

Shape Analysis on the Hypersphere of Wavelet Densities

Mark Moyou and Adrian M. Peter

Dept. of Engineering Systems, Florida Institute of Technology, Melbourne, FL, USA
mmoyou@my.fit.edu, apeter@fit.edu

Abstract

We present a novel method for shape analysis which represents shapes as probability density functions and then uses the intrinsic geometry of this space to match similar shapes. In our approach, shape densities are estimated by representing the square-root of the density in a wavelet basis. Under this model, each density (of a corresponding shape) is then mapped to a point on a unit hypersphere. For each category of shapes, we find the intrinsic Karcher mean of the class on the hypersphere of shape densities, and use the minimum spherical distance between a query shape and the means to classify shapes. Our method is adaptable to a variety of applications, does not require burdensome preprocessing like extracting closed curves, and experimental results demonstrate it to be competitive with contemporary shape matching algorithms.

1 Introduction

Shape analysis remains one of the most active research areas within computer vision, given that it plays a foundational role in the ultimate objective of general object recognition. Two key elements of shape analysis are robust representations and similarity metrics to distinguish between different shapes. We directly address these two areas by developing an efficient representation and matching procedure using wavelet expansions of density functions and the associated geometry of the space. We show this provides a rich framework that overcomes many of the drawbacks of contemporary approaches. The basic idea is to estimate density functions from shape point sets, where we use a wavelet basis expansion for the densities, and under such a representation each density becomes a point on a unit hypersphere. We then leverage the simple geometry of the hypersphere to compute intrinsic statistics on the manifold and compute shape similarity using the closed-formed distance between densities on the hypersphere. The method can easily be applied to either 2D or 3D

shapes, presently we develop the theory and detail results for 2D shape matching.

The almost de facto shape representation model for the majority of contemporary techniques is to use closed curves [3, 6, 5]. This topological restriction has mathematical advantages due to a plethora of theoretical results established for working with simple curves. However, one is constantly faced with the not so simple task of having to extract curves from imagery, where the presence of noise and occlusions severely degrade the quality of extracted contours. In addition, most complex shapes are stripped of their discriminatory qualities under such a representation. Often, additional interior features and/or disconnected regions can aid us in better distinguishing one shape from another.

We use the shape modeling approach that was first detailed in [11], which allows us to move away from the reliance of closed-curve representations. The shapes are simply represented as an unordered point sets. Given this, we estimate a probability density function $p(x)$ from the shapes. Rather than directly estimate $p(x)$, we instead estimate $\sqrt{p(x)}$ expanded in an orthogonal wavelet basis, recovering the desired density as $(\sqrt{p(x)})^2$. A key advantage afforded by using an orthogonal basis to expand the square-root of the density is that geometrically we can identify the density as a point on a unit hypersphere. Also, these shape densities visually resemble the original shape point set. Hence, if we match shape densities, we are in effect matching our original shape point sets. After describing this representation in section (2.1), we detail our matching approach that uses the intrinsic Karcher mean on hypersphere in section (2.2). Section (3) discusses promising experimental results on the MPEG-7 CE-Shape-1 part B database.

2 Theory

The idea of representing shapes as densities is usually brought to fruition in two ways. Either the density is directly estimated from the shape's discrete sam-

ples [1] or some other feature is first extracted from the shape and then the density is fit to these features [7]; this method falls in line with the former. We follow the wavelet density estimation (WDE) technique first presented in [11].

2.1 Square-root Wavelet Density Shape Representation

Many of the issues of estimating a bona fide density ($p(x) \geq 0$ and $\int p(x)dx = 1$) can be overcome by first estimating $\sqrt{p(x)}$ and then obtaining the desired density as $(\sqrt{p})^2$ [8]. For our purposes of 2D shape density estimation, the wavelet expansion of the square root of the density is given by

$$\sqrt{p(\mathbf{x})} = \sum_{j_0, \mathbf{k}} \alpha_{j_0, \mathbf{k}} \phi_{j_0, \mathbf{k}}(\mathbf{x}) + \sum_{j \geq j_0, \mathbf{k}} \sum_{w=1}^3 \beta_{j, \mathbf{k}}^w \psi_{j, \mathbf{k}}^w(\mathbf{x}) \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^2$, j_1 is some stopping scale level for the multiscale decomposition and $(k_1, k_2) = \mathbf{k} \in \mathbb{Z}^2$ is a multi-index that represents the spatial location of the basis. (The translation range of \mathbf{k} can be computed from the span of the data and basis function support size.) The father and mother basis are tensor product combinations of their one dimensional counterparts, i.e.

$$\begin{aligned} \phi_{j_0, \mathbf{k}}(\mathbf{x}) &= 2^{j_0} \phi(2^{j_0} x_1 - k_1) \phi(2^{j_0} x_2 - k_2) \\ \psi_{j, \mathbf{k}}^1(\mathbf{x}) &= 2^j \phi(2^j x_1 - k_1) \psi(2^j x_2 - k_2) \\ \psi_{j, \mathbf{k}}^2(\mathbf{x}) &= 2^j \psi(2^j x_1 - k_1) \phi(2^j x_2 - k_2) \\ \psi_{j, \mathbf{k}}^3(\mathbf{x}) &= 2^j \psi(2^j x_1 - k_1) \psi(2^j x_2 - k_2). \end{aligned} \quad (2)$$

The goal is to estimate the set of coefficients $\rho = \{\alpha_{j_0, \mathbf{k}}, \beta_{j, \mathbf{k}}^w\}$ and reconstruct the density using (1). An efficient maximum likelihood method to estimate them, with fast convergence, is discussed in [10]. Due to the increased indexing notation for two dimensional wavelet expansion, we will typically resort to one dimensional arguments, with it being understood that all results directly translate to two dimensions. Under an orthogonal wavelet expansion of $\sqrt{p(x)}$, the unit integrability requirement of all probability densities translates to a constraint on the wavelet coefficients

$$\int \left(\sqrt{p(x)} \right)^2 dx = \sum_{j_0, \mathbf{k}} \alpha_{j_0, \mathbf{k}}^2 + \sum_{j \geq j_0, \mathbf{k}} \beta_{j, \mathbf{k}}^2 = 1. \quad (3)$$

We consider only orthonormal bases such as Haar, Coiflets or Symlets. Figure 1 illustrates estimated densities for four point set shapes from the MPEG-7 database, using a single level wavelet decomposition (with only scaling functions). Notice that shapes exhibit a variety of topological properties like interior structures and disconnected components.

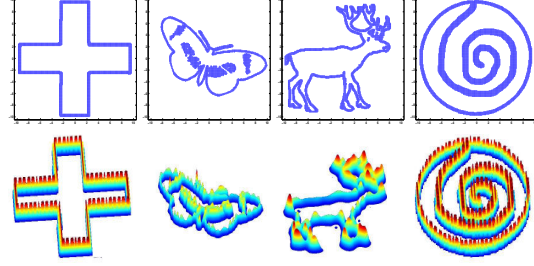


Figure 1. Example wavelet densities estimated from points-sets of MPEG-7 shapes. Top row are point sets, cardinality from left to right: 4,948;5,578;7,773;11,984. Second row is a perspective view of the estimated densities using the following wavelet families (from left to right): Haar ($j_0 = 2$), Coiflet-4 ($j_0 = 1$), Symlet-10 ($j_0 = 0$) and Haar ($j_0 = 2$). Notice how the wavelet densities accurately represent the shapes.

2.2 Shape Classification Using Intrinsic Means on the Hypersphere

Equation (3) showed that a natural by-product of working with the square root of the density and then expanding it with an orthonormal wavelet expansion was that it imposed a constraint on the basis coefficients; namely the sum of squared coefficient values must equal one. This immediately leads to the interpretation that the basis coefficients—which are unique to a particular density since wavelets serve as a true basis for the space of continuous distributions—give the coordinates for a position on the unit hypersphere S^{n-1} . The dimensionality of the hypersphere is determined by the cardinality of the set containing all the coefficients, i.e. for the set of coefficients ρ associated with a wavelet density, we have $\rho \in S^{n-1}$ where $n = |\rho|$.

Given many exemplars of a particular shape category, we can obtain a prototype representation of the category by computing the mean shape. In the proposed framework, this notion translates to computing a mean density function using the estimated wavelet densities for each of the point set exemplars in a category. Since the density functions are points on manifold, we must compute the generalized Karcher mean[4]. To do this intrinsically on the manifold, we utilize the associated Exponential and Log maps of the manifold, and implement the simple optimization procedure detailed in Algorithm 1 (more details can be found in [9]). In the present context, the Exponential map takes a vector γ on the tangent space at ρ_1 , $\gamma \in T_{\rho_1}(S^{n-1})$, and gives

Algorithm 1 Numerical computation of Karcher mean on manifold \mathcal{M} . For the present context $\mathcal{M} = S^m$, m -dimensional unit hypersphere, use the Exp and Log maps defined in (4) and (5), respectively. κ is a small step size parameter.

Input: $\rho_1, \rho_2, \dots, \rho_m \in \mathcal{M}$

Output: $\mu \in \mathcal{M}$

Let $\mu^0 = \rho_1$

While $\|\gamma^\tau - \gamma^{\tau-1}\| > \epsilon$

$$\gamma^\tau = \frac{\kappa}{m} \sum_{i=1}^m \text{Log}_{\mu^{\tau-1}}(\rho_i)$$

$$\mu^\tau = \text{Exp}_{\mu^{\tau-1}}(\gamma^\tau)$$

us a point on the hypersphere

$$\rho_2 = \text{Exp}_{\rho_1}(\gamma) = \cos(|\gamma|) \rho_1 + \sin(|\gamma|) \frac{\gamma}{|\gamma|}. \quad (4)$$

Conversely, the Log maps takes a point on the hypersphere ρ_2 and returns a vector on the tangent space at ρ_1 , by letting

$$\begin{aligned} \tilde{\rho} &= \rho_2 - \langle \rho_2, \rho_1 \rangle \rho_1 \\ \gamma = \text{Log}_{\rho_1}(\rho_2) &= \frac{\tilde{\rho} \cos^{-1}(\langle \rho_1, \rho_2 \rangle)}{\sqrt{\langle \tilde{\rho}, \tilde{\rho} \rangle}}. \end{aligned} \quad (5)$$

For a shape recognition problem where we have multiple categories of shapes, our classification approach is to compute a mean density for each of the classes using a random subset of the associated densities in each class. Then given a query shape, we estimate its wavelet density and then compute the distance from the query shape density to each of the class mean densities using the closed-form distance on the unit hypersphere

$$d(\rho, \mu_i) = \cos^{-1}(\rho^T \mu_i),$$

where ρ is a vectorized set of wavelet coefficients for the query shape density and μ_i is the set of coefficients for the mean shape density associated with the i th class. The query shape belongs to the class with the minimum distance. It is worth noting that all of our analysis is taking place intrinsically on the manifold of our shape representation, an added advantage over methodologies that decouple the representation and matching.

3 Experiments

We evaluated the proposed shape matching framework on the popular MPEG-7 CE-Shape-1 part B database [5]. It consists of 70 different categories with

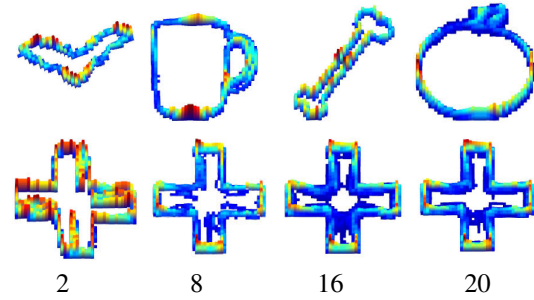


Figure 2. Top row: example mean shape densities estimated from point-sets of MPEG-7 shapes. Bottom row: effects due to the number of shapes (indicated under the image) used to compute the mean shape density for an example shape. Each mean is the intrinsic Karcher mean on the hypersphere of wavelet densities.

20 images per category for a total of 1400 binary images. Each image consists of a single shape. Each shape was represented with a subset of points. There are *no topology or equal point-set cardinality requirements* amongst shapes, allowing shapes with richer features to be represented with a greater number of points, see Figure 1 for some examples. Shapes within each category were rigid aligned to a category reference shape. Next, we estimated coefficients for the wavelet density of each shape using a Haar basis with $j_0 = 1$. This resulted in each shape density being represented with 1,764 coefficients. We then proceeded to compute a mean shape density for each category using Algorithm 1. In order to test the robustness of our approach, we randomly selected varying numbers of exemplars from each category to compute the mean, beginning with all 20 shapes per category down only 2 shapes per category. Finally, all 1400 shapes are treated as query shapes, for which we compute the distance between the query shape density and class mean shape densities. The class label for the mean shape density with the minimum distance is assigned to the query. Since the means are computed with randomly selected members from each category, we repeated the entire procedure 100 times to understand how our recognition performance varies under this random selection. Figure 2 illustrates some of examples of mean shape densities and the box plot in Figure 3 summarizes our classification results. Notice that our approach is very stable, with recognition rates exceeding 95% even when using less than half number of shapes per category to compute the mean.

Methods based on hierarchical representations [3, 6] have reported recognition rates above 85% on the MPEG-7 data set. They assume shapes are repre-

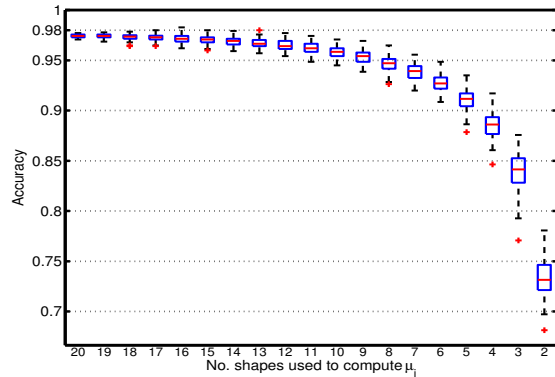


Figure 3. MPEG-7 classification accuracies.

sented by their boundary outlines and typically use less than 200 points for the shapes. These methods have the drawback of extracting oriented, boundary curves which can be a troublesome preprocessing procedure. We also lose the descriptive power afforded by allowing arbitrary shape topologies and unconstrained point set cardinalities. More recently, graph transduction approaches [12, 2] have been applied in conjunction with any existing similarity measure, yielding accuracy scores in the 92 – 99.99% range on MPEG-7. However, the underlying techniques again work with curve representation of the shapes and further more, these algorithms require one to build a $N \times N$ graph over the database of N shapes to the encode nearest-neighbor information. This can be prohibitive for large data sets. In [11], the authors also use wavelet shape densities; however, their approach differs in that they directly incorporate non-rigid alignment via “sliding” of the wavelet coefficients and also do not compute mean densities for each of shape categories. It also worth noting that the accuracies for these competing algorithms are computed using the bulls-eye criterion [5], which is a more forgiving accuracy measure compared to our strict minimum distance classification. The proposed method’s performance is on par with these, while maintaining the strong advantages of requiring less preprocessing, being easy to use, and is computationally efficient.

4 Conclusions and Discussion

We have presented a new shape matching framework using a square-root wavelet density representation of point-set shapes. This representation places each density on unit hypersphere, whose geometry we leverage to derive an intrinsic similarity measure using the Karcher mean on the density manifold. Our approach has several advantages over other contemporary shape modeling and matching schemes:

- Each shape can have an arbitrary number of points without topological restrictions. This is in sharp contrast to methods that work only on shape contours or are limited to only a few sample points.
- Limited preprocessing is required since we directly estimate the density from shape points.
- The similarity metric is in closed form and computationally efficient for large querying applications.
- The use of wavelet bases in our representation also supports sparse representations for our shape densities—something we plan to investigate in the future.

Our method was validated on a major shape database and demonstrated promising results which are competitive with the state of the art.

References

- [1] T. Chen, B. Vemuri, A. Rangarajan, and S. Eischen-schenck. Group-wise point-set registration using a novel CDF-based Havrda-Charvát divergence. *Intl. Journal of Comp. Vis.*, 86:111–124, 2010.
- [2] A. Egozi, Y. Keller, and H. Guterman. Improving shape retrieval by spectral matching and meta similarity. *IEEE Trans. on Image Proc.*, 19(5):1319–1327, 2010.
- [3] P. F. Felzenszwalb and J. D. Schwartz. Hierarchical matching of deformable shapes. In *IEEE Conf. on Comp. Vis. and Patt. Recog.*, pages 1–8, 2007.
- [4] H. Karcher. Riemannian center of mass and mollifier smoothing. *Commun. on Pure and Appl. Math.*, 30(5):509–541, 1977.
- [5] L. J. Latecki, R. Lakämper, and U. Eckhardt. Shape descriptors for non-rigid shapes with a single closed contour. In *IEEE Conf. on Comp. Vis. and Patt. Recog.*, pages 424–429, 2000.
- [6] G. McNeill and S. Vijayakumar. Hierarchical procrustes matching for shape retrieval. In *IEEE Conf. on Comp. Vis. and Patt. Recog.*, pages 885–894, 2006.
- [7] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin. Shape distributions. *ACM Trans. on Graphics*, (4):807–832, 2004.
- [8] S. Penev and L. Dechevsky. On non-negative wavelet-based density estimators. *Journal of Nonparametric Statistics*, 7:365–394, 1997.
- [9] X. Pennec. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *Journal of Math. Imaging and Vision*, 25(1):127–154, 2006.
- [10] A. Peter and A. Rangarajan. Maximum likelihood wavelet density estimation with applications to image and shape matching. *IEEE Trans. on Image Proc.*, 17(4):458–468, April 2008.
- [11] A. Peter, A. Rangarajan, and J. Ho. Shape L’Âne Rouge: Sliding wavelets for indexing and retrieval. In *IEEE Conf. on Comp. Vis. and Patt. Recog.*, pages 1–8, June 2008.
- [12] X. Yang and L. Latecki. Affinity learning on a tensor product graph with applications to shape and image retrieval. In *IEEE Conf. on Comp. Vis. and Patt. Recog.*, pages 2369–2376, 2011.