Modified Fourier Descriptors for Shape Representation – A Practical Approach

Yong Rui, Alfred C. She, and Thomas S. Huang *

Beckman Institute, University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA
E-mail: yrui@ifp.uiuc.edu, a-she@uiuc.edu, huang@ifp.uiuc.edu

Abstract. We propose a Modified Fourier Descriptor and a new distance metric for describing and comparing closed planar curves. Our method accounts for the effects of *spatial discretization* of shapes, an issue seldom mentioned, much less addressed in the literature.

The motivating application is shape matching in a content based image retrieval system. The application requires a compact and reliable shape representation, and a feature distance measure which can be computed in real time. Experimental results suggest that our method is a feasible solution for on-line shape comparisons in such a system.

1 Introduction

Content based retrieval (CBR) has gained considerable attention recently[1-5]. Color and texture features are explored in [1-5]. This paper will focus on shape matching. We propose that a useful shape representation should satisfy the following four conditions:

1. Robustness to Transformation – the representation must be invariant to translation, rotation, and scaling of shapes, as well as the starting point used in defining the boundary sequence.
2. Robustness to Noise – the representation must be robust to *spatial discretization* noise.
3. Feature extraction Efficiency – feature vectors should be computed efficiently.
4. Feature matching Efficiency – since matching is done on-line, the distance metric must require a *very* small computational cost.

We propose the Modified Fourier Descriptor (MFD), which satisfies the four conditions above. The Fourier Descriptor (FD) method is the most closely related work, so we give a brief review of it in section 2. We discuss the proposed MFD in section 3. Comparisons between MFD and existing methods are given in section 4. Experimental results and conclusions are in sections 4 and 5, respectively.

List of symbols:

* This work was supported by the NSF/DARPA/NASA Digital Library Initiative under Cooperative Agreement No. 94-11318.

- N_V: number of vertices of a polygon;
- N_B: number of boundary points of a shape;
- N_C: number of FD coefficients used in shape reconstruction;
- V_i: the ith vertex of a polygon;
- N_{dense}: number of "dense" samples used in resampling in the MFD;
- N_{unif}: number of uniformly spaced samples used in MFD method;
- α, β, γ: planar curves (shape boundaries).

2 Fourier Descriptors

There are two commonly known FD's, described in [7] and [6], which we denote as "FD1" and "FD2", respectively. FD1 has low efficiency in reconstructing the shape, so we discuss FD2 only. A point moving along the boundary γ generates the complex function $u(l) = x(l) + jy(l)$. FD2 is defined as:

$$a_n = \frac{1}{L\left(\frac{2\pi n}{L}\right)} \sum_{k=1}^{N_V} (b_{k-1} - b_k) e^{-j\left(\frac{2\pi n l_k}{L}\right)} \quad (1)$$

where L is the total length of γ; $l_k = \sum_{i=1}^{k} |V_i - V_{i-1}|$ (for $k > 0$ and $l_0 = 0$); and $b_k = \frac{V_{k+1} - V_k}{|V_{k+1} - V_k|}$.

Let $\{a_n\}$ and $\{b_n\}$ denote the FD's of two curves α and β, respectively, and assume only N_C harmonics are used; the distance metric is

$$d(\alpha, \beta) = \sqrt{\sum_{n=-N_C}^{N_C} |a_n - b_n|^2} \quad (2)$$

To account for the effects of scale (s), rotation (ϕ), and starting point (p), we must minimize the distance metric

$$d^*(\alpha, \beta) = \min_{s,\phi,p} \sum_{n=-M, n\neq 0}^{M} \left| a_n - s e^{j(np+\phi)} b_n \right|^2 \quad (3)$$

over the parameters (s, ϕ, p). This is a computationally expensive optimization problem and makes FD2 impractical for shape matching in a real-time CBR system, especially when the image database is large.

3 Proposed Method – Modified Fourier Descriptors

The reason why the feature-matching computation for FD1 and FD2 is expensive is that the length between adjacent vertices l_k is not uniform. If the starting point changes, the whole boundary sequence $\{b_k\}$ will change.

We overcome this drawback by defining the boundary as a 4-connected boundary, which has uniform length. Let $z(n) = x(n) + jy(n)$, $n = 0, ..., N_B - 1$, be the boundary sequence. The MFD is defined as the Discrete Fourier Transform of $z(n)$.

$$Z(k) = \sum_{n=0}^{N_B-1} z(n) e^{-j\frac{2\pi n k}{N_B}} = M(k) e^{j\theta(k)} \quad (4)$$

where $k = 0, ..., N_B - 1$.

Next, we examine the properties of MFD and propose a distance metric which is both reliable and easy to compute. Let $z'(n)$ be a boundary sequence obtained from $z(n)$: $z'(n)$ is $z(n)$ translated by z_t, rotated by ϕ, and scaled by α, with the starting point shifted by l. Explicitly, $z'(n)$ is related to $z(n)$ by

$$z'(n) = \alpha z(n - l) e^{j\phi} \tag{5}$$

The corresponding MFD of $z'(n)$ is

$$Z'(k) = \sum_{n=0}^{N_B-1} z'(n) e^{-j\frac{2\pi nk}{N_B}} = \alpha e^{j\phi} \sum_{n=0}^{N_B-1} z(n-l) e^{-j\frac{2\pi nk}{N_B}} \tag{6}$$

Setting $m = n - l$, we get

$$Z'(k) = \alpha e^{j\phi} \sum_{m=-l}^{N_B-l-1} z(m) e^{-j\frac{2\pi mk}{N_B}} e^{-j\frac{2\pi lk}{N_B}} = \alpha e^{-j(\phi + \frac{2\pi lk}{N})} Z(k) = M'(k) e^{j\theta'(k)} \tag{7}$$

where

$$M'(k) = \alpha M(k),\ \theta'(k) = \phi + \theta(k) - \frac{2\pi lk}{N_B} \tag{8}$$

The distance metrics for magnitude (D_m) and phase (D_p) are defined as

$$D_m = Var[ratio],\ D_p = Var[shift] \tag{9}$$

where

$$ratio(k) = \frac{M(k)}{M'(k)};\ shift(k) = \frac{\theta(k) - \theta'(k) - \phi}{k}; \tag{10}$$

$$\phi = \theta_0 - \theta'_0;\ k = -N_C, ..., N_C, k \neq 0. \tag{11}$$

θ_0 and θ'_0 are the orientations of the major axes of the two shapes, defined as

$$\theta = \frac{1}{2} \tan^{-1} \left(\frac{2cm_{11}}{cm_{20} - cm_{02}} \right) \tag{12}$$

where cm_{ij} is the $(i, j)^{\text{th}}$ central moment of the shape.

The overall similarity distance is defined as

$$D = w_m D_m + w_p D_p \tag{13}$$

where w_m and w_p are weighting constants. Empirically, we find that $w_m = 1$ and $w_p = 0.1$ will give good results to most of the images.

4 Comparisons with the Existing Methods

4.1 Computational complexity

Tables 1 and 2 show the computation operation counts for MFD, FD1, and FD2 in feature extraction and feature matching, respectively.

We can see that although MFD requires a little bit more computation during feature extraction, it is much faster during feature matching. This is because the MFD distance metric is *intrinsically* invariant to translation, rotation, scale, and starting point. This is a very important advantage for the MFD since feature extraction is done off-line while matching is done on-line.

Table 1. Operation counts for feature extraction

	FD1	FD2	MFD
Adds	$O(N_V^2)$	$O(N_V^2)$	$O(N_B \log_2 N_B)$
Mults	$O(N_V)$	$O(N_V)$	$O(N_B \log_2 N_B)$

Table 2. Operation counts for feature matching

	FD1	FD2	MFD
Adds	$O({N_C}^3)$	$Huge^*$	$O(N_C)$
Mults	$O({N_C}^3)$	$Huge^*$	$O(N_C)$

$Huge^*$: beyond comparison since it requires finding all zeros of a trigonometric polynomial of degree N_c.

4.2 Robustness: Practice and Theory

Regardless of the different computational costs, FD1, FD2 and MFD are all valid shape representations, at least theoretically. But to be of practical use, a representation must be tested using the following procedure:

1. Use a camera to take two images of the same physical object, but at different scales, rotations, and translations.
2. Segment the two input images to obtain two shape boundaries, with arbitrary starting point.
3. Compare the features obtained from the each image.
4. If the match is good, conclude that the method is valid.

Note that the segmentation occurs *after* the transformation. This is the actual situation when comparing shapes from two different images. If we use this testing procedure, none of the existing methods give good results, including our proposed MFD method. This is because the boundaries used in these methods are sensitive to *discretization noise*. The discretization noise in many cases changes the boundary enough such that the Fourier coefficients become significantly different. Both FD1 and FD2 suffer from this problem.

MFD also suffers from discretization noise. A simple example illustrates this point (see fig. 1). We discretize the triangle using two different orientations. Note that the upper figure has redundant information (staircase effect) in edge c while the lower figure has redundant information in edges a and b. The Fourier transform magnitudes, as well as $ratio(k)$ (defined in section 3) are shown in fig. 2. Note that the plot of $ratio(k)$ shows a large variance, even though the FFT coefficients were obtained from the *same* object.

We want to solve this problem of spatial discretization while keeping the invariance properties of the MFD; we propose the following procedure:

1. Compute the FFT of the boundary;
2. Use the first $(2N_C + 1)$ FFT terms to form a dense but possibly non-uniformly sampled set of points on the boundary:

$$z_{dense}(n) = \sum_{k=-N_C}^{N_C} Z(k) e^{-j\frac{2\pi nk}{N_B}},\ n = 0, ..., N_{dense} - 1 \qquad (14)$$

3. Use interpolation to trace the dense samples and form samples $z_{unif}(n)$, n = 0, ..., N_{unif} which are uniformly spaced in terms of arc length, estimated from $z_{dense}(n)$;
4. Compute the FFT of $z_{unif}(n)$ to obtain MFD coefficients $Z_{unif}(k), k = -N_C, ..., N_C$.

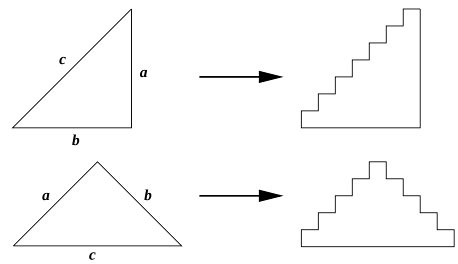
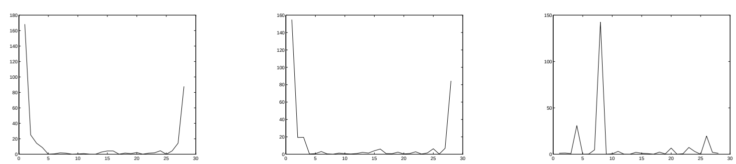
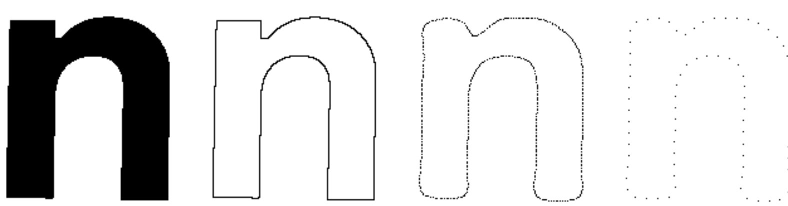


Fig. 1. Effect of spatial discretization on the chain code.



(a) (b) (c)

Fig. 2. (a) FFT magnitude of upper triangle; (b) FFT magnitude of lower triangle; (c) $ratio(k)$ vs. k.



(a) (b) (c) (d)

Fig. 3. (a) Original image; (b) Extracted boundary; (c) Low frequency reconstruction; (d) Uniform re-sampling.

5 Experimental Results

We applied the MFD to our multimedia database prototype, Multimedia Analysis and Retrieval System (MARS). However, to be consistent with other methods, we used several letters of the alphabet as our test set. Images were created by printing out the letters {m, n, u, h, l, t, f} on a laser printer and digitizing the printouts using a scanner. Letters were printed using 256 pt. Helvetica font.

5.1 Sensitivity to choice of parameters

The letters "n" and "f" are used in the following experiments. "n vs. n" denotes the distance between "n" and a rotated version of "n", where the rotation angle is 27 degrees. "n vs. f" denotes the distance between an upright "n" and an upright "f".

1. Sensitivity to N_C

Table 3 shows Distance vs. N_C, where we can see that the MFD is very robust to N_C. We have a wide range to choose N_C from – it can range from 5 to 40 without significantly affecting the matching results for the images we used.

Table 3. Distance vs. N_C

N_C	10	15	20	25	30	35
n vs. n	0.095	0.090	0.059	0.051	0.051	0.051
n vs. f	1.984	1.806	1.930	1.713	1.907	1.705

2. Sensitivity to N_{dense}

N_{dense} is defined as

$$N_{dense} = \frac{boundary\ length}{N_{step}} \tag{15}$$

where N_{step} is the sampling interval. The finer the interval, the larger the number of dense samples. From Table 4 we see that the distance is almost constant for a wide range of N_{step}.

Table 4. Distance vs. N_{step}

N_{step}	2	4	6	8	10	12	14
n vs. n	0.059	0.059	0.059	0.059	0.059	0.059	0.060
n vs. f	1.912	1.912	1.913	1.912	1.931	1.932	1.932

3. Sensitivity to N_{unif}

N_{unif} is defined as

$$N_{unif} = (2N_C + 1)multi \tag{16}$$

where $multi$ makes N_{unif} a multiple of the number of total frequencies used. $multi$ should be at least 1, which corresponds the Nyquist frequency. (see Table 5).

Table 5. Distance vs. $multi$

multi	1	2	3	4	5	6
n vs. n	0.075	0.059	0.060	0.059	0.059	0.060
n vs. f	1.705	1.912	1.911	1.911	1.911	1.911

5.2 Discriminatory ability

Tables 6-8 show the MFD distances between the shapes of each letter from the original set, rotated set (27 degrees), and scaled set (210%).

As expected, “n” and “u” match quite closely, since they are only rotated versions of each other. “h” matches “n” and “u” better than the other letters. We see that discretization (after rotation and scaling) introduces some noise and thus the distances between the same letters are not exactly zero (Tables 7, 8) as is the case in Table 6. But the results indicate that the MFD deals with the discretization effects fairly well. Distances between different letters are always much larger (10 to 100 times) than those between the same letter.

Table 6. Distances between letters – original set.

	m	n	u	h	l	t	f
m	0.000	1.802	1.809	1.625	0.893	1.802	1.512
n	1.802	0.000	0.075	1.439	1.026	1.729	1.907
u	1.809	0.075	0.000	1.483	0.991	1.747	1.852
h	1.625	1.439	1.483	0.000	1.081	1.583	1.557
l	0.893	1.026	0.991	1.081	0.000	1.109	1.077
t	1.802	1.729	1.747	1.583	1.109	0.000	1.260
f	1.512	1.907	1.852	1.557	1.077	1.260	0.000

Table 7. Distances between original and rotated letters.

	m	n	u	h	l	t	f
m	0.085	1.815	1.887	1.550	0.889	1.605	1.545
n	1.795	0.079	0.133	1.448	1.035	1.860	1.736
u	1.805	0.139	0.102	1.492	0.999	1.905	1.735
h	1.619	1.405	1.543	0.068	1.094	1.493	1.603
l	0.837	1.154	1.034	1.077	0.016	1.109	1.070
t	1.808	1.757	1.760	1.586	1.112	0.058	1.262
f	1.512	1.911	1.923	1.571	1.085	1.244	0.040

Table 8. Distances between original and scaled letters.

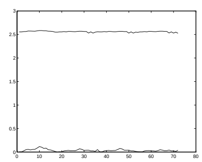
	m	n	u	h	l	t	f
m	0.025	1.873	1.848	2.081	1.762	1.674	1.602
n	1.797	0.023	0.083	1.441	2.342	1.847	1.808
u	1.804	0.080	0.023	1.487	2.374	1.685	1.806
h	1.621	1.324	1.333	0.022	2.017	1.582	1.601
l	0.895	1.028	0.991	1.080	0.012	1.112	1.088
t	1.810	1.730	1.745	1.582	1.382	0.025	1.267
f	1.518	1.911	1.884	1.574	1.232	0.891	0.034

5.3 Robustness to transformation

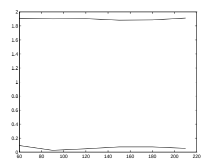
– Translation
  No discretization noise involved. Zero error.
– Rotation
  We plot the distance vs. rotation angle in fig. **??**a. The upper curve is the distance between “f” and rotated versions of “n”. The lower curve is the distance between “n” and its rotated version. The rotation step is five degrees.
– Scale
  We plot distance vs. scale factor in fig. **??**b. The upper curve is the distance between “f” and scaled versions of “n” (from 30% to 210%, with a step size

of 30%). The lower curve is the distance between “n” and scaled versions of “n”. The magnitude difference is also about a factor of 20, indicating that the MFD is scale invariant.

– Starting point
  No discretization noise involved. Zero error.



(a)



(b)

Fig. 4. “n vs. n” and “n vs. f” for various (a) rotation angles; (b) scale factors

6 Conclusions

We presented a new method of shape representation and its distance metric. We compared it with existing FD methods in terms of both computational cost and practical robustness. The main features of our method are:

1. The method is in variant to translation, rotation, scale, and starting point.
2. The method takes into account spatial discretization.
3. The computational cost for feature extraction is low, and for feature matching the cost is extremely low, making the method suitable for real-time multi-user CBR systems.
4. The representation is able to describe complex shapes while remaining relatively compact, reducing the disk space and memory required in the CBR system.

References

1. Faloutsos, C. et. al.: Efficient and Effective Querying By Image Content, IBM Research Report, Aug 3, 1993.
2. John R. Smith and Shih-Fu Chang: Tools and Techniques for Color Image Retrieval, ACM Multimedia '95, San Francisco, Nov 5-9, 1995.
3. H. J. Zhang, et. al. Video Parsing: Retrieval and Browsing: An Integrated and Content-Based Solution, ACM Multimedia'95, San Francisco, Nov 5-9, 1995.
4. Yihong Gong, et. al.: An Image Database System with Content Capturing and Fast Image Indexing Abilities, 0-8186-55-5/94, 1994, IEEE.
5. Michael J. Swain: Interactive Indexing into Image Database, SPIE Vol 1908 (1993)/95.
6. E. Persoon and K. S. Fu: Shape Discrimination Using Fourier Descriptors, IEEE Trans. Sys. Man, Cyb. SMC-7 (March 1977):170-179.
7. C. T. Zahn and R. Z. Roskies: Fourier Descriptors for Plane Closed Curves, IEEE Trans. on Computers, Vol. c-21, No. 3, Mar 1972.
8. C. T. Zahn: A Formal Description for Two-Dimen-sional Patterns, Proc. Int. Joint Conf. Artif. Intelligence, May 1969, pp 621- 628.