

ROBUST CONTOUR DESCRIPTION AND MEASUREMENT USING DISTANCE ON THE GRASSMANN MANIFOLD

Zhaozhong Wang and Zesheng Tang

Image Processing Center, Beihang University
Beijing 100191, China

ABSTRACT

We propose robust affine-invariant contour descriptor and measure for shape matching under nonlinear deformations. The descriptor is formed by orthonormal configuration matrix of local contour. The geodesic distance on Grassmann manifold is used to measure similarities of shapes under locally affine transformations, which can approximate complex deformations like articulations. A rigorous perturbation analysis proves that condition numbers of configuration matrices are critical for robustness. Then a method to improve matching stability using the condition numbers is deduced. Commonly used contour matchers, e.g., dynamical programming and others, are all applicable to the descriptor to obtain satisfied matching. Experimental evaluations are given using both synthetic and real-world images.

Index Terms— Affine invariance, contour descriptor, shape measure, Grassmann manifold, perturbation analysis.

1. INTRODUCTION

Shape or contour matching receives increased attentions in recent years. Such a problem typically consists of three parts: shape description, similarity measure, and matching based upon the measure, where the measure is typically coupled with the descriptor. Our contribution is that we propose a new shape descriptor with a new similarity measure, which can be used for different matching methods. After introducing basic affine invariance, we deduce the orthonormal configuration matrices (OCM) of local contours as descriptors, and use the geodesic distance on Grassmann manifold to measure similarity of descriptors under locally affine transformations. The proposed shape descriptor with measure is theoretically solid, robust and fully affine invariant to approximate complex deformations, especially the articulated motion.

A number of shape description and matching methods have been proposed in the literature, we mainly review those with certain invariant properties: The widely used Fourier descriptors were generalized to affine-invariant descriptions [1]. Boundary moments [2] of contours are invariant to transla-

tion, rotation and scaling. Contour described by differential geometry [3] results in affine invariant curvature, which is sensitive to noise due to high-order derivatives. The shape invariant signature [4] is based on algebraic invariants.

The shape context (SC) [5] is well-known, which has been generalized to inner-distance shape context (IDSC) [6]. A limitation of them is that they lack mathematical interpretations to predict their performance. There are researches using angles as contour descriptions, which may be invariant to scale [7] or to translation and rotation [8]. The work of [9] uses a descriptor which is also a generalization of SC by considering both the length and angle description. Our previous work [10] adopts sorted diagonals (SD) of orthogonal projection matrices to form affine invariant descriptions.

There are work using Grassmann manifold [11, 12], most of them focus on *global* shape measures. Our work is different from them in two folds: First, we establish a *local* shape measure, which is available under complex shape transformations like perspective, nonrigid and articulation. Second, we propose a detailed perturbation analysis, then deduce an approach to improve matching robustness using condition numbers of configuration matrices.

2. CONTOUR DESCRIPTOR AND MEASURE

Considering in general two shape contours $\mathcal{C}_X, \mathcal{C}_Y \subset \mathbb{R}^2$ with m_x and m_y ($m_x \neq m_y$) landmark points, we aim to establish point correspondences between them. To determine if two points $\mathbf{x}_i \in \mathcal{C}_X$ and $\mathbf{y}_j \in \mathcal{C}_Y$ are matched, we take $2 \leq n \leq \min(m_x, m_y)$ contour points neighboring respectively to \mathbf{x}_i and \mathbf{y}_j to form *configuration matrices* $X_i, Y_j \in \mathbb{R}^{n \times 2}$ [10], where each row of X_i or Y_j is the point coordinate \mathbf{x}_i^T or \mathbf{y}_j^T . We assume that X_i and Y_j are of full rank, and they are related by a nonlinear transformation like perspective or articulation in the case that \mathbf{x}_i and \mathbf{y}_j are correspondence. Since both X_i and Y_j are local, we may approximate locally the nonlinear transformations using affine transformations. By introducing the *centerized configuration matrices* [10] \check{X}_i and \check{Y}_j to remove the effect of translation, we have

$$\check{Y}_j = \check{X}_i A, \quad (1)$$

where A is a 2×2 nonsingular matrix representing affine transformations like rotation, scaling and skewing. Note that

This work was supported by the National Natural Science Foundation of China under Grant 60803071.

this equation may have the reverse-order problem, i.e., the rows of \check{X}_i may not correspond directly to the rows of \check{Y}_j ; see Section 4 later for detailed treatments.

Eq. (1) indicates that the column spaces $\mathcal{R}(X_i)$ of X_i and $\mathcal{R}(Y_j)$ of Y_j are identical [13], i.e.,

$$\mathcal{R}(Y_j) = \mathcal{R}(X_i). \quad (2)$$

Thus we may use the subspaces as invariant description to local contours, and determine the similarity of \mathbf{x}_i and \mathbf{y}_j by measuring a distance between the subspaces:

$$\gamma(\mathbf{x}_i, \mathbf{y}_j) := \text{dist}[\mathcal{R}(X_i), \mathcal{R}(Y_j)], \quad \mathbf{x}_i \in \mathcal{C}_X, \mathbf{y}_j \in \mathcal{C}_Y. \quad (3)$$

but we need concrete and specific forms of the above description and distance for practical computations. Let $\Theta_i, \Omega_j \in \mathbb{R}^{n \times 2}$ be the orthonormal configuration matrices (OCM), whose columns form orthonormal bases for the subspaces $\mathcal{R}(X_i)$ and $\mathcal{R}(Y_j)$ respectively, we use Θ_i and Ω_j as local *contour descriptors* of points \mathbf{x}_i and \mathbf{y}_j . Next we require to choose a specific form of the subspace distance. The subspaces $\mathcal{R}(X_i)$ and $\mathcal{R}(Y_j)$ are both 2-dimensional spaces of the n -dimensional space \mathbb{R}^n where the column vectors of X_i and Y_j lie in. Thus $\mathcal{R}(X_i)$ and $\mathcal{R}(Y_j)$ are elements of the Grassmann manifold $\mathcal{G}(2, n)$ [14], on which the geodesic distance is given by

$$\text{dist}[\mathcal{R}(X_i), \mathcal{R}(Y_j)] = \sqrt{\theta_1^2 + \theta_2^2}, \quad (4)$$

where $\theta_1 \geq \theta_2$ are the two *principal angles* between $\mathcal{R}(X_i)$ and $\mathcal{R}(Y_j)$. Numerical methods to compute the principal angles are well studied [15]: By computing the following *thin* singular value decomposition (SVD)

$$\Theta_i^T \Omega_j = U \Sigma V^T, \quad \Sigma = \text{diag}(\sigma_1, \sigma_2),$$

where Θ_i and Ω_j are contour descriptors, i.e., the orthonormal matrices from \check{X}_i and \check{Y}_j ; $U \in \mathbb{R}^{n \times 2}$, $V \in \mathbb{R}^{2 \times 2}$, $U^T U = V^T V = V V^T = I$, σ_1 and σ_2 are the two singular values with $\sigma_1 > \sigma_2$, we have the relation

$$\cos \theta_i = \sigma_i, \quad i = 1, 2. \quad (5)$$

Though the SVD is required, the computation of principal angles is not very time-consuming since the space dimension is only two. Fig. 1 illustrates the use of the descriptor and measure to match articulated shapes.

3. PERTURBATION ANALYSIS

This section analyzes the stability of the descriptor and measure under perturbations of contour points; this will reveal key factors to affect robustness of contour matching, then deduce approaches to improve the robustness. We consider how the principal angles between $\mathcal{R}(X_i)$ and $\mathcal{R}(Y_j)$ change when the elements in X_i and Y_j are subject to perturbations, where

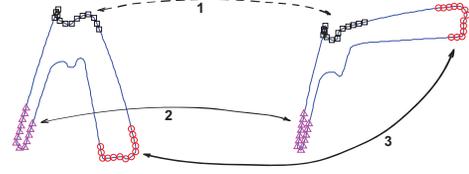


Fig. 1. Illustrating the proposed work for matching articulated shapes in [6]. The pair 1 of local configurations locate at joints, whose deformation cannot be estimated by affine transform (with the distance measure γ a high value of 0.9603). The rigid transforms of pairs 2 and 3 can be approximated by affine, whose γ are 0.1626 and 0.4233, respectively.

the perturbations of configuration matrices are caused by the perturbations of contour points. Assume that X_i and Y_j are perturbed to ΔX_i and ΔY_j respectively, and

$$\frac{\|\Delta X_i\|_2}{\|X_i\|_2} \leq \epsilon_X, \quad \frac{\|\Delta Y_j\|_2}{\|Y_j\|_2} \leq \epsilon_Y, \quad (6)$$

then the following theorem reveals the perturbation bound for principal angles.

Theorem 1 (Bjorck and Golub [15]). *Let the perturbations of X_i and Y_j be given in (6), then the perturbations of principal angles are bounded by*

$$|\Delta \theta_k| \leq g_{\max} [\kappa(X_i) \epsilon_X + \kappa(Y_j) \epsilon_Y] + O(\delta^2), \quad (7)$$

where $k = 1, 2$, and

$$g_{\max} = \sqrt{\sin^2 \theta_{\max} + \cos^2 \theta_{\min}} \leq \sqrt{2}, \quad (8)$$

$\kappa(X_i)$ and $\kappa(Y_j)$ are respectively the condition numbers of the matrices X_i and Y_j , θ_{\max} and θ_{\min} denote respectively the maximal and minimal principal angles between the column spaces of X_i and Y_j , and $\delta = \kappa(X_i) \epsilon_X + \kappa(Y_j) \epsilon_Y$.

A conclusion deduced from this theorem is that if both $\kappa(X_i)$ and $\kappa(Y_j)$ are small, then the principal angles θ_k are well determined. Next we derive a perturbation bound for the measure γ based on the principal angles.

Corollary 1. *For contour matching between 2-dimensional images, the perturbation $\Delta \gamma$ of the measure γ defined in Eq. (3) is bounded by*

$$|\Delta \gamma| \leq 2\sqrt{2} [\kappa(X_i) \epsilon_X + \kappa(Y_j) \epsilon_Y] + O(\delta^2), \quad (9)$$

where $\kappa(X_i)$, $\kappa(Y_j)$ and δ are defined in the Theorem 1.

We omit the simple proof here due to page limitation. This corollary indicates that the measure is also well determined under the condition that both the matrices X_i and Y_j are with small condition number. A contour configuration of straight line has a very large condition number (in theory it is infinity), but a local contour configuration has a much smaller condition number if it is far from a straight line; such a configuration would be better for more robust matching. This motivates us to use the condition number to select better contour points, as given in the next section.

4. MATCHING ISSUES

Here we discuss some issues in contour matching arose from the proposed descriptor and measure. First, following Section 3 we deduce an approach to improve robustness using condition numbers. We use condition numbers as criteria to select subsets of contour points

$$C'_X = \{\mathbf{x}_i : \kappa(X_i) < \lambda\}, \quad C'_Y = \{\mathbf{y}_i : \kappa(Y_j) < \lambda\}, \quad (10)$$

for matching, where λ is a given threshold to condition numbers (we set $\lambda \leq 5.0$). Eq. (10) states that points with smaller condition numbers are left as stable points. Such a selection is theoretically reasonable, and will improve algorithm's robustness to perturbations in terms of the previous analysis.

After the above step we can perform matching of contour points. The proposed descriptor with measure is suitable to many contour matchers. We mainly use dynamical programming (DP) and the Hungarian matching [5] in experiments, which will show that the proposed obtains better performance than many others under different matchers.

There is a point reverse-order problem due to sampling of contours. Eq. (1) implicitly assumes that the k -th point (row) in \check{X}_i corresponds to the k -th point (row) in \check{Y}_j ; this is false if corresponding objects in two images are shown as reflection (see the horse images in Fig. 4) but with same sampling direction (e.g., both clock-wise). In this case the contours are corresponded in reversed order, and Eq. (1) should be $\check{Y}_j = J\check{X}_iA$ with J the reverse-identity matrix. To taking into account this effect, we apply a simple approach by matching contours twice in the same and reversed order respectively, then choose the one with larger number of correspondences as the better.

5. EXPERIMENTAL RESULTS

In this section we test accuracy and robustness of the proposed descriptor with measure using synthetic and real-world data. For fairly comparisons, we use different combinations of descriptors and matchers. The descriptors we compared include the proposed OCM (formed by $n = 25$ points), the SC and the IDSC; the matchers are the DP, the Hungarian and a simple nearest-neighbor (NN) matching method.

Fig. 2 shows part of the synthetic data used in experiments, where contour matchings are between the first column and one of the rest columns. The first experiment is mainly for global affine transformations like anisotropic scaling. The left of Fig. 3 shows the matching results against various scaling ratios, depicting that the proposed descriptor results in the highest matching score under different matchers. Next we compare the algorithms' robustness to noise. The middle of Fig. 3 depicts the matching accuracy against noise variances, showing the good robustness of the proposed OCM with measure. We also test the computational loads of different algorithms. The result in the right of Fig. 3 indicates that for the

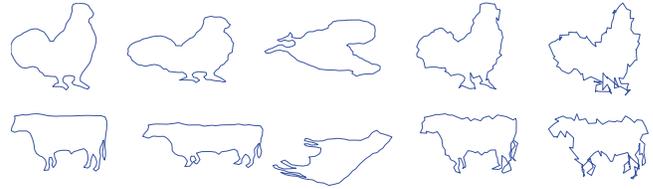


Fig. 2. Part of synthetic data used for testing. The contours in first column is the original shapes obtained from the MPEG-7 database. Other columns are transformed from the first by affine or noise.

Descriptor + Matcher	Top 1	Top 2	Top 3	Top 4
SC + DP	20/40	10/40	11/40	5/40
IDSC + DP [6]	40/40	34/40	35/40	27/40
SD + DP [10]	39/40	39/40	37/40	21/40
OCM + DP	40/40	39/40	39/40	30/40
SC + Hungarian [5]	18/40	7/40	8/40	5/40
IDSC + Hungarian	40/40	33/40	30/40	27/40
SD + Hungarian	39/40	37/40	33/40	20/40
OCM + Hungarian	40/40	40/40	37/40	26/40

Table 1. Retrieval results for the articulation database in [6].

matchers of the DP and the NN, the algorithms using the SC and IDSC are about two to three times faster than those using the proposed OCM. But it is interesting that the algorithm using the OCM followed by the Hungarian matcher is faster than those using the SC and the IDSC.

As proposed, our descriptor with measure is suitable for articulated shape matching. We use the articulation shape data in [6] to test the descriptors of SC, IDSC, SD [10] and OCM combined with the matchers of DP and Hungarian. The matching results are summarized as the numbers of the 1st, 2nd, 3rd and 4th most similar matches coming from the correct object, as shown in Table 1. The proposed OCM obtains better matching accuracy.

Finally we use real-world images for testing. Contours in the images are extracted by the GrabCut algorithm [16], with perturbations in contour points. We use the proposed OCM with the DP matcher for experiments. Fig. 4 shows the matching performance for articulated and deformable shapes, where points without correspondences are mainly those with higher condition numbers and are removed by the scheme in Section 4. The DP matcher also has some capability to reject matching outliers [6], but its effect is minor.

6. CONCLUSION

The proposed local contour descriptor with distance measure from Grassmann manifold is theoretically invariant to affine transformations and can well approximate articulated and other nonlinear shape deformations. Condition numbers of contour configuration matrices are key for robustness; removing those with high condition numbers can improve

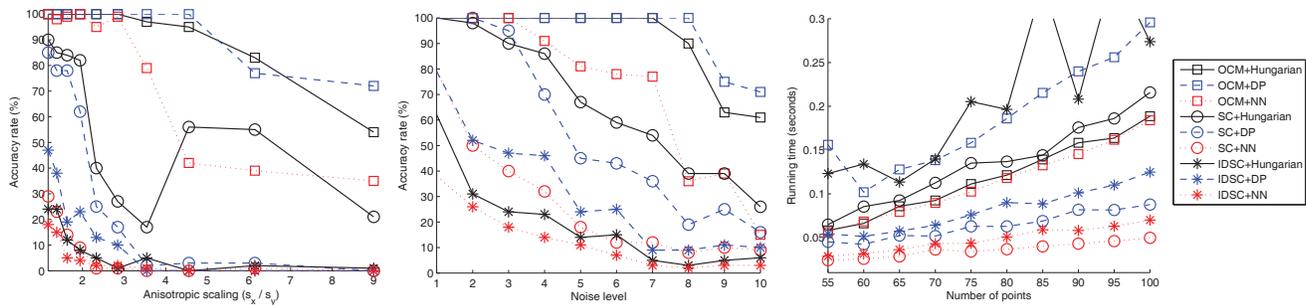


Fig. 3. Experimental results for comparisons of descriptors. Left: Comparison under different degrees of anisotropic scaling. Middle: Comparison under different noise levels. Right: Comparison of running times.

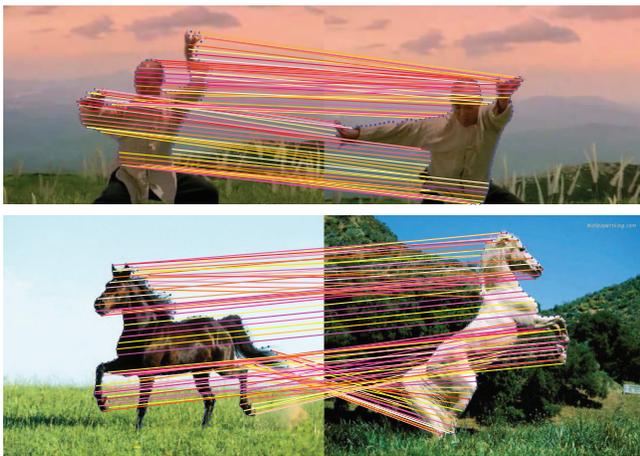


Fig. 4. Contour matching for real-world articulated images using the proposed method. Top: Jet Li's movie frames. Bottom: Horse images searched from the Web.

robustness for contour matching under the proposed descriptor and measure. A drawback is that the computational load is relative large; using a faster estimation of the distance measure will be studied further.

7. REFERENCES

- [1] K. Arbter, W. E. Snyder, H. Burkhardt, and G. Hirzinger, "Application of affine-invariant Fourier descriptors to recognition of 3-D objects," *IEEE Trans. PAMI*, vol. 12, no. 7, pp. 640–647, 1990.
- [2] J. Heikkila, "Pattern matching with affine moment descriptors," *Pattern Recognition*, vol. 37, no. 9, pp. 1825–1834, 2004.
- [3] F. Mokhtarian and S. Abbasi, "Affine curvature scale space with affine length parametrisation," *Pattern Anal. Appl.*, vol. 4, no. 1, pp. 1–8, 2001.
- [4] D.M. Squire and T.M. Caelli, "Invariance signature: Characterizing contours by their departures from invariance," *Computer Vision and Image Understanding*, vol. 77, pp. 284–316, 2000.
- [5] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. PAMI*, vol. 24, no. 24, pp. 509–522, 2002.
- [6] H. Ling and D. W. Jacobs, "Shape classification using the inner-distance," *IEEE Trans. PAMI*, vol. 29, no. 2, pp. 286–299, 2007.
- [7] C.E. Lu, L.J. Latecki, N. Adluru, X.W. Yang, and H.B. Ling, "Shape guided contour grouping with particle filters," in *ICCV*, 2009, pp. 2288–2295.
- [8] H. Riemenschneider, M. Donoser, and H. Bischof, "Using partial edge contour matches for efficient object category localization," in *ECCV*, 2010, pp. V: 29–42.
- [9] T. Ma and L. J. Latecki, "From partial shape matching through local deformation to robust global shape similarity for object detection," in *CVPR*, 2011, pp. 1441–1448.
- [10] Z. Wang and M. Liang, "Locally affine invariant descriptors for shape matching and retrieval," *IEEE Signal Processing Letters*, vol. 17, no. 9, pp. 803–806, 2010.
- [11] E. Begelfor and M. Werman, "Affine invariance revisited," in *CVPR*, 2006, pp. II: 2087–2094.
- [12] P.K. Turaga, A. Veera., and R. Chellappa, "Statistical analysis on stiefel and grassmann manifolds with applications in computer vision," in *CVPR*, 2008, pp. 1–8.
- [13] G.H. Golub and C.F. van Loan, *Matrix Computations*, Johns Hopkins University Press, 3rd edition, 1996.
- [14] P.-A. Absil, R. Mahony, and R. Sepulchre, "Riemannian geometry of Grassmann manifolds with a view on algorithmic computation," *Acta Appl. Math.*, vol. 80, no. 2, pp. 199–220, 2004.
- [15] A. Bjorck and G.H. Golub, "Numerical methods for computing angles between linear subspaces," *Mathematics of Computation*, vol. 27, no. 123, pp. 579–594, 1973.
- [16] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake, "'GrabCut': Interactive foreground extraction using iterated graph cuts," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 309–314, 2004.