamma_x^{obs}(t))^2 + (f(\Gamma_y^{k-1}(t)) - \Gamma_y^{obs}(t))^2 dt \quad (6)$$

In an early version of the work [10] formlets were selected by exhaustive dictionary search, however this was later updated to a combination of exhaustive search and gradient descent which the authors dubbed 'dictionary descent' [3]. Given the location $\mathbf{x_0}$ and scale σ, the optimal gain α^* can be determined analytically:

$$\alpha^* = \frac{\langle\Gamma^{obs}-\Gamma^{k-1}, g(\Gamma^{k-1}-\mathbf{x_0};\sigma)\rangle}{\|g(\Gamma^{k-1}-\mathbf{x_0};\sigma)\|_2^2}, \quad \text{where}$$
$$g(\mathbf{x}-\mathbf{x_0};\sigma) = \frac{\mathbf{x}-\mathbf{x_0}}{r}\sin\left(\frac{2\pi r}{\sigma}\right)\exp\left(\frac{-r^2}{\sigma^2}\right) \quad (7)$$

and the inner product is defined as: $\langle\Gamma^a,\Gamma^b\rangle = \int_0^1\left(\Gamma_x^A(t)\Gamma_x^B(t) + \Gamma_y^A(t)\Gamma_y^B(t)\right)dt \quad \forall\Gamma^a,\Gamma^b:[0,1]\to\mathbb{R}^2$.

A. Initialization

In the original shapelet [2] and formlet [10], [3] work, target shapes were first normalized to have unit variance in horizontal and vertical dimensions, and the model was initialized as the best-fitting ellipse to the resulting normalized target shape. To facilitate comparison, we adopt the same initialization method for the majority of this paper.

This initialization method generally produces an embryonic ellipse with a balance of positive and negative error, i.e., some regions where the model incorporates part of the background, and other regions where it excludes some of the target shape. As a result, pursuit involves a rough balance between expansion and contraction of the model as it converges to the target shape.

In the last sections of the paper, we explore an alternative initialization, namely, the maximal circle inscribing the target shape. In this initialization scheme, the error starts out as strictly positive (the model never includes background), and as a result pursuit largely involves a series of expansion or growth processes, which we conjecture may yield more natural formlets with a stronger correspondence to perceived part structure. The maximal inscribing circle is computed using a standard distance transform method [8].

III. Oriented formlets

While isotropic deformations are good for representing shapes with roughly circular features, reconstruction converges slowly for shapes with elongated parts.

An image point $\mathbf{x} \in \mathbb{R}^2$ can be expressed in a polar coordinate system centred at $\mathbf{x_0}$: $(x, y) \leftrightarrow (r, \theta)$, where $r = \|\mathbf{x} - \mathbf{x_0}\|$ and $\theta = \arg(\mathbf{x} - \mathbf{x_0})$. The formlet transformation [10], [3] can then be written as:

$$\mathbf{x} \mapsto f(\mathbf{x}) \iff (r, \theta) \mapsto (\rho(r), \theta)$$

where $\rho : \mathbb{R}^+ \mapsto \mathbb{R}^+$.

Since the radial deformation $\rho(r)$ is a function only of the distance of the point $\mathbf{x}$ from the centre of the deformation, and not the angle, the formlet is isotropic. We can generalize the definition to allow anisotropic deformations by making the radial deformation a function of angle as well:

$$(r, \theta) \mapsto (\rho(r, \theta), \theta) \tag{8}$$

where now $\rho : \mathbb{R}^+ \times \mathbb{S} \mapsto \mathbb{R}^+$.

Following Grenander et al. [5], [4], we factor the radial deformation into two components: $R(r)$, the *radial deformation function* (RDF) depending only upon r, and $A(\theta)$, the *angular deformation function* (ADF), depending only upon θ. While Grenander et al. assumed a nonparametric form for the ADF, here we use a von-Mises ADF, so that our oriented formlet can be written as:

$$f(\mathbf{x}; \mathbf{x_0}, \sigma, \alpha, \kappa, \theta_0) = \mathbf{x_0} + \frac{\mathbf{x} - \mathbf{x_0}}{\|\mathbf{x} - \mathbf{x_0}\|} \rho(\|\mathbf{x} - \mathbf{x_0}\|, \theta), \tag{9}$$

$$\rho(r, \theta) = r + \alpha \cdot R(r; \sigma) \cdot A(\theta; \theta_0, \kappa) \tag{10}$$

$$R(r; \sigma) = \sin\left(\frac{2\pi r}{\sigma}\right) \exp\left(\frac{-r^2}{\sigma^2}\right) \tag{11}$$

$$A(\theta; \theta_0, k) \propto \exp(\kappa \cos(\theta - \theta_0)) \tag{12}$$

As shown in the appendix, a formlet f is a diffeomorphism if the conditions on α (Eqn. II) are satisfied and $0 \leq A(\theta) \leq 1 \quad \forall \theta \in [0, 2\pi]$. Since $A(\theta)$ is everywhere nonnegative, it suffices to normalize it by its mode, $A(\theta_0) = \exp(\kappa \cos(0)) = \exp(\kappa)$. Therefore,:

$$A(\theta; \theta_0, k) = \exp(-\kappa) \exp(\kappa \cos(\theta - \theta_0)) \tag{13}$$

$$= \exp(\kappa \cos(\theta - \theta_0) - \kappa) \tag{14}$$

Fig. 2 contrasts the action of an isotropic formlet and an oriented formlet on an ellipse. Fig. 3 compares mean convergence rates for shapelets, isotropic formlets and oriented formlets on the 391 animal shape dataset. While the improvement in mean convergence rate for oriented formlets is relatively modest, we will see in Section VI that the impact on shape completion is quite substantial. Run time on a 2.7 GHz Intel Core i5 processor is roughly 2 sec per iteration for the isotropic formlet model, and 7.2 sec per iteration for the oriented formlet model.

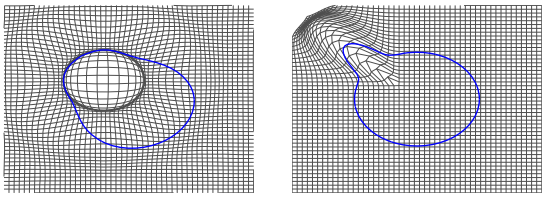


Figure 2: Example of an isotropic formlet (left) and an oriented formlet (right)

IV. Reparameterization

To assess potential formlets at each stage of pursuit, we need a 1:1 correspondence between points on the model and points on the target shape. In earlier work [10], [3], the correspondence was established at initialization, when the embryonic ellipse was fitted to the target shape, and this correspondence remained fixed over the duration of formlet pursuit. This is problematic: as the model shape evolves, the fixed correspondence is generally no longer optimal (i.e., produces a much larger residual than a Hausdorff distance), and could lead to slower convergence and the selection of unnatural formlets. To solve this problem, we need an efficient method for reparameterizing the model shape to reestablish an accurate 1:1 mapping with the target shape after the selection and application of each formlet.

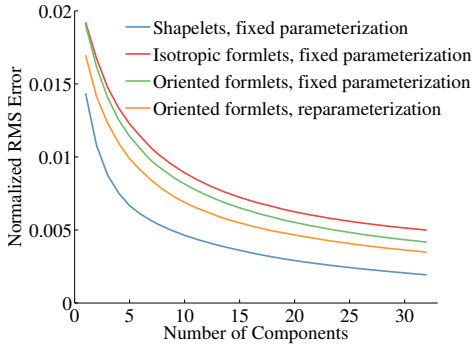


Figure 3: Comparing mean convergence rates for formlet system with correspondence determined by the contour mapping measure (CM) and the baseline model.

A. The contour mapping measure

Here we explore the use of a shape distance measure introduced by Movahedi et al. [9] for evaluating segmentation algorithms. They considered a number of methods for measuring the distance between two simple, closed curves, and demonstrated that most methods, including some used extensively in the literature, exhibit serious failure modes. Based upon this analysis, they proposed a novel *contour mapping measure* for measuring the distance between two curves, and showed that it is more consistent with human judgement of shape similarity than competing methods.

Given two sequences of points of length m and n, the contour mapping mapping measure finds the monotonic (but not generally 1:1) mapping between them that minimizes $\mathcal{L}^2$ error. The mapping is computed by dynamic programming in $\mathcal{O}(mn \log n)$ time.

The contour mapping measure allows 1:n and n:1 mapping between the two contours. These matchings occur when there is a different degree of detail in the parts being matched (e.g. a circular arc on the model shape is matched to an elongated limb on the target shape). It is critical that the measure have this flexibility, otherwise the correspondence between the two shapes will become misaligned. However, a 1:1 correspondence is also crucial for formlet pursuit, otherwise convergence may stall. For example, consider a case where a point on the model is matched to two points on the target shape that are equidistant but in opposite directions: there is no formlet that can reduce the residual.

We address this problem as follows. After using the contour mapping measure to optimize the parameterization of the model, we execute a series of split and merge operations on the model points to restore the mapping to 1:1, using the following two rules:

1) *(1:n)* If a point $\Gamma^k(t)$ on the model curve is matched to n points on the target curve, replace it with n points linearly interpolated between its neighbours.
2) *(n:1)* If n points $\Gamma^k(t_1) \ldots \Gamma^k(t_n)$ on the model curve are matched to a single point on the target curve, replace the n points by the midpoint of the two points flanking the n points, $\Gamma^{k-1}(t_0), \Gamma^k(t_{n+1})$.

Fig. 4 illustrates the reparameterization process.

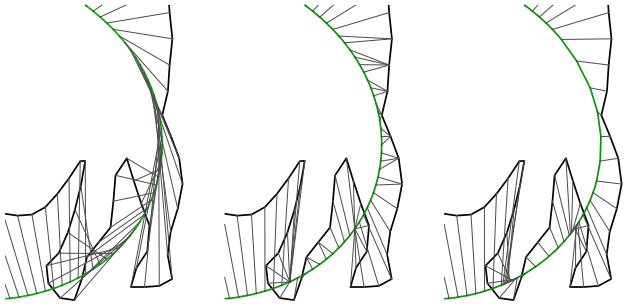


Figure 4: Reparameterization algorithm. (a) Sub-optimal correspondence before reparameterization. (b) Optimal correspondence after applying the contour mapping measure. (c) Restoration of the 1:1 mapping after split/merge stage.

Fig 3 shows that reparameterization substantially speeds convergence for the oriented formlet representation. We find empirically that the reparameterization has negligible impact on run time (less than 100 msec per iteration).

V. Regularization

In GRID/formlet theory, shape is modeled through a sequence of deformations of the image. The objective function used to select these deformations, however, is only sensitive to properties of the bounding contour. This serves to make identification of the formlet model ill-posed, as two very different formlets may have very a similar effect on the model shape (Fig. 5).

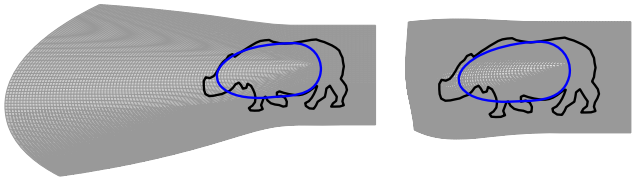


Figure 5: Two formlets applied to the inscribed circle. The scales of the formlets on the left and right are 1.5×10^9 pixels and 1568 pixels respectively. The image is 800×600 pixels.

This problem can be addressed by regularizing the objective function. Given a formlet with parameters $\mathbf{p} = (x_0, y_0, \sigma, \alpha, \kappa, \theta_0)$, we augment the objective function E to include both a fidelity term F and a regularizing term G:

$$E(\Gamma^{obs}, f(\Gamma^{k-1}; \mathbf{p})) = F(\Gamma^{obs}, f(\Gamma^{k-1}; \mathbf{p})) + \lambda G(\mathbf{p}) \tag{15}$$

Since the parameter space is heterogenous and the mapping from parameter space to error is highly nonlinear, it is convenient to define the regularizing term in terms of the data space (the image) rather than the parameter space. In particular, we define the penalty G as the energy of the deformation field, which is obtained by integrating the magnitude of the deformation over the plane :

$$G(\mathbf{p}) = \int_{\mathbb{R}^2} \|f(\mathbf{x};\mathbf{p}) - \mathbf{x}\|_2^2 \quad d\mathbf{x} \tag{16}$$

$$= C\alpha^2\sigma^2 I_0(2\kappa)e^{-2\kappa}, \quad \text{where } C \approx 0.2568\pi \tag{17}$$

and $I_0(\kappa)$ is the modified Bessel function of order 0:

$$I_\alpha(\kappa) = \sum_{m=0}^{\infty} \left(\frac{1}{m!}\left(\frac{\kappa}{2}\right)^m\right)^2 \tag{18}$$

(See Appendix for this derivation.) Note that the fidelity term is a line integral over the shape whereas the penalty term is an area integral over the plane. To establish a common frame of reference, we multiply the fidelity term by a constant L approximating the average thickness of the residual: $R = (S^{obs} \cup S^k) - (S^{obs} \cap S^k)$, where $S^{obs} = \text{interior}(\Gamma^{obs})$ and $S^k = \text{interior}(\Gamma^k)$:

$$F = L \cdot \|f(\Gamma^{k-1}) - \Gamma^{obs}\|_2^2. \tag{19}$$

A. *Selecting the regularization parameter λ*

Regularization will slow convergence but will improve identifiability of the model. To achieve a balance, we estimate the maximum λ that will slow convergence by no more than one iteration:

$$\lambda^{opt} = \max\{\lambda : \xi_\lambda^k \leq \xi_{\lambda_0}^{k-1}\} \tag{20}$$

where ξ_λ^k is the Hausdorff error at iteration k with regularization parameter λ, and $\xi_{\lambda_0}^{k-1}$ is the error at iteration $k-1$ with no regularization ($\lambda = 0$). We tested 32 values of λ ranging over the interval [0.0002, 0.006] on a set of 128 randomly selected training shapes. The interval was chosen with a coarse-to-fine approach. Fig. 6 shows how formlet pursuit using this regularized objective alters the parameters of the selected formlets, greatly reducing the average gain and scale, and thus producing a much more localized model.

VI. Evaluation: Contour Completion

We have seen that generalizing formlet theory to include oriented deformations, and reparameterizing to maintain a natural correspondence between target and model shapes both lead to faster convergence. Also, regularization in the data space can effectively dampen the scale and gain of selected formlets, which should improve the identifiability of the model, without substantial impact on convergence. Now we examine how these innovations impact the model's ability to complete fragmented shapes, which provides an indication of the utility of the model as a global guide for perceptual organization in cluttered scenes [3].

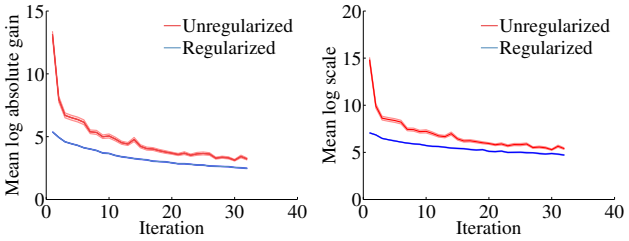


Figure 6: Average gains ($|\alpha|$) and scales (σ) selected by the oriented formlet model, with and without regularization. Shaded areas indicate standard error of the mean.

Fig. 7 shows an example of shape completion by shapelet, isotropic and oriented formlet models. Notice the topological errors induced by the shapelet model, and the improvement in the estimated completion for the oriented formlet model.

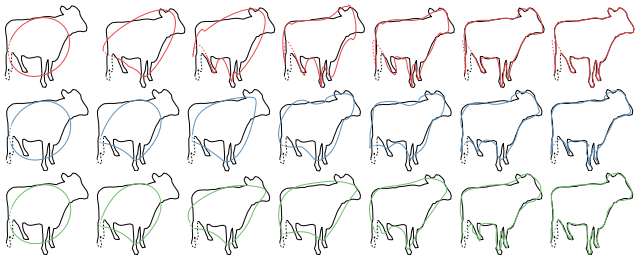


Figure 7: Pursuit of a partially occluded cow shape using shapelets (top row), isotropic formlets (middle row) and oriented formlets (bottom row). Iterations 0,1,2,4,8,16,32 are shown from left to right. The dash-dot line indicates the occluded portion of the contour.

Fig. 8 shows how these three models compare quantitatively on this task. As previously reported [10], [3], formlets are more accurate than shapelets on the task of shape completion. The new result here is that not only do oriented formlets converge faster than isotropic formlets, but they also lead to even more accurate shape completion. Faster convergence is not too surprising, since the oriented formlet has more parameters and in fact subsumes the isotropic formlet as a degenerate case. However, the improvement in completion performance is not so obvious, and suggests that oriented formlets somehow more accurately reflect the structure of natural shapes.

It is also of interest to consider how regularization might affect shape completion accuracy. Fig. 9 illustrates how unregularized pursuit leads to very large deformations in the image, well beyond the neighbourhood of the visible shape. If these deformations are not somehow reigned-in, they could lead to wild completions in the model contour for

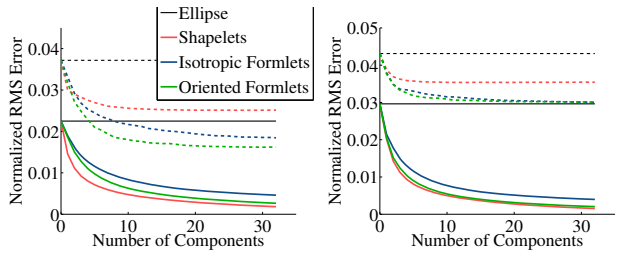


Figure 8: Completion error at 10% (left) and 30% occlusion (right). Solid and dashed lines indicate performance on the visible and occluded parts of the shape respectively. The black lines indicate the accuracy of the least-squares fit ellipse used as the embryonic shape Γ^0 for all methods.

occluded regions, where formlet selection is not constrained by the visible data. Fig. 10 seems to bear this out: regularized completion is more accurate for this particular shape.

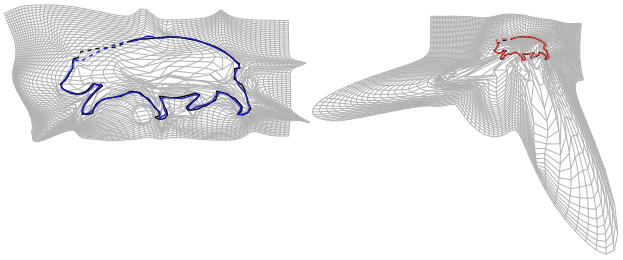


Figure 9: Total image deformation after 32 iterations of oriented formlet pursuit, with (left) and without (right) regularization.

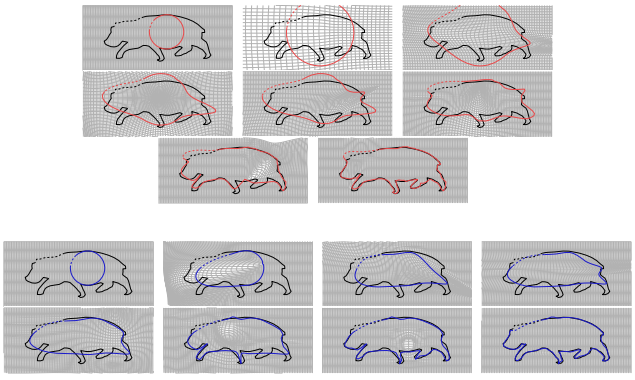


Figure 10: Pursuit of rhinoceros shape with unregularized (top) and regularized (bottom) formlet systems. Iterations 0,1,2,3,4,8,16,32 are shown from left to right.

Consistent with this example, Fig. 11 shows that quantitatively, over the database of 391 animal shapes, regularization improves shape completion for the oriented formlet model, while having minimal effects on convergence.

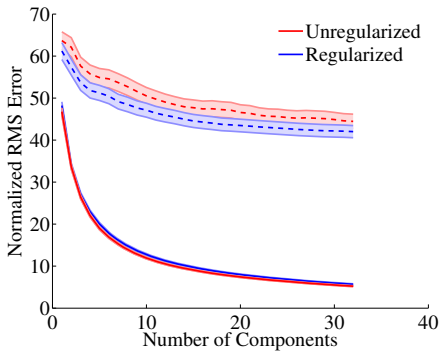


Figure 11: Completion error at 10% occlusion for regularized and unregularized models, using the maximal inscribed circle as the embryonic shape. Dashed lines indicate performance on the occluded component. Error bars ($\pm\frac{s}{\sqrt{n}}$) are indicated by the shaded area. (The units of error differ from previous plots as we do not normalize shape coordinates to [0,1] with this initialization)

VII. Conclusion

The GRID/formlet model for planar shape representation has a number of nice properties, but also limitations, and our goal in this paper has been to address these limitations. We have shown that by generalizing the isotropic model to include oriented formlets, and by introducing a novel and efficient form of reparameterization during formlet pursuit, we can substantially speed convergence and improve the accuracy of shape completion. We have shown that a novel method for regularizing formlet pursuit in the data domain further improves shape completion performance, while having minimal impact on convergence.

A natural direction for future work is to explore alternatives to the strictly sequential structure of the model, such as a hierarchical organization that captures the part-based structure of natural shapes.

VIII. Appendix

A. *Diffeomorphism Constraint*

Versions of the diffeomorphism constraint for the isotropic case can be found in [4], [10], [3]. Derivation of the constraint for oriented formlets is similar; we included it here for completeness.

Since f is radial, it will be a injective if and only if $\rho(r,\theta)$ is monotonically increasing or decreasing. ρ cannot be everywhere decreasing, since $\rho(0,\theta) = 0$ and ρ is everywhere nonnegative. Therefore, it to determine the conditions under which $\rho(r,\theta)$ is monotonically increasing. In other words, we require that

$$\frac{\partial}{\partial r}\rho(r,\theta) > 0 \quad \forall\theta \in [-\pi,\pi] \tag{21}$$

From equation (10) we have that:

$$\frac{\partial}{\partial r}\rho(r;\theta) = 1 + \alpha A(\theta) R'(r;\theta) \tag{22}$$
$$\equiv 1 + \alpha A \psi(r) \tag{23}$$

where we defined:

$$\psi(r) = e^{-r^2/\sigma^2}\left[\frac{2\pi}{\sigma}\cos\left(\frac{2\pi r}{\sigma}\right) - \frac{2r}{\sigma^2}\sin\left(\frac{2\pi r}{\sigma}\right)\right] \tag{24}$$

We assume that $0 \leq A(\theta) \leq 1$. First, we consider the case where $\alpha < 0$:

$$0 < 1 + \alpha A \psi(r) \tag{25}$$
$$< 1 + \alpha \psi(0) \tag{26}$$
$$= 1 + \alpha A \frac{2\pi}{\sigma} \tag{27}$$
$$\iff -\frac{\sigma}{2\pi} < \alpha A, \quad \text{for } 0 \leq A(\theta) \leq 1 \tag{28}$$
$$\iff -\frac{\sigma}{2\pi} < \alpha \tag{29}$$

where in equation 27 we used the fact that $\arg\max_{r>0} \psi(r) = 0$. If $\alpha > 0$, we have:

$$0 < 1 + \alpha A \psi(r) \tag{30}$$
$$0 < 1 + \alpha A \psi_{min}, \quad \text{for } 0 \leq A(\theta) \leq 1 \tag{31}$$
$$0 < 1 + \alpha \psi_{min} \tag{32}$$
$$\alpha < \frac{-1}{\psi_{min}} \approx \frac{-1}{-5.112/\sigma} = 0.1956\sigma \tag{33}$$

B. Regularizing term

Given the transformation in Equation 9, the squared displacement of a given point $\mathbf{x} \in \mathbb{R}^2$ is:

$$\|f(\mathbf{x}) - \mathbf{x}\|_2^2 = \|\mathbf{x} + \frac{\mathbf{x} - \mathbf{x_0}}{r}\alpha A(\theta) R(r) - \mathbf{x}\|_2^2 \tag{34}$$
$$= |\alpha A(\theta) R(r)|^2 \cdot \|\frac{\mathbf{x} - \mathbf{x_0}}{r}\|_2^2 \tag{35}$$
$$= |\alpha A(\theta) R(r)|^2 \tag{36}$$

To find the energy of a formlet, we integrate the displacement over r, θ space:

$$G = \int_{r=0}^{\infty}\int_{-\pi}^{\pi} |\alpha A(\theta) R(r)|^2 r d\theta dr \tag{37}$$
$$= \alpha^2 \int_{-\pi}^{\pi} A^2(\theta) d\theta \int_0^{\infty} R^2(r) r dr \tag{38}$$

To compute the integral over θ, we will make use of the property that the von-Mises pdf is normalized:

$$\int_{\theta=-\pi}^{\pi} \frac{1}{2\pi I_0(\kappa)} \exp(\kappa\cos\theta) = 1 \tag{39}$$

where $I_0(\kappa)$ is the modified Bessel function of order 0.

$$\int_{\theta=0}^{2\pi} A^2(\theta) d\theta = \int_{-\pi}^{\pi} (\exp(\kappa\cos\theta - \kappa))^2 d\theta \tag{40}$$
$$= e^{-2\kappa}\int_{-\pi}^{\pi} \exp(2\kappa\cos\theta) d\theta \tag{41}$$
$$= 2\pi I_0(2\kappa) e^{-2\kappa} \tag{42}$$

We find the integral over r numerically, using the substitution $u = \frac{r}{\sigma}$:

$$\int_{r=0}^{\infty} R^2(r) r dr = \dots$$
$$= \int_0^{\infty} \left|\exp\left(\frac{-r^2}{\sigma^2}\right)\sin\left(\frac{2\pi r}{\sigma}\right)\right|^2 r dr$$
$$= \int_0^{\infty} \left|\exp(-u^2)\sin(2\pi u)\right|^2 u\sigma^2 du \tag{43}$$

Using an adaptive Simpson quadrature method, we find that $\int_{r=0}^{\infty} R^2(r) r dr \approx 0.1284\sigma^2$

Substituting these expressions back into equation (38), we arrive at the following expression for the energy of a formlet:

$$G = \alpha^2 \int_{-\pi}^{\pi} A^2(\theta) d\theta \int_0^{\infty} R^2(r) dr \tag{44}$$
$$\approx \alpha^2 \cdot 2\pi I_0(2\kappa) e^{-2\kappa} \cdot 0.1284\sigma^2 \tag{45}$$
$$= C\alpha^2\sigma^2 I_0(2\kappa) e^{-2\kappa} \tag{46}$$

where $C \approx 0.2568\pi$

C. Optimal gain computation

Derivation of the optimal gain for the isotropic case can be found in [10], [3]. The derivation for oriented formlets is similar; we included it here for completeness.

Here we derive the optimal gain α^* given the remaining formlet parameters $p = \{x_0, y_0, \sigma, \kappa, \theta\}$. The regularized objective function in equation 15 is a quadratic function of α. Therefore, the optimal gain can be computed by differentiating the objective function with respect to α and setting to zero:

$$0 = \frac{\partial}{\partial\alpha} E(\Gamma^{obs}, f(\Gamma^{k-1}; p)) \tag{47}$$
$$= \frac{\partial}{\partial\alpha}\left[F(\Gamma^{obs}, f(\Gamma^{k-1}; p)) + \lambda G(p)\right] \tag{48}$$
$$= \frac{\partial}{\partial\alpha}\left[L\|f(\Gamma^{k-1}; p) - \Gamma^{obs}\|_2^2 + \lambda C\alpha^2\sigma^2 I_0(2\kappa)e^{-2\kappa}\right] \tag{49}$$

First, note that we can express the residual as:

$$\|\Gamma^{obs} - f(\Gamma^{k-1}; p)\|_2^2 = \|\Gamma^{res} - \alpha g(\Gamma^{k-1} - \mathbf{x_0})\|_2^2$$
$$= \|\Gamma^{res} - \alpha g(\Gamma^{k-1} - \mathbf{x_0})\|_2^2$$
$$= \|\Gamma^{res}\|_2^2 + \alpha^2\|g\|_2^2 - 2\alpha\langle\Gamma^{res}, g\rangle \tag{50}$$

where $\Gamma^{res} = \Gamma^{obs} - \Gamma^{k-1}$. Therefore,

$$\begin{aligned} 0 &= \frac{\partial}{\partial \alpha}(L\|\Gamma^{res}\|_2^2 + L\alpha^2\|g\|_2^2 - 2L\alpha\left\langle\Gamma^{res}, g\right\rangle + \dots \\ &\qquad + \lambda C\alpha^2\sigma^2 I_0(2\kappa)e^{-2\kappa}) \\ &= 2\alpha L\|g\|_2^2 - 2L\left\langle\Gamma^{res}, g\right\rangle + 2\lambda C\alpha\sigma^2 I_0(2\kappa)e^{-2\kappa} \\ &= \alpha\left[\|g\|_2^2 + \frac{1}{\lambda L}C\sigma^2 I_0(2\kappa)e^{-2\kappa}\right] - \left\langle\Gamma^{res}, g\right\rangle \quad (51) \end{aligned}$$

We arrive at the following expression for the optimal gain:

$$\alpha^* = \frac{\left\langle\Gamma^{res}, g\right\rangle}{\|g\|_2^2 + \frac{1}{\lambda L}C\sigma^2 I_0(2\kappa)e^{-2\kappa}} \quad (52)$$

REFERENCES

[1] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *PAMI, IEEE Transactions on*, 24(4):509–522, 2002.

[2] Alexander Dubinskiy and Song Chun Zhu. A multi-scale generative model for animate shapes and parts. *Proceedings of the 9th IEEE ICCV*, 1:249–256, October 2003.

[3] James H Elder, Timothy D Oleskiw, Alex Yakubovich, and Gabriel Peyré. On growth and formlets: Sparse multi-scale coding of planar shape. *Image and Vision Computing*, 31:1–13, 2013.

[4] U. Grenander, A. Srivastava, and S. Saini. A pattern-theoretic characterization of biological growth. *Medical Imaging, IEEE Trans.*, 26(5):648–659, 2007.

[5] Ulf Grenander, Anuj Srivastava, and Sanjay Saini. Characterization of biological growth using iterated diffeomorphisms. In *ISBI*, pages 1136–1139. IEEE, 2006.

[6] A.K. Jain, Y. Zhong, and S. Lakshmanan. Object matching using deformable templates. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(3):267–278, 1996.

[7] S.G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *Signal Processing, IEEE Transactions on*, 41(12):3397–3415, 1993.

[8] Calvin R Maurer Jr, Rensheng Qi, and Vijay Raghavan. A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions. *PAMI, IEEE Transactions on*, 25(2):265–270, 2003.

[9] Vida Movahedi and James H Elder. Design and perceptual validation of performance measures for salient object segmentation. In *CVPR Workshops, 2010 IEEE Computer Society Conference on*, pages 49–56. IEEE, 2010.

[10] T.D. Oleskiw, J.H Elder, and G. Peyré. On growth and formlets: Sparse multi-scale coding of planar shape. In *CVPR*, pages 459–466. IEEE, June 2010.

[11] Theodosios Pavlidis. Structural pattern recognition. *Springer Series in Electrophysics, Berlin: Springer, 1977*, 1, 1977.

[12] Nataliya Portman, Ulf Grenander, and Edward R Vrscay. Direct estimation of biological growth properties from image data using the "grid" model. In *Image Analysis and Recognition*, pages 832–843. Springer, 2009.

[13] E. Sharon and D. Mumford. 2d shape analysis using conformal mapping. In *CVPR*, volume 2, pages 350–357. IEEE, 2004.

[14] Darcy Wentworth Thompson. *On growth and form*. Cambridge Univ. Press, 1942.