

AFFINE-INVARIANT CURVE MATCHING

Marco Zuliani, Sitaram Bhagavathy, B. S. Manjunath, and C. S. Kenney

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106.

{zuliani, sitaram, manj, kenney}@ece.ucsb.edu

ABSTRACT

In this paper, we propose an affine-invariant method for describing and matching curves. This is important since affine transformations are often used to model perspective distortions. More specifically, we propose a new definition of the shape of a curve that characterizes a curve independently of the effects introduced by affine distortions. By combining this definition with a rotation-invariant shape descriptor, we show how it is possible to describe a curve in an intrinsically affine-invariant manner. To validate our procedure we built a database of shapes subject to perspective distortions and plotted the precision-recall curve for this dataset. Finally an application of our method is shown in the context of wide baseline matching.

1 Introduction

It is trivial to state that shape is an important visual feature in computer vision. This naturally follows since human beings greatly rely on shape to recognize objects. Shape description techniques can be classified broadly into contour-based ones and region-based ones [1, 2]. In the former, only the contour of a region is used for extracting its shape features, whereas in the latter, the interior is also used. Several contour-based features have been proposed in the literature [1, 2]. All of these involve the description and matching of 2-D curves that represent shape contours. One of the persisting problems has been the efficient description and matching of 2-D curves in the presence of perspective distortion. Many existing methods are RST (rotation, scaling, and translation) invariant, but fail under perspective distortion.

Why is affine invariance important? The goal of this paper is to study the problem of matching curves between views related by perspective distortion. It is well known that a perspective transformation can be approximated by an affine transformation as long as the

This project was supported by the grants ONR #N00014-02-1-0121 (Zuliani and Manjunath), ONR #N00014-02-1-0318 (Kenney) and NSF-DLI #IIS-49817432 (Bhagavathy).

imaged object is planar and the camera optical center is far enough from such plane [3]. This explains why we want to develop a methodology to compare curves related by an affine transformation.

In this paper, we propose a method that will lead us to a compact and intrinsically affine invariant shape descriptor that is shown to be effective in matching curves that are affine equivalent. Previously, the curvature scale space (CSS) descriptor [4] of MPEG-7 has been shown to be robust under perspective distortion. However, this descriptor is not intrinsically affine invariant. In order to achieve robustness to affine transformation, it uses the *affine length* curve parametrization method that requires derivatives up to the fourth order. In a digital implementation, higher-order derivatives (a) are difficult to implement, and (b) may cause instability in the presence of noise. Our method merits from not using any derivatives.

2 Affine Invariant Curve Matching

2.1 Describing The Shape of a Curve

Consider a planar curve Γ represented by the set of points $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$. What is the shape of Γ ? Kendall [5] gave the following definition of shape:

[shape is] what is left of the configuration [of a set of points] after the effects of translation, rotation and scaling have been factored out.

Unfortunately this definition does not allow us to deal with curves that are subject to affine transformations, since in these cases it is necessary to factor out not only rotation, scaling and translation but also shear.

Consider a Jordan curve Γ that is the boundary of the domain $\Omega \subset \mathbb{R}^2$: to extend Kendall's notion of shape we introduce the following quantities:

- Let $V(\Omega) \stackrel{\text{def}}{=} \int_{\Omega} d\mathbf{x}^2$ be the area of Ω , where $d\mathbf{x}^2$ is the infinitesimal area element.
- Let $\mathbf{m}(\Omega) \stackrel{\text{def}}{=} \frac{1}{V(\Omega)} \int_{\Omega} \mathbf{x} d\mathbf{x}^2$ be the centroid of Ω .

- Let $\Sigma(\Omega) \stackrel{\text{def}}{=} \frac{1}{V(\Omega)} \int_{\Omega} [\mathbf{x} - \mathbf{m}(\Omega)] [\mathbf{x} - \mathbf{m}(\Omega)]^T d\mathbf{x}^2$ be the covariance of Ω .

We now have all the ingredients to define the shape of a Jordan curve:

Definition 1 Let Γ be a Jordan curve. The shape of Γ is defined as:

$$S(\Gamma) \stackrel{\text{def}}{=} \left\{ \mathbf{s} \in \mathbb{R}^2 : \mathbf{s} = \Sigma(\Omega)^{-\frac{1}{2}} [\mathbf{x} - \mathbf{m}(\Omega)] \text{ for } \mathbf{x} \in \Gamma \right\} \quad (1)$$

This definition is important because it allows us to relate affine-transformed curves, as stated in the following theorem:

Theorem 1 Let Γ_1 and Γ_2 be two Jordan curves related by an affine transformation:

$$\Gamma_2 = \left\{ \mathbf{x}_2 \in \mathbb{R}^2 : \exists \mathbf{x}_1 \in \Gamma_1 \text{ such that } \mathbf{x}_2 = A\mathbf{x}_1 + \mathbf{b} \right\}$$

where $A \in \mathbb{R}^{2 \times 2}$ is a non-singular matrix and $\mathbf{b} \in \mathbb{R}^2$. Then the shapes of Γ_1 and Γ_2 are geometrically congruent via a 2-dimensional rotation.

Sketch of the Proof. Let $\Gamma_1 = \partial\Omega_1$ and $\Gamma_2 = \partial\Omega_2$. We want to show that the matrix $R \stackrel{\text{def}}{=} \Sigma(\Omega_1)^{\frac{1}{2}} A^T \Sigma(\Omega_2)^{-\frac{1}{2}}$ establishes the congruence relation between $S(\Omega_1)$ and $S(\Omega_2)$. The first step consists in verifying that R is actually a rotation matrix. Then the proof is completed observing that (a) for any $\mathbf{s}_1 \in S(\Gamma_1)$ there exists $\mathbf{s}_2 \in S(\Gamma_2)$ such that $\mathbf{s}_1 = R\mathbf{s}_2$ and that (b) for any $\mathbf{s}_2 \in S(\Gamma_2)$ there exists $\mathbf{s}_1 \in S(\Gamma_1)$ such that $\mathbf{s}_2 = R^{-1}\mathbf{s}_1$. The details of the proof can be found in [6].

In other words we can say that the notion of shape we propose is unique modulo a rotation. Therefore, in order to describe a curve Γ in an affine-invariant manner, we only need to describe its shape $S(\Gamma)$ in a rotation-invariant manner. This is made feasible by the abundance of rotation-invariant shape descriptors.

2.2 A Modified Shape Matrix Descriptor

With regard to complexity, we choose not to use an intrinsically rotation-invariant shape descriptor. Instead, we obtain a partial invariance by using a combination of a shape matrix-based descriptor and the distance function used to compare curves. Because of quantization effects, we cannot achieve a true rotation-invariance with this descriptor. However, this descriptor allows us to represent a curve compactly and to easily cope with the problems posed by curve sampling and reflection.

Shape matrices have been used in the past to represent shapes in an RST-invariant manner. Goshtasby

[7] proposed a shape matrix built from a circular quantization grid. A polar raster of concentric circles and radial lines is positioned in the center of mass of the shape (Fig. 1(b)). The maximum radius of the shape is equal to the radius of the circle. Rotation invariance is achieved by aligning the axis of the coordinate system along the maximum radial line OA. The binary value of the shape is sampled at the intersections of the circles and radial lines. The shape matrix is formed so that the circles correspond to the columns and the radial lines correspond to the rows. Taza and Suen [8] compare two shapes by comparing the corresponding shape matrices. The matrix entries are weighted for more objective comparison because the sampling density is higher at the center of the circle than at the periphery. Parui et al. [9] implicitly impose this weighting by using the relative areas of the shape contained in concentric rings around the shape center of mass.

We use a variation of Goshtasby's method [7]. Given any curve Γ (Fig. 1(a)), we arrive at its shape $S(\Gamma)$ (Fig. 1(c)) using Definition 1. We then use a non-uniform circular grid (Fig. 1(d)) to build the shape matrix. We sample uniformly along the angular coordinates and non-uniformly along the radial lines so that the area of each sector is independent of its distance from the center. The grid is centered at the center of mass O of the shape and the maximum radius is equal to the distance of the furthest point of the shape from O. Note that we do not require the axis to be aligned along the maximum radial line.

Suppose the grid has N_R radial divisions and N_θ angular divisions. Using this grid, we build a binary matrix $D \in \{0, 1\}^{N_R \times N_\theta}$ where the entry $D_{i,j}$ is set to one if and only if the shape $S(\Gamma)$ intersects the sector corresponding to the i^{th} radial division and to the j^{th} angular division. The coordinate axis is arbitrary and the sampling direction is counter-clockwise. The matrix D can be stored using $N_R N_\theta$ bits. The distance¹ between two curve descriptors A and B is computed using the following function:

$$d(A, B) = \min_{-N_\theta+1 \leq i \leq N_\theta-1} \sum_{h=1}^{N_R} \sum_{k=1}^{N_\theta} A_{h,k} \text{ XOR } B_{h, < k-i >_{N_\theta}}$$

where the notation $< \cdot >_{N_\theta}$ denotes circular shift. The negative circular shift indices indicate a reversal in the sampling direction (clockwise) and thus account for reflections. The distance function counts the minimum number of different bits between A and all the possible column circular shifts of B . The larger is the number of angular divisions the greater is the degree of invariance

¹With a little abuse of language we will call d distance even though it does not always satisfy the triangle inequality.

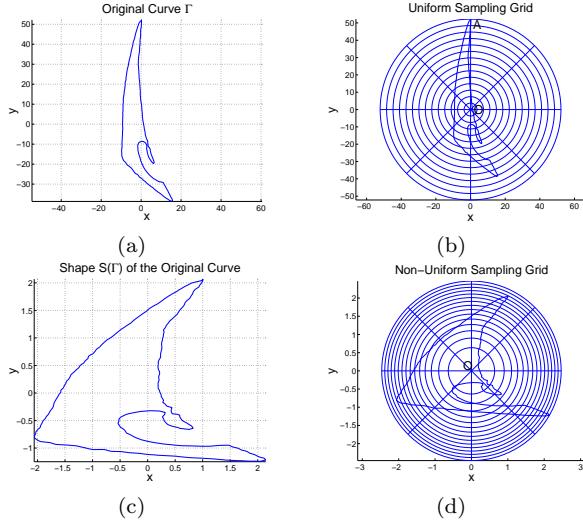


Fig. 1. (a) A closed curve; (b) uniform circular quantization grid; (c) The shape corresponding to the closed curve (having removed the effect of affine transformation); (d) non-uniform circular quantization grid (all sectors have the same area).

with respect to rotations and reflections. In our experiments we used a sampling grid with 12 radial divisions and 15 angular divisions.

3 Experimental Results

The MPEG-7 *Similarity* test dataset [10] has been widely used to evaluate the performance of shape descriptors. However, it does not contain many shapes with perspective distortion and hence is not suitable for evaluating affine-invariant descriptors. In order to fill this void, we created the Multiview Curve Dataset (MCD) to evaluate descriptors intended for comparing curves under perspective distortion.

3.1 The Multiview Curve Dataset

This dataset² comprises 40 shape categories, each corresponding to a shape drawn from an MPEG-7 shape category. Each category in the new dataset contains 14 curve samples that correspond to different perspective distortions of the original shape. The original MPEG-7 shapes were printed on white paper and 7 samples were taken using a digital camera from various angles. The contours were extracted from the iso-intensity level set decomposition of the images [11]. Adding random rotations and reflections to these samples, the number of samples in each category is doubled to 14. Fig. 2(a) shows the samples from the category “key”.

²To obtain this dataset please send an email to {sitaram, zuliani}@ece.ucsb.edu

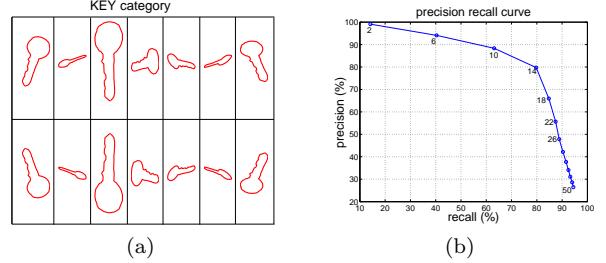


Fig. 2. (a) The samples for the category “key” in the Multiview Curve Dataset; (b) Precision-Recall over the MCD.

3.2 Precision vs. Recall Curve

The proposed method is evaluated by plotting a precision vs. recall curve over the above dataset. Each curve Q is used in turn as the query. Let $A(Q)$ denote the set of T retrievals (based on the smallest distances from Q in the descriptor space) and $R(Q)$, the set of 14 images in the dataset relevant to Q . The *precision* is defined by $P(Q) = \frac{|A(Q) \cap R(Q)|}{|A(Q)|}$, and the *recall* by $C(Q) = \frac{|A(Q) \cap R(Q)|}{|R(Q)|}$, where $|\cdot|$ denotes cardinality. Fig. 2(b) shows the precision-recall curve for the proposed descriptor. The curve is plotted by averaging precision and recall over all Q , for different values of T (shown below the curve).

When $T = 2$, the precision is higher than 99%. This means that in more than 99% of cases, the best match to the query (i.e. the second retrieval³) is from the same category. When $T = 14$, i.e. the number of retrievals is equal to the number of samples per category, the precision is about 80% at a recall of 80%. This suggests that our approach is effective in matching curves despite perspective distortion. In the future, we intend to use the MCD to compare our method with the CSS approach (using affine length curve parametrization) in terms of both performance under perspective distortion and complexity.

3.3 Finding Wide Baseline Correspondences

Establishing correspondences between images taken from different points of view is a central problem in computer vision. Our approach can be used to match planar curves extracted from image pairs related by a perspective distortion. Fig. 3 shows the best 15 matches between the set of curves (whose length is larger than a certain threshold) extracted from the “Graffiti” pair. Only two curves are mismatched. However, in this specific example such mismatches are due to the fact that not all the curves extracted from one image have their counterpart in the other image. Note also that the mismatched curves are very similar in shape.

³The first retrieval is always the query itself (zero distance).

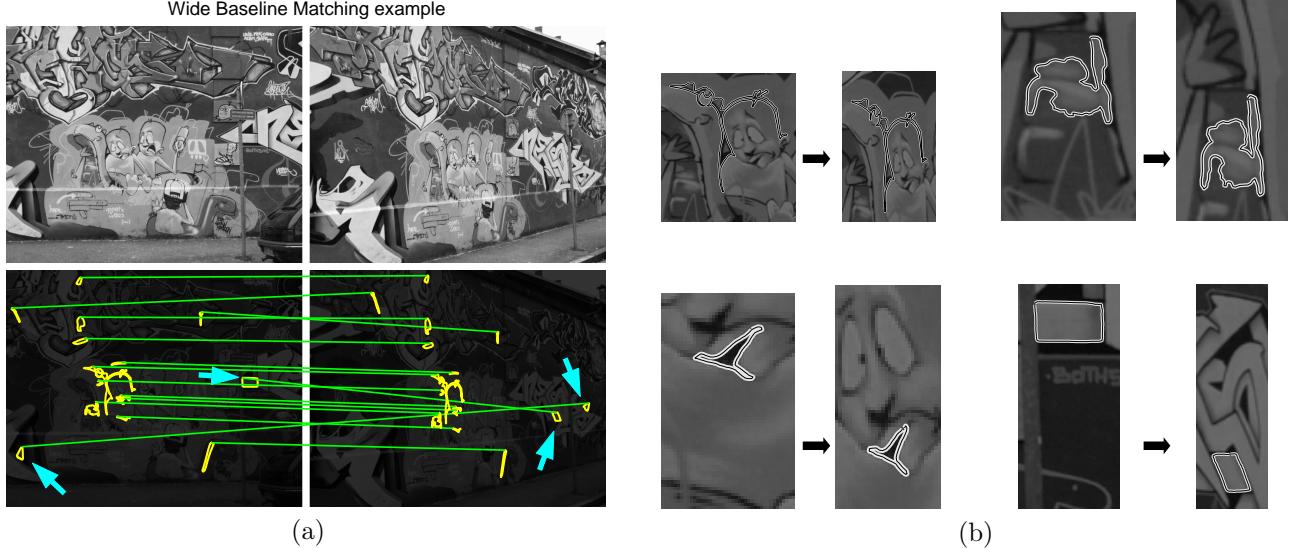


Fig. 3. (a) Two views of the “Graffiti” scene. The lines connect corresponding curves and the arrows indicate the curve mismatches; (b) Zoomed versions of some of the matches and one mismatch (bottom right).

4 Conclusions

We have proposed a method for describing and matching curves in an affine-invariant way. This is important because affine transformation is often used to model perspective distortion. The novelty of our approach is a new definition for the shape of a planar curve that is independent of affine transformations modulo rotation. We achieve affine-invariant curve matching by combining this concept with a modified shape matrix descriptor and a rotation-invariant distance function.

To evaluate our method, we built a Multiview Curve Dataset (MCD) consisting of curves under different perspective distortions. The retrieval performance was visualized as a precision-recall curve. Furthermore, we demonstrate the application of this method in the context of wide baseline matching.

In the future, we intend to make the approach truly affine-invariant by eliminating the quantization effects of the shape matrix descriptor. We plan to achieve this goal by combining our shape extraction approach with an intrinsically rotation-invariant descriptor such as the CSS descriptor. However, preliminary results indicate that the CSS descriptor⁴ does not improve the performance significantly.

Acknowledgement: We would like to thank prof. Shiv Chandrasekaran for helpful discussions, and Dr. Miroslaw Bober for permitting us to use the MPEG-7 shape dataset.

⁴The CSS descriptor was computed using the implementation provided in the MPEG-7 XM [12].

References

- [1] S. Loncaric, “A survey of shape analysis techniques,” *Patt. Recog.*, 31(8):983–1001, 1998.
- [2] D. Zhang and G. Lu, “Review of shape representation and description techniques,” *Patt. Recog.*, 37(1):1–19, 2004.
- [3] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000.
- [4] F. Mokhtarian and M. Bober, *Curvature Scale Space Representation: Theory, Applications, and MPEG-7 Standardization*, Kluwer Academic Publishers, 2003.
- [5] D. G. Kendall, “Shape manifolds, procrustean metrics and complex projective spaces.,” *Bulletin of the London Mathematical Society*, 16:81–121, 1984.
- [6] M. Zuliani, C. Kenney, and B. S. Manjunath, “A definition of shape for jordan manifolds,” Tech. Rep. TR-VRL-03-2004, VRL, ECE Department, UCSB, 2004.
- [7] A. Goshtasby, “Description and discrimination of planar shapes using shape matrices,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 7(6):738–743, Nov 1985.
- [8] A. Taza and C. Y. Suen, “Discrimination of planar shapes using shape matrices,” *IEEE Trans. Syst. Man Cybern.*, 19(5):1281–1289, 1989.
- [9] S. K. Parui, E. Sarma, and D. Majumder, “How to discriminate shapes using the shape vector,” *Pattern Recognit. Lett.*, 4:201–204, 1986.
- [10] M. Bober, F. Preteux, and W. Y. Kim, *Introduction to MPEG7: Multimedia Content Description Interface*, chapter 15, pp. 234–238, John Wiley & Sons Ltd., 2002.
- [11] J. L. Lisani, *Shape Based Automatic Images Comparison*, Ph.D. thesis, Université Paris IX-Dauphine, 2001.
- [12] Institute for Integrated Systems http://www.lis.ei.tum.de/research/bv/topics/mmdb/e_mpeg7.html, Feb 2004.