

Bayesian Active Contours with Affine-Invariant, Elastic Shape Prior

Darshan Bryner and Anuj Srivastava
 Department of Statistics
 Florida State University, Tallahassee, FL
 {dbryner, anuj}@stat.fsu.edu

Abstract

Active contour, especially in conjunction with prior-shape models, has become an important tool in image segmentation. However, most contour methods use shape priors based on similarity-shape analysis, i.e. analysis that is invariant to rotation, translation, and scale. In practice, the training shapes used for prior-shape models may be collected from viewing angles different from those for the test images and require invariance to a larger class of transformation. Using an elastic, affine-invariant shape modeling of planar curves, we propose an active contour algorithm in which the training and test shapes can be at arbitrary affine transformations, and the resulting segmentation is robust to perspective skews. We construct a shape space of affine-standardized curves and derive a statistical model for capturing class-specific shape variability. The active contour is then driven by the true gradient of a total energy composed of a data term, a smoothing term, and an affine-invariant shape-prior term. This framework is demonstrated using a number of examples involving the segmentation of occluded or noisy images of targets subject to perspective skew.

1. Introduction

An object of interest in an image can be characterized to some extent by the shape of its external boundary. It is therefore important to develop procedures for boundary extraction in problems of detection, tracking, and classification of objects in images. Active contour algorithms have become an important tool in image segmentation for object detection [5, 6]. As segmentation algorithms become more sophisticated, they are tested in more difficult imaging environments of real-world scenarios where images do not have enough contrast to provide sharp boundaries, there is some occlusion of the target, or there exists target-like clutter or noise. Thus, it is of increasing importance that boundary extraction algorithms make use of prior knowledge about the expected target class in order to help compensate for the lack of clear data. This is accomplished by influencing

the contour evolution in part with a *shape prior*, a statistical model derived from a set of known training shapes, in a *Bayesian active contour* approach [12, 16, 8, 4].

Most of the past Bayesian segmentation methods use a shape prior designed to be invariant to similarity transformations of translation, rotation, and global scaling. However, in situations when the image plane of a camera is not parallel to the plane containing the defining part of the shape, perspective effects can transform the observed shapes in a more complicated manner than what can be modeled by the similarity group alone. The affine group is commonly used to approximate such shape deformations, and thus it is our goal to develop a segmentation algorithm that uses a shape prior built from affine-invariant shape statistics. In other words, *training and test shapes can be at random affine transformations from each other and the segmentation results will be invariant to those transformations*. Thus, our segmentation will be robust not only to poor image quality but also to perspective skews due to different viewing angles, either in test or training.

1.1. Past Work on Prior-Driven Active Contours

There are two broad categories of active contour methods: *parametric* methods that evolve an explicitly defined parameterized curve, and *geometric* methods that evolve implicitly defined zero-level sets of higher-dimensional functions. Due to the popularity and versatility of geometric methods, pioneered by the works of [13, 5] among others, most Bayesian methods have been applied in the geometric realm and follow the ideas presented in Leventon *et al.* [12], which uses PCA of level-set functions to form a shape prior. Tsai *et al.* in [16] incorporate a similar shape prior in an improved level-set segmentation framework given by Chan and Vese [6]. Others, e.g. [10], improve on Leventon's Gaussian shape prior by applying non-parametric density estimation techniques in \mathbb{L}^2 space. Yezzi and Soatto in [18] propose a shape prior based on an average shape that is invariant to any finite-dimensional group transformation, which includes the affine group.

There have been a few Bayesian active contour mod-

els that take a parametric approach. One such example, [9] makes use of “landmark-based” shape analysis [7] to impose a shape prior based on the statistics of similarity-invariant point sets. Recent advancements in the modeling of shapes as continuous curves, given by elastic shape analysis, e.g. [15], have allowed for more accurate and parsimonious shape models compared to those of [7, 12, 16, 10, 18]. Elastic shape analysis offers the important advantage of simultaneous registration and deformation of curves with an optimal combination of stretching and bending. Joshi *et al.* [8] create a shape prior from an intrinsic density on elastic shape space, but the method uses an older, more computationally expensive representation. The work of Bryner *et al.* [4] incorporates recent simplifications for elastic shape analysis provided in [15] for a computational speed-up. The works of [8] and [4] only formulate an intrinsic similarity-invariant shape prior rather than allowing for an affine-invariant, elastic shape model, which has been developed in [3] and [2].

1.2. Our Approach and Contributions

Our goal is to develop a method for *representing, modeling, and incorporating prior information* about shapes of closed curves, invariant to affine transformation and re-parameterization, in a parametric boundary extraction algorithm. Using the mathematical representation presented in [3] on affine-invariant elastic shape, we develop an intrinsic statistical model on the space of canonical, or affine-standardized, closed curves that will serve as a shape prior for the segmentation. Due to elastic matching of curves, this shape model captures the underlying shape variation of a shape class more accurately and leads to a more parsimonious shape model than its extrinsic counterparts often used in previous geometric Bayesian contour models. Furthermore, an invariance to affine transformation allows the model to be robust to perspective skews. With respect to many state-of-the-art Bayesian contour models that use intrinsic shape statistics (e.g. [8, 4]), we make one further key advancement to their approach aside from our novel, affine-invariant shape prior term. In our work we compute a true gradient descent flow for energy minimization [17]; that is, for each energy functional in the active contour model, we compute its gradient with respect to the same (\mathbb{L}^2) metric. Previous methods tend to mix gradients by computing the shape prior energy gradient with respect to the intrinsic shape metric and the remaining energy gradients with respect to the \mathbb{L}^2 metric [8, 4].

In summary, there are three important elements of our approach in this paper: (1) Bayesian active contours with elastic shape priors, (2) shape analysis that is invariant to both the affine and re-parameterization transformations, and (3) the concept of true gradient. After reviewing the current literature, we find that (i) while item (1) was introduced in

both [8] and [4], these papers did not have items (2) and (3); (ii) item (2) – elastic, affine-invariant shape analysis – was introduced in [3] and developed further in [2], but those papers did not have items (1) or (3); (iii) item (3) was proposed in [17] but without either (1) or (2). To reiterate, there is no paper currently in the literature that performs *even two of the three* items together, which underlines the novelty of our approach.

The organization of the rest of the paper is as follows. Section 2 reviews the affine-invariant, elastic shape analysis method from [3], and then develops the algorithmic tools to build intrinsic statistical shape models on the affine shape space. Section 3 describes the Bayesian active contour model, focusing on computation of the shape prior energy gradient with respect to the \mathbb{L}^2 metric. Section 4 shows a variety of experimental results that showcase the effectiveness of using our affine shape prior with \mathbb{L}^2 gradient compared to other methods. Section 5 is the conclusion and summary of future efforts.

2. Affine-Invariant, Elastic Shape Statistics

Here, we summarize the affine-invariant, elastic shape analysis method from [3] and then develop the procedures to form intrinsic statistical models on such a shape space.

2.1. Affine-Invariant, Elastic Shape Analysis

Let $\beta \in \mathcal{B}$ where \mathcal{B} is the set of all closed, parameterized, absolutely continuous curves. The action of the orientation-preserving affine group $G_a = GL_+(2) \times \mathbb{R}^2$ on β results in the orbits $[\beta] = \{A\beta + b | A \in GL_+(2), b \in \mathbb{R}^2\}$, where the group $GL_+(2)$ represents all matrices in $GL(2)$ with positive determinant. Let Γ be the set of all re-parameterizations of the form $\gamma : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ such that γ is a diffeomorphism. We wish to analyze the space of all equivalence classes $\mathcal{B}/(G_a \times \Gamma)$, yet, with respect to the standard \mathbb{L}^2 metric, the group $G_a \times \Gamma$ does not act on \mathcal{B} by isometries. Thus, it is not possible to impose a proper distance between equivalence classes and perform affine-invariant shape analysis of curves with this metric in this representation space. The solution proposed in [3] is to define a space M/G_0 with the following properties: (a) $M \subset \mathcal{B}$ and G_0 is a subgroup of $G_a \times \Gamma$, (b) there exists a bijection between M/G_0 and $\mathcal{B}/(G_a \times \Gamma)$, and (c) G_0 acts by isometries on M with respect to the chosen metric. Thus, a proper statistical analysis on M/G_0 is possible and represents, implicitly, a statistical analysis on $\mathcal{B}/(G_a \times \Gamma)$. The space M/G_0 is called a *section* of affine orbits.

The section is defined in the following manner. Let $L_\beta = \int_0^1 |\dot{\beta}(t)| dt$ be the length of the curve β , where $|\cdot|$ is Euclidean 2-norm. (Please contrast it from $\|\cdot\|$ which is used to denote the \mathbb{L}^2 -norm of a curve or function.) The *centroid* of β is defined as $C_\beta = \frac{1}{L_\beta} \int_0^1 \beta(t) |\dot{\beta}(t)| dt \in \mathbb{R}^2$.

The *covariance* of β is defined as $\Sigma_\beta = \frac{1}{L_\beta} \int_0^1 (\beta(t) - C_\beta)(\beta(t) - C_\beta)^T |\dot{\beta}(t)| dt \in \mathbb{R}^{2 \times 2}$. It can be shown that for any $\beta \in \mathcal{B}$ there exists a canonical, or standard, element $\beta_0 \in [\beta]$ that satisfies the following three conditions: (1) $L_{\beta_0} = 1$, (2) $C_{\beta_0} = 0$, and (3) $\Sigma_{\beta_0} \propto I$. Furthermore, for any two curves $\beta^{(1)}$ and $\beta^{(2)}$ within an affine transformation of each other, the corresponding standard elements, $\beta_0^{(1)}$ and $\beta_0^{(2)}$, are related by a rotation and a re-parameterization, i.e. $\beta_0^{(2)} = O(\beta_0^{(1)} \circ \gamma)$, where $O \in SO(2)$ and $\gamma \in \Gamma$. If \mathcal{B}_0 is the space of all such affine-standardized curves, then the quotient space $\mathcal{B}_0/(SO(2) \times \Gamma)$ satisfies properties (a) and (b) of a section as defined in the previous paragraph. Property (c) is not yet satisfied since the group Γ does not act on \mathcal{B}_0 by isometries with respect to the \mathbb{L}^2 metric.

In order to achieve the isometry property (c), we make the following transformation as given by [15]. Define $q(t) = \dot{\beta}(t)/\sqrt{|\dot{\beta}(t)|}$ as the square-root velocity function (SRVF) of β . The action of a $\gamma \in \Gamma$ on q is given by $(q, \gamma) = (q \circ \gamma)\sqrt{\dot{\gamma}}$, and now the group Γ acts by isometries on the space of SRVF's with respect to the \mathbb{L}^2 metric. Therefore, if \mathcal{Q}_0 is the set of all SRVF's of \mathcal{B}_0 , the space $\mathcal{Q}_0/(SO(2) \times \Gamma)$ is in one-to-one correspondence with $\mathcal{B}_0/(SO(2) \times \Gamma)$ and satisfies all three properties of a section. Furthermore, the paper [15] shows that the \mathbb{L}^2 metric of SRVF's is equivalent to the elastic metric of curves. Hence, an analysis on $\mathcal{Q}_0/(SO(2) \times \Gamma)$ is equivalent to that of an affine-invariant analysis on $\mathcal{B}/(G_a \times \Gamma)$, and it is in fact an elastic shape analysis framework with respect to the standard \mathbb{L}^2 metric. We denote $\mathcal{S} = \mathcal{Q}_0/(SO(2) \times \Gamma)$ as affine-invariant, elastic shape space.

Fig. 1 provides an illustration of this affine standardization of curves. The image on the left shows four rows of curves in \mathcal{B} , where objects in each row are within the same affine orbit, and the image on the right shows the standardized versions in \mathcal{B}_0 of each respective curve in the left image. Within each row on the right, the curves are all the same modulo rotation and re-parameterization.

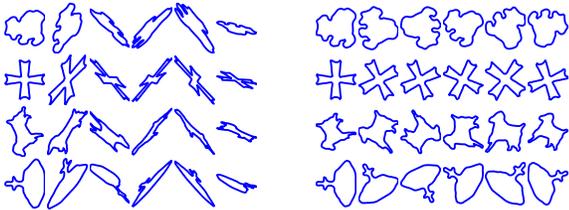


Figure 1. Affine standardization example.

2.2. Statistical Modeling

In order to build intrinsic statistical models on \mathcal{S} , one must first develop a set of algorithmic tools that use the geometry of \mathcal{Q}_0 to compute pertinent statistical quantities.

Bryner *et al* [3] describe geometry of \mathcal{Q}_0 , and we will not repeat it here. Instead, we discuss the algorithms and sub-routines necessary to compute an intrinsic sample mean and covariance on \mathcal{S} . These two statistics are then used to define a Gaussian-type probability model of elastic shapes invariant to affine transformation and re-parameterization. This probability model is in turn used as a shape prior in our Bayesian active contour framework, described later.

Statistical modeling on \mathcal{S} requires the use of two algorithms, one for mean calculation and one for covariance calculation. Both make use of five subroutines that each require as input a basis for $N_q(\mathcal{Q}_0)$, the normal space of \mathcal{Q}_0 at any point q . The five subroutines are (1) Projection onto the manifold, (2) Projection onto a tangent space, (3) Parallel translation of a tangent vector, (4) Exponential mapping, and (5) Inverse exponential mapping. An expression for the basis of $N_q(\mathcal{Q}_0)$ as well as descriptions of Subroutines (1)–(3) appear in [3]. Now, we provide Subroutines 4 and 5.

Given an element $[q] \in \mathcal{S}$ and the shooting vector $v \in T_{[q]}(\mathcal{S})$, the exponential mapping computes a point $[p] = \exp_{[q]}(v)$ in \mathcal{S} that represents the point that is reached by traveling along a constant-speed geodesic starting at $[q]$, and with the initial velocity v . Subroutines 4 and 5 make use of the fact that $\mathcal{Q}_0 \subset \mathbb{S}^\infty$, the unit hypersphere in \mathbb{L}^2 space, which has well-known analytical formulas for the exponential and inverse exponential mappings.

Subroutine 4 – Exponential Map: Given $[q] \in \mathcal{S}$, $v \in T_{[q]}(\mathcal{S})$, an integer n , and $\epsilon > 0$,
(i) If $\|v\| < \epsilon$, return $[p] = [q]$, else
(ii) Let $\delta = 1/n$. For $i = 1, \dots, n$
(a) Compute $q_{adv} = \exp_q(\delta v)$ on \mathbb{S}^∞ via the formula $q_{adv} = \cos(\delta\|v\|)q + \sin(\delta\|v\|)\frac{v}{\|v\|}$.
(b) Project $[q_{adv}]$ to \mathcal{S} using Subroutine 1.
(c) Parallel translate v from $T_{[q]}(\mathcal{S})$ to $T_{[q_{adv}]}(\mathcal{S})$ using Subroutine 3. Let $q = q_{adv}$.
(iii) Return $[p] = [q]$.

The opposite of the exponential map, the inverse exponential map, computes a vector $v = \exp_{[q]}^{-1}([p])$ in $T_{[q]}(\mathcal{S})$ that represents the shooting vector that satisfies $[p] = \exp_{[q]}(v)$ given $[p], [q] \in \mathcal{S}$. Since we wish to compute statistics modulo rotation and re-parameterization on the quotient space \mathcal{S} , when computing this shooting vector v , we must select either q or p and optimally rotate and re-parameterize it to the other. This registration is accomplished using a combination of Procrustes rigid body alignment and dynamic programming (see [15]).

Subroutine 5 – Inverse Exponential Map: Given $[q], [p] \in \mathcal{S}$ and $\epsilon > 0$,

(i) Optimally rotate/register WLOG q to p .
(ii) Compute the arclength $\theta = \cos^{-1}(\langle q, p \rangle)$.
(iii) If $\theta < \epsilon$, let $v = 0$, else compute $v = \exp_q^{-1}(p)$ in $T_q(\mathbb{S}^\infty)$ via the formula $v = \frac{\theta}{\sin(\theta)}(q - p \cos(\theta))$.
(iv) Project v to $T_{[q]}(\mathcal{S})$ via Subroutine 2.

Now we are ready to present the algorithms to compute the mean $[\mu]$ and the covariance K of a set of n shapes $\{[q_i]\}$ in \mathcal{S} . A popular intrinsic mean calculation is the Karcher mean, which is defined as $[\mu] = \operatorname{argmin}_{[q] \in \mathcal{S}} \sum_{i=1}^n d_{\mathcal{S}}([q], [q_i])^2$, where $d_{\mathcal{S}}(\cdot, \cdot)$ is the geodesic distance on shape space. An iterative algorithm to find the Karcher mean of a set of shapes is outlined below. The general idea is to update the current estimate $[\mu_j]$ in the direction of the average shooting vector from $[\mu_j]$ to each of data points $\{[q_i]\}$.

Algorithm (Karcher Mean): Let $[\mu_0] \in \mathcal{S}$ be an initial estimate of the mean of $\{[q_i]\}$, e.g. let $[\mu_0] = [q_1]$. Set $j = 0$.

(i) For each $i = 1, \dots, n$, register/rotate q_i to μ_j , and compute $v_i = \exp_{[\mu_j]}^{-1}([q_i])$ using Subroutine 5.

(ii) Compute the average direction $\bar{v} = \frac{1}{n} \sum_{i=1}^n v_i$.

(iii) If $\|\bar{v}\|$ is small, stop. Else, update $[\mu_j]$ by $[\mu_{j+1}] = \exp_{[\mu_j]}(\delta \bar{v})$ via Subroutine 4, where $\delta \approx 0.5$.

(iv) Set $j = j + 1$ and return to step 1.

Once we have found a Karcher mean $[\mu]$, we obtain the Karcher covariance matrix via $K = \frac{1}{n-1} \sum_{i=1}^n v_i v_i^T$, where the v_i 's are shooting vectors from μ to the respective q_i 's, each optimally rotated and registered to μ . While in theory $v : [0, 1] \rightarrow \mathbb{R}^2$ is a vector valued function, in practice it is computed using T equally spaced samples on the interval $[0, 1]$. Therefore, $v \in \mathbb{R}^{2 \times T}$ or re-arranged to be $\mathbb{R}^{1 \times 2T}$ and K is a $2T \times 2T$ covariance matrix.

Algorithm (Karcher Covariance): Given a set of shapes $\{[q_i]\}$ and its Karcher mean $[\mu]$,

(i) For $i = 1, \dots, n$, register/rotate q_i to μ , and calculate the shooting vector $v_i = \exp_{[\mu]}^{-1}([q_i])$ using Subroutine 5.

(ii) Compute $K = \frac{1}{n-1} \sum_{i=1}^n v_i v_i^T$.

Now that we have the tools to compute a sample mean and covariance of data on \mathcal{S} , we can speak of defining a probability density function from the shape class $(\{[q_i]\}, \mu, K)$. There are many densities one can define on \mathcal{S} from $(\{[q_i]\}, \mu, K)$, but for this research we only consider a truncated wrapped-normal density [11], which is formed as follows. First, obtain the singular value decomposition of K as $[U, S, V] = \operatorname{svd}(K)$, and let U_m be the m -dimensional principal subspace of $T_{[\mu]}(\mathcal{S})$ defined as the first m columns of U . The truncated wrapped-normal density is given as

$$f_m([q]) = \frac{1}{Z} e^{-\frac{1}{2}(v_{\parallel}^T S_m^{-1} v_{\parallel} + \|v_{\perp}\|^2 / \delta^2)} J_{[\mu]} \mathbf{1}_{\|v\| < \pi}, \quad (1)$$

where $v = \exp_{[\mu]}^{-1}([q])$, $v_{\parallel} = U_m^T v$ is the projection of v into U_m , $v_{\perp} = v - U_m v_{\parallel}$, S_m is the diagonal matrix containing the first m singular values, $J_{[\mu]}$ is the Jacobian of the exponential mapping, and Z is the normalizing constant. The scalar value δ is chosen to be less than the smallest singular value in S_m . In other words, this density is defined as a multivariate Gaussian density on $U_m \subset T_{[\mu]}(\mathcal{S})$ wrapped onto the manifold via the exponential mapping.

After computing the density in Eqn. 1, it is rather

straightforward to randomly sample from it. Since it is not possible to visualize the density function itself, in order to illustrate the shape variation explained in the density, we show a number of random samples instead. Fig. 2 displays four shape models in this manner formed from the same training data – 20 crown shapes from the well-known MPEG-7 database – in different shape spaces. Applied to each shape is a random orientation-preserving affine transformation, and the resulting curves are then sampled to have uniform speed parameterization. The shape spaces from top to bottom are the level-set similarity-invariant space [12], landmark similarity space [7], elastic similarity space [15], landmark affine space [1], and elastic affine space. The samples in the two affine cases are displayed in two stages. The left set of samples are in standard form, created in affine space, while the right set are the same samples but with a random orientation-preserving affine transformation re-applied to them. In this manner we separate the shape variability and affine variability when forming random samples.

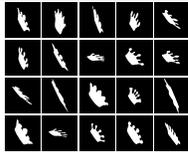
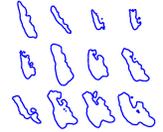
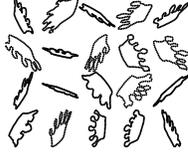
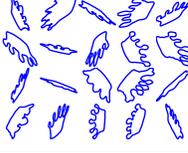
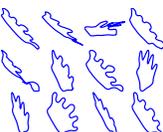
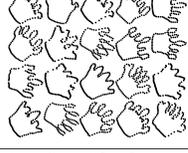
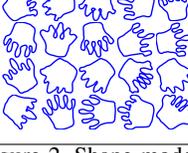
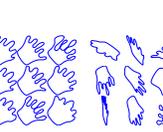
Data	Mean	Random Samples
		
		
		
		
		

Figure 2. Shape models from top to bottom: [12], [7], [15], [1], and our elastic affine-invariant model.

From Fig. 2, one can see that the ability of the elastic shape models to describe the underlying shape variation in a complicated shape class is superior to the non-elastic ones [12, 7, 1]. The mean shapes in the non-elastic models wash

out the defining crown features, such as the five points and the flat bottom, while the elastic models do not. Furthermore, the random samples from all non-elastic models as well as the elastic similarity model are not representative of the training data. The elastic affine model is more accurate and parsimonious than the elastic similarity model since affine variability and shape variability are separated. Thus, our elastic affine model is the superior choice in general for a shape prior in a Bayesian active contour model.

3. Bayesian Active Contour Model

The problem of boundary extraction in our active contour model can be posed as a maximum *a posteriori* (MAP) estimation via the energy minimization

$$\hat{\beta} = \underset{\beta \in \mathcal{B}}{\operatorname{argmin}} (\lambda_1 E_{\text{image}}(\beta) + \lambda_2 E_{\text{smooth}}(\beta) + \lambda_3 E_{\text{prior}}(\beta)), \quad (2)$$

where the λ_i 's are user-defined constant weights. The terms E_{image} and E_{smooth} are common energy terms based on given image pixel data and the smoothness of the evolving boundary curve, respectively, while the term E_{prior} is the novel term representing the shape prior energy functional built from a wrapped-normal density on \mathcal{S} . If E_{total} represents the weighted sum of all the energy functionals in the model, the active contour evolution is given by the gradient descent iteration $\beta_{n+1} = \beta_n - \epsilon \nabla E_{\text{total}}$, where $\epsilon > 0$ is a step size selected to maintain numerical stability.

Our research focuses on the influence of E_{prior} on active contour evolution, and thus the formulation of E_{image} and E_{smooth} are less important for our overall exposition. For this reason we have selected the relatively simple and well-known E_{image} term based on ‘‘region competition’’ and E_{smooth} term based on curvature smoothing. The explicit formula and reference for each of these functionals are provided in [4]. The important point here is that E_{image} and E_{smooth} are functionals of the curve β , and their gradients ∇E_{image} and ∇E_{smooth} are taken with respect to the \mathbb{L}^2 metric of curves. Thus, in order for our active contour evolution to be a true gradient descent flow [17], we must take the gradient ∇E_{prior} also with respect to the \mathbb{L}^2 metric of curves.

Given a prior shape class $(\{[q_i]\}, [\mu], K)$ in \mathcal{S} , $E_{\text{prior}} : \mathcal{B} \rightarrow \mathbb{R}$ is defined as such:

$$E_{\text{prior}}(\beta) = \frac{1}{2} v^T (U_m S_m^{-1} U_m^T) v + \frac{1}{2\delta^2} \|v - U_m U_m^T v\|^2, \quad (3)$$

where all terms are defined in the description of Eqn. 1. Thus, minimizing E_{prior} corresponds to maximizing the log-likelihood of the prior shape density. The global minimizer of this functional is the curve representation of the mean $[\mu]$. Note that $v = \exp_{[\mu]}^{-1}([q])$, where q is the SRVF of the standardized curve $\beta_0 \in [\beta]$. Even though the cal-

ulation of E_{prior} is based on an elastic shape distance between SRVF's, we ultimately treat it as any black box functional on \mathbb{L}^2 space. A numerical technique to approximate the gradient of such a functional is given as follows. Select an orthonormal basis for $T_{\beta}(\mathcal{B})$, say $\{b_i, i = 1, 2, \dots\}$. The standard Fourier basis functions that are periodic on $[0, 1]$ serve as a basis for this \mathbb{L}^2 space. After truncating to the first N basis elements for practical implementation, a first order numerical approximation is given as

$$\nabla E_{\text{prior}}(\beta) \approx \sum_{i=1}^N \frac{E_{\text{prior}}(\beta + \epsilon b_i) - E_{\text{prior}}(\beta)}{\epsilon} b_i, \quad (4)$$

where $\epsilon > 0$ is sufficiently small. The full algorithm for computing $\nabla E_{\text{prior}}(\beta)$ is given below.

Algorithm (∇E_{prior} Calculation): Given a curve $\beta \in \mathcal{B}$,
(i) Standardize β to \mathcal{B}_0 and convert to SRVF representation to obtain $[q] \in \mathcal{S}$.
(ii) Register/rotate μ to q to obtain $\mu^* = O^*(\mu, \gamma^*)$, and calculate $v = \exp_{[\mu]}^{-1}([q])$ via Subroutine 5. Calculate $E_{\text{prior}}(\beta)$ via Eqn. 3.
(iii) For each $i = 1, \dots, N$,
(a) Compute q_i , the SRVF of $\beta + \epsilon b_i$.
(b) Register/rotate μ to q_i using O^* and γ^* from step 2, and approximate $v = \exp_{[\mu]}^{-1}([q_i])$ via Subroutine 5. Calculate $E_{\text{prior}}(\beta + \epsilon b_i)$ via Eqn. 3.
(iv) Compute $\nabla E_{\text{prior}}(\beta)$ via Eqn. 4.

Note that in step (iii-b) above, we compute an approximation of the term $E_{\text{prior}}(\beta + \epsilon b_i)$ since we do not require for each i an optimization over $SO(2) \times \Gamma$. This would be quite expensive computationally. The approximation is valid since we assume ϵ small enough, *i.e.* a small enough perturbation of β , that the values O^* and γ^* obtained from optimizing μ to q can be used for optimizing μ to q_i .

This completes our calculation of the shape prior gradient with respect to the \mathbb{L}^2 metric. Our formulation advances the works of [8, 4] because in each of these papers, the authors compute ∇E_{prior} with respect to the elastic metric on shape space, which is inconsistent with the remaining energy functional gradients.

4. Experimental Results

Here, we evaluate the segmentation performance of multiple active contour models on various datasets. The contour models use all possible combinations of the following tools to formulate E_{prior} and its gradient: similarity-invariant shape statistics (from [15]), affine-invariant shape statistics (developed here as an extension of [3]), elastic gradient (from [8]), and \mathbb{L}^2 gradient (developed here). The E_{prior} scenarios are thus given as (1) no shape prior, (2) similarity-invariant with elastic gradient, (3) similarity-invariant with \mathbb{L}^2 gradient, (4) affine-invariant with elastic gradient, and

(5) affine-invariant with \mathbb{L}^2 gradient. Scenario (2) is exactly the model presented in [4], while scenarios (3)–(5) are novel to this paper. It is our proposition that scenario (5) will yield the best segmentation results due to its robustness to perspective effects as well as consistency of gradients.

In order to evaluate the accuracy of any segmentation result, we compare the converged contour to the associated ground truth curve via two metrics: $d_{geod}(\cdot, \cdot)$ and $d_{bin}(\cdot, \cdot)$. The distance d_{geod} is the geodesic distance on similarity-invariant, elastic shape space [15]. The distance d_{bin} is a binary image metric that measures the area of non-overlapping regions and is defined in the following manner. If \hat{B} is the binary image obtained by the segmentation and B is the ground truth binary image, then we define $d_{bin}(\hat{B}, B) = \text{area}(\hat{B} \cup B - \hat{B} \cap B) / \text{area}(\hat{B} \cup B)$. The values of these two metrics together show how accurately our segmentation result matches the correct shape as well as the correct location, orientation, and scale in the image. Note that $d_{geod} \in [0, \pi/2]$ and $d_{bin} \in [0, 1]$, and in each case a lower distance value corresponds to greater accuracy.

The average computational cost for 100 iterations of each of the five scenarios, as computed in Matlab on a 2.8 GHz processor, are given in seconds as (1) 0.367, (2) 20.8, (3) 26.0, (4) 38.9, and (5) 38.9. The affine cases are more complex than the similarity cases due to the necessity to standardize the active contour at each iteration. Most results shown in this section converged within 100–200 iterations with an initialization fairly close to the true boundary; thus, the computational complexity remains in the realm of practicality. Now, we present our segmentation results.

4.1. Multiview Curve Database (MCD)

The MCD [19] has been constructed from the MPEG-7 shape database. Here, a number of shapes were selected and printed on white paper as binary images, where the region enclosed by the shape was colored black. Variations of each shape were recorded by photographing the printed shapes under seven different camera angles. Since each shape in the MCD comes from one shape class in the MPEG-7 database, we use that shape class to build a shape prior for segmentation. Thus, we construct a scenario where the test image is of a different perspective than the training shapes and show the necessity of an affine-invariant shape prior for accurate segmentation. In the following experiment we introduce some occlusion to each skewed test image and segment under the five given E_{prior} scenarios.

Fig. 3 shows the segmentation results from two test images, and for each case we show eight images. From left to right, top to bottom, the eight images are as follows: the original image under centered camera view, the skewed and occluded test image, segmentation under scenarios (1)–(5) respectively, and the ground truth segmentation. Table 1 lists the values (d_{geod}, d_{bin}) for each of the five scenarios

averaged over the seven different test images (camera angles) for each shape. The segmentation is best in both instances under scenario (5), the affine prior with \mathbb{L}^2 gradient.

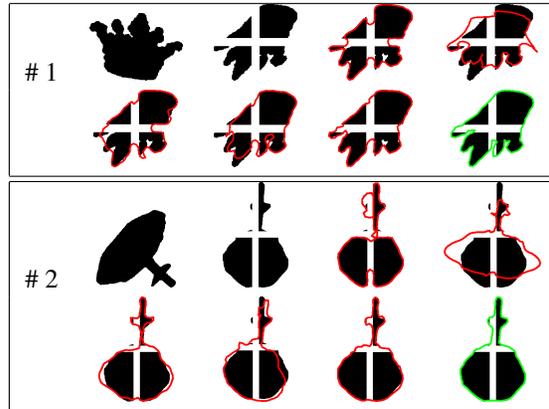


Figure 3. Segmentation results for occluded versions of the MCD shapes “crown” (camera angle 5) and “fountain” (camera angle 4).

Prior	Crown (# 1)	Fountain (# 2)
(1) None	(0.61, 0.17)	(0.60, 0.19)
(2) El. Sim	(0.37, 0.23)	(0.31, 0.30)
(3) \mathbb{L}^2 Sim	(0.40, 0.14)	(0.23, 0.12)
(4) El. Aff	(0.55, 0.16)	(0.33, 0.20)
(5) \mathbb{L}^2 Aff	(0.35, 0.10)	(0.17, 0.067)

Table 1. Averaged segmentation results of occluded MCD shapes.

4.2. Leaf Segmentation

Another useful application of our affine-invariant contour algorithm is the segmentation of leaves in images. In the upper-left image of Fig. 4 we see the ground truth boundary curves of 15 tulip poplar leaves that were extracted from images found in a Google Image search. Notice that in addition to the inherent leaf shape variability, this data additionally exhibits an approximate affine variability due to imaging from different camera angles. Fig. 4 shows the resulting similarity and affine invariant elastic shape models, and one can see from the random samples that the affine-invariant model eliminates the perspective variability while the similarity-invariant model does not.

Using the statistical models in Fig. 4 to build E_{prior} , we segment two test images of tulip poplar leaves in a somewhat noisy background with target-like clutter. Fig. 5 from left to right and top to bottom shows the test image and the segmentation results from scenarios (1)–(5) respectively. As in Table 1, Table 2 lists the pair (d_{geod}, d_{bin}) comparing the segmentation result to ground truth in each of the 5 scenarios for both of the test images. Again as predicted, scenario (5), the curve evolution using the \mathbb{L}^2 gradient of the affine-invariant E_{prior} , yields the best results.

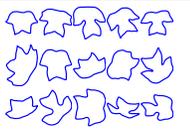
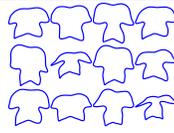
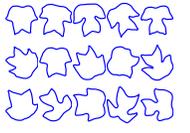
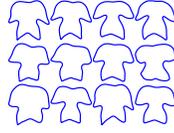
Data	Mean	Random Samples
		
		

Figure 4. Elastic shape models of the tulip poplar leaf. Top: Similarity-invariant. Bottom: Affine-invariant.

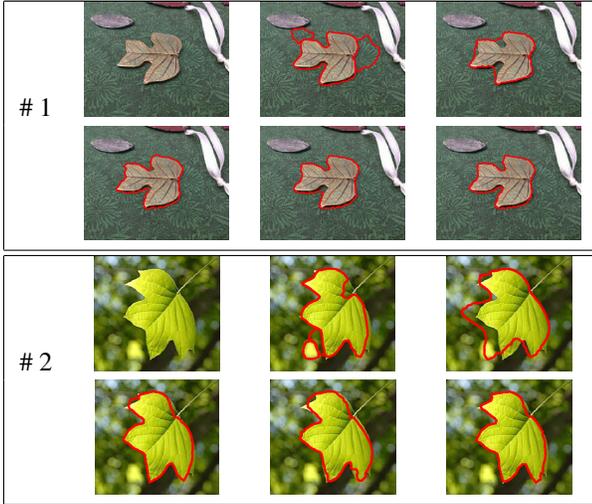


Figure 5. Segmentation results for tulip poplar leaves.

Prior	Leaf 1	Leaf 2
(1) None	(0.83, 0.32)	(0.52, 0.22)
(2) El. Sim	(0.35, 0.13)	(0.28, 0.25)
(3) L^2 Sim	(0.23, 0.11)	(0.26, 0.13)
(4) El. Aff	(0.22, 0.11)	(0.29, 0.14)
(5) L^2 Aff	(0.18, 0.096)	(0.21, 0.12)

Table 2. Segmentation results of the tulip poplar leaves.

4.3. SAS Shadow Segmentation

Here, we consider a dataset of imagery collected beyond the visible spectrum, where we segment the shadows of a cylinder target in synthetic aperture sonar (SAS) imagery. Segmentation is typically difficult in the synthetic aperture imaging modalities due to background noise, clutter, and imaging artifacts. The SAS images were created from the Shallow Water Acoustics Toolkit (SWAT), a program developed by the Naval Surface Warfare Center Panama City Division (NSWC PCD) that synthesizes SAS imagery of various targets in seabed environments [14]. The SWAT simulator is considered accurate to reality and is widely used to test automatic target detection and recognition algorithms in place of real SAS data. This particular dataset consists

of imagery of the same cylinder target at different aspect angles and ranges, which yields shadow signatures that exhibit a shape variability that can be modeled by an affine transformation. Fig. 6 shows an example of five images in the SAS dataset, and Fig. 7 shows the similarity and affine invariant shape models from 10 training shapes.

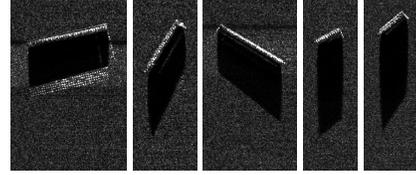


Figure 6. Example images from the PC SWAT cylinder database

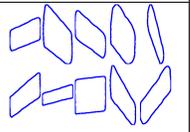
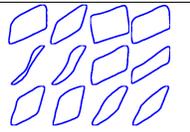
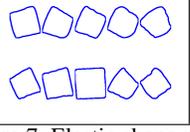
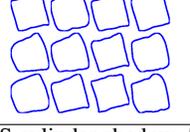
Data	Mean	Random Samples
		
		

Figure 7. Elastic shape models of the SAS cylinder shadow. Top: Similarity-invariant. Bottom: Affine-invariant.

Assuming that the ground truth shadow boundary curves are available, we perform a cross-validation experiment for each of the five scenarios. In each cross-validation iteration, we select 10 images at random for training, form the shape prior density on the appropriate shape space from the corresponding ground truth curves, and segment the remaining 90 test images with the influence of that shape prior. After each segmentation we calculate the values (d_{geod}, d_{bin}) to ground truth. The averages of these two distance values across all cross-validation iterations in each scenario are as follows: (1) (0.28, 0.14), (2) (0.21, 0.12), (3) (0.17, 0.099), (4) (0.18, 0.11), and (5) **(0.17, 0.090)**. Fig. 8 shows results from two test images. The six segmentations in each case shown from left to right and top to bottom are from scenarios (1)–(5), respectively, and ground truth. Although in many instances, segmentation with scenarios (1)–(4) yields acceptable results, there yet remain a few cases where it fails. Notably, segmentation with (4) fails often as a result of the segmentation flowing towards a different affine skew than ground truth (see case # 1 in Fig. 8). Segmentation with (5) corrects this issue due to its true gradient flow.

5. Conclusion

We present a Bayesian active contour model for image segmentation that improves the state-of-the-art in two key

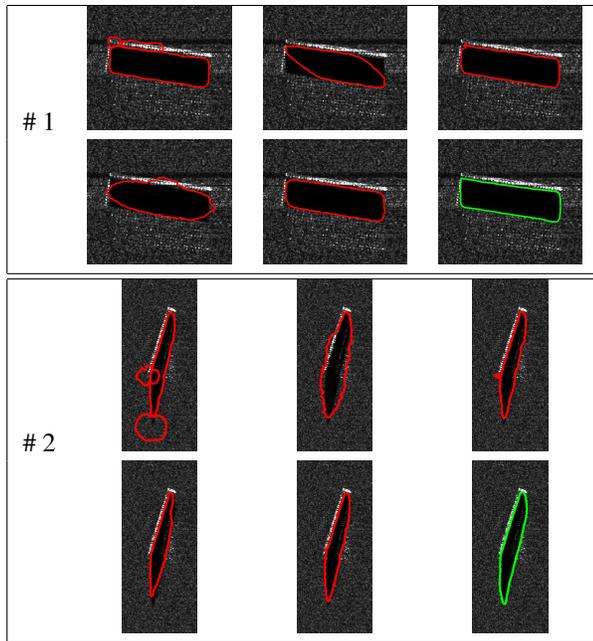


Figure 8. Segmentation results for the SAS database

aspects: (1) we use a shape prior term based on intrinsic affine-invariant, elastic shape statistics, and (2) we perform a true gradient descent flow to minimize the total energy functional. Elastic shape analysis allows us to build shape models that more accurately capture the underlying variation of complicated shape classes when compared to common extrinsic methods used in geometric active contours. Furthermore, an affine-invariant shape model is robust to perspective skew, allowing us to accurately segment when either test or training images are taken with respect to different camera angles. By computing the gradient of each of the three energy functionals – E_{image} , E_{smooth} , and E_{prior} – with respect to the same (\mathbb{L}^2) metric, the active contour evolution becomes a true gradient descent flow along E_{total} . With such a flow, segmentation results in an optimal fitting of both image and shape data. experimental results.

Acknowledgements

This work was supported in part by the grants ONR N00014-09-1-0665 and NSF DMS 09-15003 to AS. The authors would also like to acknowledge the NSWC PCD ILIR Program and the support of Dr. Quyen Huynh.

References

[1] E. Begelfor and M. Werman. Affine invariance revisited. In *CVPR*, volume 2, pages 2087–2094, 2006. 4

[2] D. Bryner, E. Klassen, H. Le, and A. Srivastava. 2D affine and projective shape analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99, 2013. 2

[3] D. Bryner, E. Klassen, and A. Srivastava. Affine-invariant, elastic shape analysis of planar contours. *CVPR*, pages 390–397, 2012. 2, 3, 5

[4] D. Bryner, A. Srivastava, and Q. Huynh. Elastic shapes models for improving segmentation of object boundaries in synthetic aperture sonar images. *Computer Vision and Image Understanding*, 2013. 1, 2, 5, 6

[5] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, Feb. 1997. 1

[6] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Trans. on Image Processing*, 10(2):266–277, Feb 2001. 1

[7] I. L. Dryden and K. V. Mardia. *Statistical Shape Analysis*. John Wiley & Son, 1998. 2, 4

[8] S. H. Joshi and A. Srivastava. Intrinsic bayesian active contours for extraction of object boundaries in images. *International Journal of Computer Vision*, 81(3):331–355, Mar. 2009. 1, 2, 5

[9] M. Kamandar and S. Seyedin. Procrustes - based shape prior for parametric active contours. In *International Conference on Machine Vision*, pages 135–140, Dec. 2007. 2

[10] J. Kim, M. Çetin, and A. S. Willsky. Nonparametric shape priors for active contour-based image segmentation. *Signal Processing*, 87(12):3021–3044, Dec. 2007. 1, 2

[11] S. Kurttek, A. Srivastava, E. Klassen, and Z. Ding. Statistical modeling of curves using shapes and related features. *Journal of American Statistical Association*, 107(499):1152–1165, 2012. 4

[12] M. E. Leventon, W. E. L. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *CVPR*, volume 1, pages 316–323, 2000. 1, 2, 4

[13] S. Osher and J. A. Sethian. Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988. 1

[14] G. Sammelmann, J. Christoff, and J. Lathrop. Synthetic images of proud targets. *MTS/IEEE Oceans Conference*, pages 1–6, Sept. 2006. 7

[15] A. Srivastava, E. Klassen, S. H. Joshi, and I. H. Jermyn. Shape analysis of elastic curves in Euclidean spaces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 33(7):1415–1428, Jul. 2011. 2, 3, 4, 5, 6

[16] A. Tsai, J. Yezzi, A., W. Wells, C. Tempny, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Trans. on Medical Imaging*, 22(2):137–154, 2003. 1, 2

[17] A. Yezzi and A. Mennucci. Conformal metrics and true “gradient flows” for curves. In *ICCV*, volume 1, pages 913–919, 2005. 2, 5

[18] A. Yezzi and S. Soatto. Deformation: Deforming motion, shape average and the joint registration and approximation of structures in images. *International Journal of Computer Vision*, 53(2):153–167, 2003. 1, 2

[19] M. Zuliani, S. Bhagavathy, B. S. Manjunath, and C. S. Kenney. Affine invariant curve matching. In *ICCV*, 2004. 6