Rotation Invariant Shape Contexts based on Feature-space Fourier Transformation

Su Yang[1], Yuanyuan Wang[2]
Dept. of Computer Science and Engineering, Fudan University, Shanghai 200433, China
Dept. of Electronic Engineering, Fudan University, Shanghai 200433, China
E-mail: {*suyang,yywang*}*@fudan.edu.cn*

Abstract

We propose a new pixel-level shape descriptor. First, shape contexts are computed. Then, 2D FFT is performed on each 2D histogram from shape contexts. Such a scheme solves the rotation-invariance problem of shape contexts based on the shift theorem of Fourier Transformation while does not increase the computational complexity. Theoretical proof and experimental validation are provided.

1. Introduction

Shape matching plays an important role in a variety of applications in computer vision and pattern recognition. The key problem for shape analysis is how to capture and describe the characteristics of a shape. Shape context is a recently proposed descriptor [1], which has received much attention to date. An ideal shape descriptor should possess the following properties: (1) good discriminative power, (2) rotation and scale-invariant, (3) robust against deformations and occlusions. The advantages of shape contexts are as follows. (1) The computation of shape contexts is performed directly on pixels, avoiding the preprocessing to detect structures in pixels, which is usually regarded an error-prone process. This promises the robustness of shape contexts. The discriminative power of shape contexts are very good because it figures out how the other point configure in reference to every point. The weakness of shape contexts lies in the rotation invariance. The rotation invariance of shape contexts is dependent on the tangent at every boundary point. To compute the tangent at every pixel, the boundary points must be computed and organized in order. This is against the fundamental spirit of shape contexts. One main advantage of shape contexts is that it can be directly applied to pixels without any error-prone preprocessing such as the perceptual organization of the boundary points into a point sequence. The tangent based rotation invariance is unstable in that perturbations may arise from both the outside outliers and the error-prone perceptual organization of the boundary points. Moreover, the perceptual organization of boundary points prevents the shape contexts from being applied to more general cases other than boundary based shape representations. It is known that shape can be represented in many ways, not just the boundary of objects. For example, skeleton points are the more generally applied primitives in binary image classification [2]. Besides, point set matching, which is a classical problem in computer vision and pattern recognition, tackles more general cases, not just boundary points. In view of such limit of shape contexts, we propose a new descriptor in [2], namely, statistical integration of pixel-level constraint histograms. It preserves the main advantages of shape contexts, discriminative power and robustness, while solves the rotation-invariance problem. However, this descriptor is limited in that its computational complexity is high, roughly $O(N^3)$. To improve the rotation-invariance of shape contexts while not introduce higher computational complexity, we propose in study a new scheme for pixel-level shape description. We refer to this new descriptor as Rotation-invariant Shape Contexts based on FFT (FFT-RISC). The key is to perform 2-dimensional FFT on the original Shape Contexts. Then, let the modulus of the FFT transformation of Shape Contexts be the signature to characterize how the other points distribute around every point. Based on the shift theorem of Fourier Transformation, in this paper, we have theoretically proved that the proposed FFT-RISC feature is invariant under any affine transformation. Some examples are also provided to experimentally validate the invariance of the FFT-RISC feature. A related work can be found in [3]. The difference between this study and [3] lies in two aspects: (1) The descriptor proposed in [3] relies on the computation of a center for a point set, which degrades its

0-7695-2929-1/07 $25.00 © 2007 IEEE
DOI 10.1109/ICIG.2007.18

discriminative power. (2) This study proposes to use FFT to solve the shift problem encountered in [3].

2. Rotation-invariant shape contexts implemented by FFT

In the following, we first present how to compute the rotation-invariant shape contexts via FFT (FFT-RISC). Then, we describe the method to match two point sets using the FFT-RISC feature. Finally, we prove the invariance of the FFT-RISC feature under affine transformations.

2.1. Computation of the FFT-RISC feature

Definition 1: int(r) functions to preserve only the integer portion of a real number r.

Definition 2: $|c|$ means the modulus of a complex value c.

Notes: The comments are enclosed between "/*" and "*/".

Subroutine 1: Compute the feature matrix for every point in a given point set P={P$_i$|i=1,2,…,L}.

Input: $P=\{P_i=(x_i,y_i)|i=1,2,\ldots,L\}$.
/*A point set of interest.*/

Output: $\{F(P_i)|i=1,2,\ldots,L\}$.
/*The feature matrix in association with every point P_i, $i=1,2,\ldots,L$.*/

Step 1: For $i,j=1,2,\ldots L$ and $j\neq i$, compute

$$l_{ij}=\sqrt{(x_j-x_i)^2+(y_j-y_i)^2}\,. \quad (1)$$

Step 2: Compute

$$r_0=\min\{\log(l_{ij})\,|\,i,j=1,2,...,L;i\neq j\}\,. \quad (2)$$

Step 3: For $i,j=1,2,\ldots L$ and $j\neq i$, compute

$$r_{ij}=\log l_{ij}-r_0\,. \quad (3)$$

Step 4: For $i,j=1,2,\ldots L$ and $j\neq i$, compute

$$\alpha_{ij}=\arctan\frac{y_j-y_i}{y_j+y_i}\,. \quad (4)$$

Step 5: Call subroutine 2 to compute the feature matrix $F(P_i)$ of every point P_i, $i=1,2,\ldots L$.

Step 6: Return $\{F(P_i)|i=1,2,\ldots,L\}$.

Note that: In subroutine 2, M will affect the resolution of FFT.

Subroutine 2: Compute the feature matrix for a given point P$_i$ in a point set P={P$_i$|i=1,2,…,L}.

Input: $\{(r_{ij},\alpha_{ij})|j=1,2,\ldots,L;j\neq i\}$.
Parameters: M, Δ_r and Δ_α.
/*M must be greater than int(r_{mi}/Δ_r)+1, where $r_{mi}=\max\{r_{ij}|\,j=1,2,\ldots,L;j\neq i\}$.*/

Output: $F(P_i)$.
/*The feature matrix of P_i.*/

Step 1: Let $h_{pq}=0$ for $p=1,2,\ldots,M$; $q=1,2,\ldots,N$.
/*Construct a 2D histogram $[h_{pq}]$.*/

Step 2: For $j=1,2,\ldots,L$ and $j\neq i$, compute $p=\text{int}(r_{ij}/\Delta_r)$ and $q=\text{int}(\alpha_{ij}/\Delta_\alpha)$, and then let $h_{pq}\leftarrow h_{pq}+1$.
/*Δ_r and Δ_α are two parameters determining the size of every block of the 2D histogram. Correspondingly, p_j and q_j are the indices to the 2D histogram.*/

Step 3: For $p=1,2,\ldots,M$ and $q=1,2,\ldots,N$, let $h_{pq}\leftarrow h_{pq}/(L\text{-}1)$.
/*Normalize the 2D histogram.*/

Step 4: Perform 2D FFT on the 2D histogram $[h_{pq}]$, which results in a new matrix $[f_{pq}]$ with the same dimension $M\times N$. Note that f_{pq} is a complex number, $p=1,2,\ldots,M$; $q=1,2,\ldots,N$.

Step 5: Let $F(P_i)=[a_{pq}]$, where a_{pq} is the modulus of f_{pq}, $p=1,2,\ldots,M$; $q=1,2,\ldots,N$.

Step 6: Return $F(P_i)$.

For a given point P_i, the step 1 to step 3 in subroutine 2 results in a 2D histogram as follows.

$$h_{pq}^{i}=\sum_{\substack{j=1\\ j\neq i}}^{L}\delta(r_{ij},\alpha_{ij},\Delta_r,\Delta_\alpha,p,q)\,, \quad (5)$$

$p=1,2,\ldots,M$ and $q=1,2,\ldots,N$, where

$$\delta(r_{ij},\alpha_{ij},\Delta_r,\Delta_\alpha,p,q)=\begin{cases}1 & \begin{matrix}r_{ij}\in[(p-1)\Delta_r,p\Delta_r]\wedge\\ \alpha_{ij}\in[(q-1)\Delta_\alpha,q\Delta_\alpha]\end{matrix}\\ 0 & else\end{cases}. \quad (6)$$

Following subroutine 1 and 2, a shape descriptor $F(P_i)$ can be obtained with regard to each point in a point set. Since the computation of FFT is very fast, it can be regarded that the above shape descriptor implemented by FFT does not lead to obviously higher computing cost in contrast to the original Shape Contexts. Because $F(P_i)$ figures our how the other points in the same point set distribute around P_i, the point correspondences between two point sets can then be computed based on such a pixel-level shape descriptor.

2.2. Matching of point sets

Subroutine 3: Compute the point correspondence between two given point sets P={P$_i$|i=1,2,…,L} and Q={Q$_i$|i=1,2,…,L′}, where we assume L′≥L without losing the generality.

Input: $\{F(P_i)|i=1,2,\ldots,L\}$ and $\{F(Q_j)|j=1,2,\ldots,L'\}$, where $F(P_i)=[a^i_{pq}]$ and $F(Q_j)=[b^j_{pq}]$, $p=1,2,\ldots,M$; $q=1,2,\ldots,N$.

/*Input the feature matrices of the two given point sets.*/

Output: ψ.

/*A set consists of the indices of matched pairs

Step 1: Initialize ψ as an empty set.

Step 2: Compute

$$D=\{d_{ij}=\sqrt{\sum_{p=1}^{M}\sum_{q=1}^{N}(a^i_{pq}-b^j_{pq})^2}\ |\ i=1,2,\ldots,L;\ j=1,2,\ldots,L'\}. \quad (7)$$

/*The distance between the two feature matrices $F(P_i)$ and $F(Q_j)$.

Step 3: Search $(s,t)=\arg\min_{i,j}\{d_{ij}\}$; Add a new element (s,t) to ψ; Update D by removing the elements $\{d_{ik}|k=1,2,\ldots,L\}$ and $\{d_{kj}|k=1,2,\ldots,L'\}$.

Step 4: If $|\psi|<L$, go to step 2. Else, return ψ.

In subroutine 3, the distance between two points is defined as the Euclidean distance between the corresponding two feature matrices. In each iteration in subroutine 3, the two points with the closest distance are selected to match. Of course, there must exist $|L-L'|$ unmatched points in Q.

2.3. Proof of invariance

In the following, we will show that $F(P_i)$ is a rotation-invariant pixel-level shape descriptor. If we apply an affine transformation to every pixel $P_i=(x_i,y_i)$ in point set P, then, we can obtain a new point set $Q=\{Q_i=(X_i,Y_i)|i=1,2,\ldots,L\}$, where (X_i,Y_i) is transformed from (x_i,y_i) via

$$\begin{bmatrix}X_i\\Y_i\end{bmatrix}=s\begin{bmatrix}\cos\theta & \sin\theta\\-\sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}x_i\\y_i\end{bmatrix}+\begin{bmatrix}x_0\\y_0\end{bmatrix}=\begin{bmatrix}s\cos\theta x_i+s\sin\theta y_i+x_0\\-s\sin\theta x_i+s\cos\theta y_i+y_0\end{bmatrix}. \quad (8)$$

In the above transformation, (x_0,y_0), s, and β are the parameters to determine the translation, scale, and rotation, respectively. According the above equation, it is easy to see that

$$L_{ij}=\sqrt{(X_i-X_j)^2+(Y_i-Y_j)^2}=s\sqrt{(x_i-x_j)^2+(y_i-y_j)^2}=l_{ij}. \quad (9)$$

It follows that

$$R_0=\min_{i,j}\{\log(L_{ij})\}=\min_{i,j}\{\log(l_{ij})+\log(s)\}=\min_{i,j}\{\log(l_{ij})\}+\log(s)=r_0+\log(s). \quad (10)$$

Due to Eq. (9) and Eq. (10),

$$R_{ij}=\log(L_{ij})-R_0=[\log(l_{ij})+\log(s)]-[r_0+\log(s)]=\log(l_{ij})-r_0=r_{ij}. \quad (11)$$

Eq. (11) shows that scaling does not affect the computation of the feature matrix.

According to Eq. (4),

$$\begin{aligned}\tan\beta_{ij}&=\frac{Y_j-Y_i}{X_j-X_i}\\&=\frac{-\sin\theta(x_j-x_i)+\cos\theta(y_j-y_i)}{\cos\theta(x_j-x_i)+\sin\theta(y_j-y_i)}\\&=\frac{-\sin\theta+\cos\theta\tan\alpha_{ij}}{\cos\theta+\sin\theta\tan\alpha_{ij}}\\&=\frac{-\sin\theta\cos\alpha_{ij}+\cos\theta\sin\alpha_{ij}}{\cos\theta\cos\alpha_{ij}+\sin\theta\sin\alpha_{ij}}\\&=\frac{\sin(\alpha_{ij}-\theta)}{\cos(\alpha_{ij}-\theta)}=\tan(\alpha_{ij}-\theta)\\&\Rightarrow\beta_{ij}=\alpha_{ij}-\theta\pm2\pi\end{aligned} \quad (12)$$

It is know that $\alpha_{ij}\in[0,2\pi]$, $\beta_{ij}\in[0,2\pi]$, and $\theta\in(0,2\pi)$. So, $\alpha_{ij}-\theta\in(-2\pi,2\pi)$. Eq. (12) shows that the difference between β_{ij} and α_{ij} is a constant if $\alpha_{ij}-\theta\geq0$. If $\alpha_{ij}-\theta<0$, then, $\beta_{ij}=\alpha_{ij}-\theta+2\pi$. Because of Eq. (11) and Eq. (12), we know that

$$(R_{ij},\beta_{ij})=(r_{ij},\alpha_{ij}-\theta+C), \quad (13)$$

where $C=2\pi$ or $C=0$ for $j=1,2,\ldots,L$ and $j\neq i$. Due to Eq. (11)~(13), we have

$$(R_{ij}/\Delta_r,\beta_{ij}/\Delta_\alpha)=(r_{ij}/\Delta_r,(\alpha_{ij}-\theta+C)/\Delta_\alpha), \quad (14)$$

Denote the Fourier transformation (FT) of $(r_{ij}/\Delta_r,\alpha_{ij}/\Delta_\alpha)$ as

$$f(r_{ij}/\Delta_r,\alpha_{ij}/\Delta_\alpha)\Leftrightarrow F(u,v). \quad (15)$$

In accordance with the shift theorem of Fourier transformation, we hold

$$\begin{aligned}&f(R_{ij}/\Delta_r,\beta_{ij}/\Delta_\alpha)=f(r_{ij}/\Delta_r,(\alpha_{ij}-\theta+C)/\Delta_\alpha)\\&\Leftrightarrow\exp\{-[2j\pi(\theta-C)/\Delta_\alpha]v\}F(u,v).\end{aligned} \quad (16)$$

This leads to

$$|f(R_{ij},\beta_{ij})|=|F(u,v)|=|f(r_{ij},\alpha_{ij})| \quad (17)$$

where $|*|$ means the modulus of a complex value. On account of subroutine 2 and Eq. (17), we can conclude that the proposed feature is invariant under affine transformation. Because we apply $(\mathrm{int}(r_{ij}/\Delta_r)$ and $\mathrm{int}(\alpha_{ij}/\Delta_\alpha))$ in constructing the 2D histogram and we use 2D FFT instead of 2D FT in computing the feature matrix, the affine invariance property of the proposed feature only holds approximately.

3. Experiments

We demonstrate the invariance of the FFT-RISC shape descriptor via the following experiments. We conducted 3 tests in total. The image to be tested is shown in Fig. 1, which is a Chinese character consisting of a point set. In the first test, we generate a new point set by applying an affine transformation to the point set shown in Fig. 1. The new point set and the original point set are illustrated in Fig. 2. In the second test, we firstly add 50 noisy points to the point set shown in Fig. 1. The point set containing noisy points is shown in Fig. 3. Then, we apply the same affine transformation to the point set shown in Fig. 3. Fig. 4 illustrates both the point set shown in Fig. 3 and the new point set transformed from this point set. Fig. 5 and Fig. 6 are similar to Fig. 3 and Fig. 4 while the number of noisy points is increased to be 100. In each of the above 3 tests, we compute the FFT-RISC feature with two groups of parameters. The parameters of the first group are: M=64, Δ_r=0.1, and Δ_α=10°. The parameters of the second group are: M=128, Δ_r=0.1, and Δ_α=10°. We only alert M in the two groups since we want to observe how the FFT computation affects the feature. After we obtained the feature matrix for each point following the procedure described in subroutine 2 and 3, we matched the two point sets in each test via the method described in subroutine 3. The experimental results are shown in Table 1, which shows the ratio of the number of incorrectly matched pairs to the total pairs in every test. It can be seen that invariance can be guaranteed by the proposed shape descriptor if there are no noisy points. When the number of noisy points increase, the matching becomes worse. Also, bigger M results in better matching of points because the resolution of FFT increases with M. These tests confirm the invariance of the FFT-RISC descriptor initially. Yet, there are a lot of works to be done to refine the proposed scheme. Nevertheless, it is an interesting step in improving the invariance of Shape Contexts. The advantage of the propose scheme is: It promises rotation-invariance with little additional computational cost because the computation of FFT is very fast. In the experiments, the computation of the FFT-RISC feature is very fast. So, the proposed scheme merits further investigations for refinement.

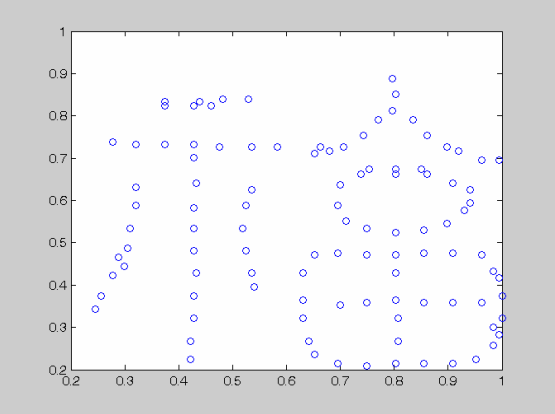

Fig. 1: Point set A

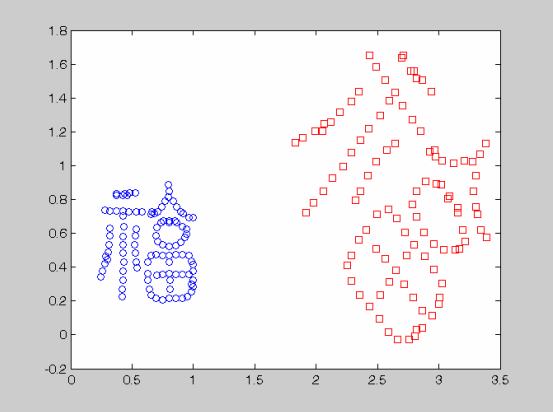

Fig. 2: Point set A and its affine transformation

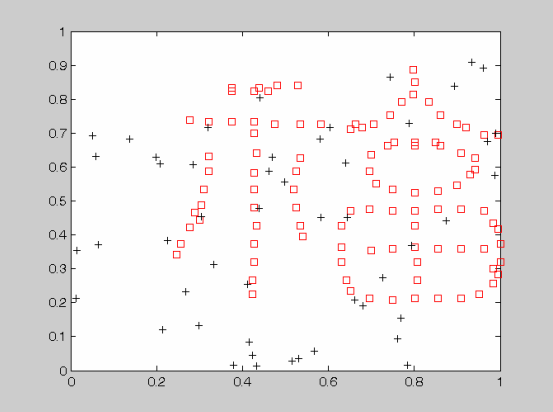

Fig. 3: Point set B (Point set A plus 50 additional noisy points)

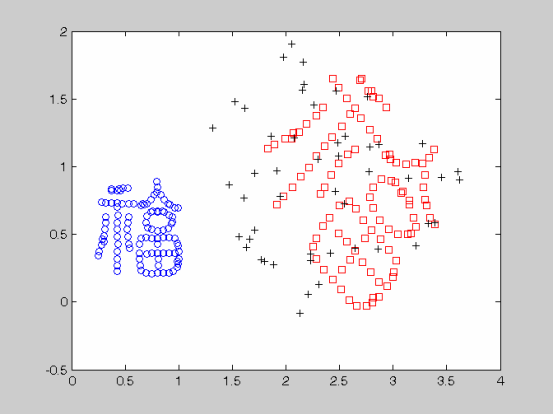

Fig. 4: Point set B and its affine transformation

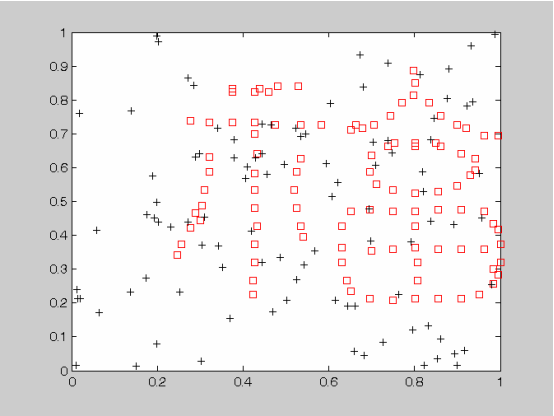

Fig. 5: Point set C (Point set A plus 100 additional noisy points)

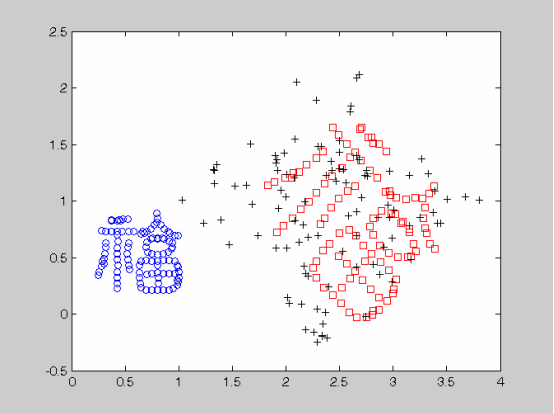

Fig. 6: Point set C and its affine transformation

Table 1. Ratio of erroneous matching

Noisy points	0	50	100
M=64	0	7/105	31/105
M=128	0	1/105	15/105

4. Summary

This paper proposes a FFT based scheme to improve the rotation-invariance of shape contexts. The main contribution is: The rotation invariance is achieved at very little cost because of the fast computation of FFT. The theoretical proof of the invariance of the proposed descriptor is provided. The experiments initially evaluated the proposed FFT-RISC shape descriptor. Yet, there are a lot of works to be done to refine the proposed scheme, which will be our future endeavor.

Acknowledgements: This work is supported by Natural Science Foundation of China under grant 60305002, China/Ireland Science and Technology Research Collaboration Fund under grant CI-2004-09, and National Basic Research Program of China under grant 2006CB705700.

References

[1] S. Belogie, J. Malik, J. Puzicha, "Shape Matching and Object Recognition using Shape Contexts". IEEE Trans. PAMI, 2002, vol. 24, pp. 509-520

[2] S. Yang: "Symbol Recognition via Statistical Integration of Pixel-Level Constraint Histograms: A New Descriptor", IEEE Trans. PAMI, 2005, Vol. 27, No. 2, pp. 278-281

[3] W.-P. Choi, K.-M. Lam, and W.-C. Siu: "A Robust Line-feature-based Hausdorff Distance for Shape Matching", PCM 2001, LNCS 2195, 2001, pp. 764-771