

# Fourier descriptors and handwritten digit recognition

Yi Lu\*, Steven Schlosser, and Michael Janeczko

Environmental Research Institute of Michigan, P.O. Box 134001, Ann Arbor, MI 48113-4001, USA

**Abstract.** This paper presents the results of a comparative study of various Fourier descriptor representations and their use in recognition of unconstrained handwritten digits. Certain characteristics of five distinct Fourier descriptor representations of handwritten digits are discussed, and illustrations of ambiguous digit classes introduced by use of these Fourier descriptor representations are presented. It is concluded that Fourier descriptors are practically effective only within the framework of an intelligent system, capable of reasoning about digit hypotheses. We describe a hypothesis-generating algorithm based on Fourier descriptors which allows a classifier to associate more than one digit class with each input. Such hypothesis-generating schemes can be very effective in systems employing multiple classifiers. We compare the performance of the five Fourier descriptor representations based on experiment results produced by a particular hypothesis-generating classifier for a test set of 14 000 handwritten digits. It is found that some Fourier descriptor formulations are more successful than others for handwritten digit recognition.

**Key words:** Fourier descriptors – Handwritten digit recognition – Boundary features – Rotational transformation – Symmetric transformation – Hypothesis generation

## 1 Introduction

It is well known that a closed curve can be represented by a periodic function and, consequently, by a Fourier series. The Fourier coefficients of this function contain information about the size, shape and orientation of the curve; they are often referred to as *Fourier descriptors* of the curve. The Fourier coefficients and the function representing the closed curve uniquely specify each other. Thus, the original closed curve can be reconstructed from its Fourier coefficients (Ballard and Brown 1982; Crimmins 1982). In many applications, a closed curve is represented by a *Fourier descriptor vector* whose components are derived as functions of

a subset of the Fourier coefficients. These frequency-domain descriptions provide an increasingly accurate characterization, as more coefficients are included. Fourier descriptors have proven effective in many problems of pattern classification and computer vision (Arbter and Snyder 1990; Ballard and Brown 1982; Ghorbel et al. 1988; Gonzalez and Wintz 1987; Granlund 1972; Impedovo et al. 1978; Lai and Suen 1981; Lin and Chellappa 1987; Persoon and Fu 1977; Shridhar and Badreldin 1984; Zahn and Roskies 1972). This paper presents a comparative study on the application of Fourier descriptors to the recognition of handwritten digits.

The recognition of unconstrained handwritten digits has received much attention in the last two decades and remains a difficult problem. The application of handwritten digit recognition is broad. Typical uses include recognizing handwritten ZIP codes and reading personal bank checks. The recognition of handwritten digits, like other problems in pattern recognition, consists of two major problems: feature selection and representation, and pattern classification. The classification approaches can be divided into three categories: statistical, syntactic or structural (Fu 1980), and neural network (Le Cun et al. 1989; Wassermann 1989). Feature selection is problem-dependent and considered most central to the final result of a recognition system. Features commonly used in the recognition of handwritten digits are boundaries, skeletons, orientations, strokes and/or cavity features (Ahmed and Suen 1980; Duran and Odell 1974; Fu 1980; Ghorbel et al. 1988; Gonzalez and Wintz 1987; Granlund 1972; Impedovo et al. 1978; Mitchell and Gillies 1989; Schlosser et al. 1991). Fourier descriptors are often used to represent boundary features of digits. Since handwritten digits of the same digit class can occur in great variety, it is desirable to generate a representation that is invariant or, at least, somewhat insensitive to the commonly encountered variations and distortions of digits within the same digit class. Considerable effort has been devoted to the design of Fourier descriptor representations that are invariant under various geometric transformations such as rotation, dilation, and translation. The aim of this study is not to derive a new Fourier descriptor representation, but rather to provide some perspective on the characteristics of various representations of Fourier descriptors with respect to handwritten digits, and to discuss effective digit classification approaches using these perspectives.

\* *Present address:* Department of Electrical and Computer Engineering, University of Michigan-Dearborn, Dearborn, MI 48128-1491, USA

*Correspondence to:* S. Schlosser

In this paper we present five formulations of Fourier descriptors and illustrate ambiguous digit classes associated with these representations. A method for reducing the extent of the digit class ambiguity associated with well-known descriptors is described, and it is concluded that a Fourier descriptor approach may be practically effective for handwritten digit recognition when it is used for focused hypothesis generation within the framework of an intelligent system capable of reasoning about each hypothesis. Consequently, we describe a hypothesis-generating scheme for Fourier descriptor feature-based classifiers. Under this scheme, a classifier generates a hypothesis list for each unknown input. The hypothesis list is constructed to contain the most probable digit classes to which the input belongs, and a numeric value indicating the reliability of each hypothesized classification. The hypothesis list of an unknown input can be passed to other digit recognizers, or it can be further processed by a high-level decision module in combination with results provided by other digit recognizers to produce a unique and reliable system response. The hypothesis-generating scheme can be very useful if a system employs more than one recognition algorithm – then this strategy typically reduces errors by allowing the overall classification process to combine results based on complementary feature sets.

Five Fourier descriptor representations are described in the following, and each combined with the same classification process. Of these, one was proposed by Shridhar and Badreldin (1984), one by Ghorbel et al. (1988), and the other three have been defined by ourselves. Each representation is based on the external boundary curves of digits, but they have different boundary functions and different transforms from Fourier coefficients to Fourier descriptor vectors which result in different ambiguity class characteristics. The five Fourier descriptor representations are evaluated on the basis of experiments which use a common hypothesis-generating classifier. The five Fourier descriptor representations and the hypothesis-generating classifier were all implemented in the C programming language. For each of the five Fourier descriptor representations, the algorithm was trained on a set of 5200 digits. The system results were produced from a test set of over 14 000 well-segmented digits. The digits used are binary images of comparatively low resolution (212 pixels per inch). The digits were manually cropped from handwritten ZIP codes contained within address block images provided by the U.S. Postal Service.

## 2 Fourier descriptor representations of handwritten digits

Fourier descriptors can be used to represent either boundaries or skeletons of digits (Crimmins 1982; Impedovo et al. 1978; Lin and Chellappa 1987). Figure 1 illustrates the boundary and skeleton of a digit. In this paper we confine ourselves to Fourier descriptors that represent digit boundaries. Figure 2 shows the different processing steps involved in obtaining a Fourier descriptor vector from a digit. In the discrete case,

the boundary function of a digit can be written as  $\Gamma(m)$ , where  $m$  represents the  $m$ th point on the boundary. The Fourier series of  $\Gamma(m)$  is defined as

$$\Gamma(m) = \sum_{n=-\infty}^{\infty} a(n)e^{2j\pi nm/L}, \quad \text{and}$$

$$a(n) = \frac{1}{L} \sum_{m=1}^L \Gamma(m)e^{-2j\pi mn/L}$$

where  $L$  is the length of the boundary. A Fourier descriptor vector  $FD = (\nu_1, \nu_2, \dots, \nu_p)$  can be obtained through a set of transform functions  $T_i$  such that

$$\nu_i = T_i(a(0), a(\pm 1), a(\pm 2), \dots, a(\pm k)), \quad i = 1, \dots, p,$$

when  $2k + 1$  initial Fourier coefficients are utilized.

Ideally, digits of the same class should correspond to similar Fourier descriptor vectors and digits of different classes should correspond to dissimilar Fourier descriptor vectors. Handwritten digits, however, are extremely diverse. Digits belonging to the same class can be very different from each other, while on the other hand, digits belonging to different classes can be very similar. Figure 3 shows some examples. The open four in Fig. 3a and the closed four in Fig. 3b are very different from each other, but the open four is similar to an open nine and the closed four is similar to a closed nine. This example shows that a digit class can have many subclasses, and the subclasses can be very different from each other. When the subclasses of each digit class are identified properly, the ambiguity between digit classes can be reduced. In addition to the ambiguous forms of digits, class confusion may result from the use of boundary features which are invariant under certain geometric transformations. Of particular interest in this regard are similarity transforms (euclidean transformations) and symmetric transforms (reflections). The following subsections will examine ambiguous digit classes related to Fourier descriptor features.

### 2.1 Boundary features and functions

Handwritten digits belonging to the same class can be partitioned into subclasses based on boundary features. However, boundary features are limited: digits of different classes can have similar boundary features, although their stroke or cavity features may be different. Classifications based on boundary features alone commonly result in misclassifications in digit-class pairs such as eight and zero, seven and nine, eight and nine and four and eight. Figure 4 shows examples. The boundary of the eight in Fig. 4a is similar to the boundary of a zero, the boundary of the digit nine is similar to the boundary of a seven, the boundary of the eight in Fig. 4c is similar to the boundary of a nine, the boundary of the eight in Fig. 4d is similar to the boundary of an open four, the boundary of the eight in Fig. 4e is similar to the boundary of a closed four, and the boundary of the eight in Fig. 4f is similar to the boundary of a one. These examples indicate that boundary features should be used in combination with

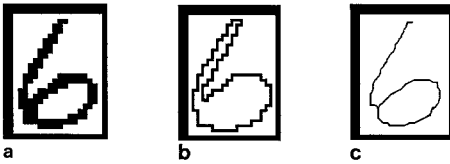


Fig. 1a-c. An illustration of a digit 6, b its boundary, and c the corresponding skeleton

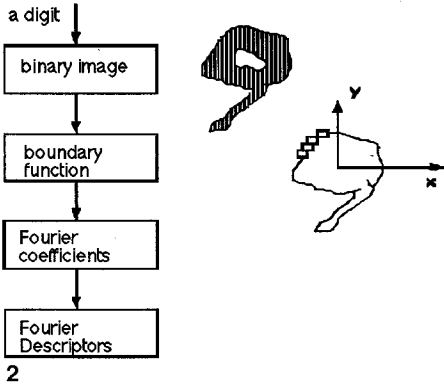


Fig. 2. From a digit image to Fourier descriptors



Fig. 3a-c. The boundaries of a the open four and b the closed four are very different. c The closed four can be confused with a six if the features are invariant under a half-turn

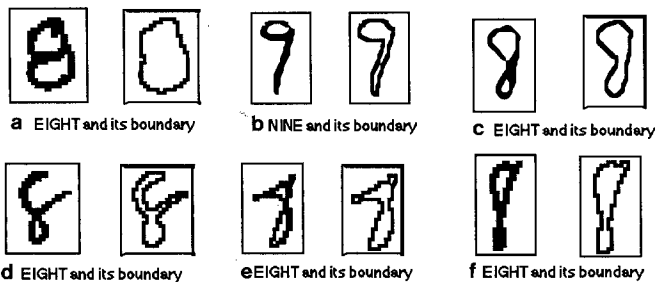


Fig. 4a-f. Digit confusion classes arising from exclusive use of boundary features

other features that are sufficient to differentiate the above ambiguous digit pairs.

Additionally, it is significant that many Fourier descriptor representations have been designed to be insensitive to changes in the starting point of the boundary curve of a digit. This is important for *on-line* applications; for *off-line* applications, it is always possible to normalize the starting point of the boundary curve of a digit.

## 2.2 Similarity transforms

Many of the Fourier descriptor representations applied to recognition of handwritten digits were designed to be in-

variant, or significantly insensitive, with respect to the commonly encountered variations and distortions in digits of the same classes. Typically, Fourier descriptor representations have been formulated to be invariant under similarity transforms – translation, dilation and rotation.

A translationally invariant Fourier descriptor representation maps digits of the same boundary shape, at either the same or different locations, to a unique Fourier descriptor vector. Translational invariance does not cause confusion between different digit classes, and hence is a helpful property of features. Translational invariance of a Fourier descriptor representation can be obtained either by using a translationally invariant function to represent the boundary features, or by applying a translationally invariant mapping from Fourier coefficients to Fourier descriptor vectors.

If a Fourier descriptor representation is invariant under dilation, then digits having the same boundary shape that are of the same or different size will correspond to the same Fourier descriptor vector. Dilational invariance does not cause confusion between different digit classes, and can normally be obtained through a dilationally invariant mapping from Fourier coefficients to Fourier descriptor vectors.

A rotationally invariant Fourier descriptor representation associates digits having the same boundary shape but different orientation with the same Fourier descriptor vector. The rotational invariance of a Fourier descriptor representation can be obtained through either a boundary function or a transform from Fourier coefficients to Fourier descriptor vectors. A rotationally invariant transform can generate ambiguous digit class pairs such as six and four, six and nine, two and seven, nine and three and five and nine. Figure 3c shows an example of a rotated four matching a six. The ambiguity caused by rotational invariance between digit classes six and nine is a serious problem. Almost all sixes can be identified as nines after rotation through 180 deg (see Fig. 5a), and all closed nines can be identified as sixes after rotation through 180 deg (see Fig. 5b). Figure 5c shows that a two can be identified as a seven after rotation by 180 deg, while Fig. 5d shows that a five can be identified as a nine after rotation by 180 deg, and Fig. 5e shows that a three can be identified as a nine after rotation by 180 deg.

## 2.3 Symmetric transforms

The symmetric transforms of interest operate about either the  $x$ -axis or the  $y$ -axis. Invariance of digit representations under symmetric transforms can cause ambiguity between different digit classes. Since no digits in the same class are symmetric maps of the others, symmetric invariance of Fourier descriptors is not desirable.

If a Fourier descriptor representation is invariant under a symmetric transform about the  $x$ -axis, the Fourier descriptor vectors are identical for certain digits in the class pairs of two and five, zero and six, two and six, three and seven, and seven and five. If a Fourier descriptor representation is invariant under a symmetric transform about the  $y$ -axis, the Fourier descriptor vectors are identical for certain digits

in the class pairs of two and five, zero and nine, and two and nine. We use symmetric images of digits to illustrate these ambiguous digit classes. Figure 6a–e shows examples of ambiguous digit classes caused by a symmetric transform about the *x*-axis. A two becomes a five in Fig. 6a, a zero becomes a nine in Fig. 6b, a two becomes a nine in Fig. 6c, a three becomes a seven in Fig. 6d, and a seven becomes a five in Fig. 6e. Figure 6f–h shows examples of ambiguous digit classes caused by a symmetric transform about the *y*-axis. A two becomes a five in Fig. 6f, a zero becomes a six in Fig. 6g, and a two becomes a six in Fig. 6h.

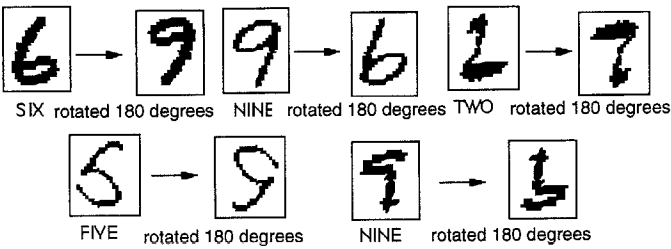


Fig. 5a–e. Examples of ambiguous digit classes arising from rotational invariance of boundary descriptions

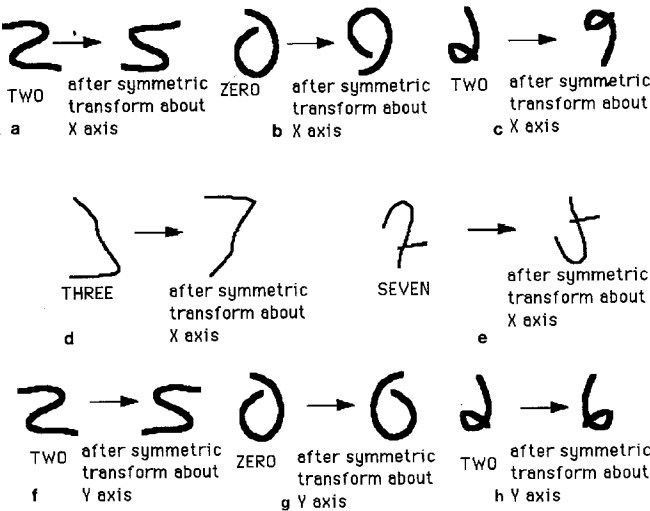


Fig. 6a–h. Examples of ambiguous digit classes arising from symmetric invariance of boundary descriptions

### 2.4 Cataloging ambiguous digit classes

As noted above, Fourier descriptor representations typically give rise to ambiguous digit-class pairs. It is possible to systematically catalog the digit-class ambiguities of a particular application by enumerating the different input forms for the digits, and then systematically applying geometric transformations (primitives and their composites) to the digit boundaries. The transitive nature of geometric similarity may be used to abbreviate the cataloging process.

Figure 7a–c shows ambiguity pairs that can arise from basic geometric transformations such as rotation. Figure 7d,e

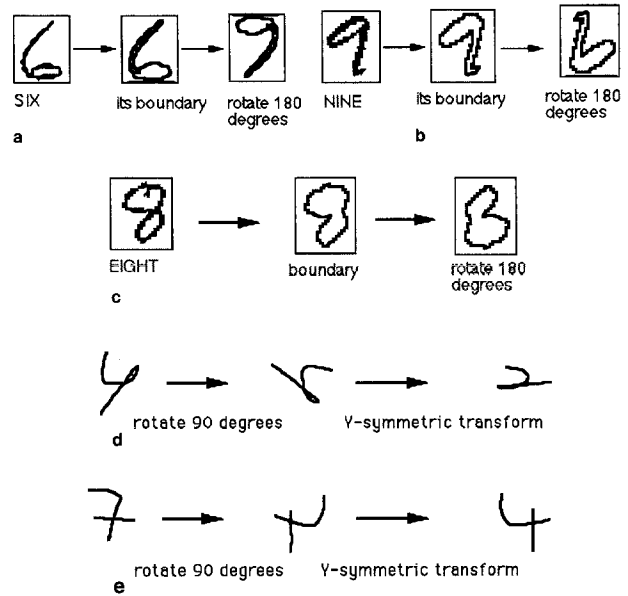


Fig. 7a–e. Examples of ambiguous digit classes caused by the combination of boundary features and various transforms

illustrates ambiguity pairs which could be identified through composition of transformations. These examples show seven and (a particular type of) four and (a different type of) four and two to be possible confusion pairs. Additional confusion pairs can be identified by relating such digit subclasses through transitivity. These examples also illustrate that the existence of ambiguity is dependent upon the form of the input data – different writers present different ambiguity pairs.

### 2.5 Recommendations for the use of Fourier descriptors

The invariant properties of a Fourier descriptor representation should be carefully examined in relation to each application's particular requirements. The similarity transforms are important when the input digits to a Fourier descriptor representation are not normalized, or the Fourier descriptor representation is used for preprocessing. For this type of application, Fourier descriptors can be used to group digits into subclasses while other features of digits can be used to classify the ambiguous digit subclasses caused by similarity invariance. Fourier descriptor representations invariant under similarity transformations usually have fewer different feature vector components; hence, they require less storage space and computational time during classification. If a recognition system has a preprocess which normalizes digits in size, location and orientation (Casey 1970), then a Fourier descriptor representation that is sensitive to rotation should be considered in order to avoid ambiguity between the rotationally invariant digit (sub)classes. The symmetrically invariant transforms should be avoided whenever possible.

### 3 Classification algorithm based on Fourier descriptors

Our discussion in the previous section suggests that Fourier descriptor features should be used in conjunction with other, complementary features. A handwritten digit classifier based on Fourier descriptor features alone will generate a high error rate if a unique response is required for every input. In the following sections we describe a hypothesis-generating strategy which is effective for use in recognition systems having multiple classifiers and for use with other classifiers that cannot uniquely identify all digit classes.

#### 3.1 Generating hypotheses for classifications

This approach requires a digit classifier to generate a hypothesis list for each input digit. A hypothesis list is a list of digit classes which strongly match the unknown input. Additionally, each digit class on the list is associated with a numeric value indicating the confidence of the classification. Such a digit classifier can be evaluated according to two measures: reliability and efficiency. Reliability is the frequency with which the correct class is included in the hypothesis list. Efficiency is the average length of the hypothesis list. An effective classifier will produce short hypothesis lists which are correct. That is, good classifiers are efficient and reliable.

A handwritten digit classifier should be able to capture the dissimilarity between digits of different classes and the similarity between digits of the same class, as well as describe possible misclassifications and indicate "I don't know" for inputs for which it cannot find good classifications. The hypothesis-generating strategy provides these capabilities. It reduces erroneous responses by allowing a classifier to produce multiple responses when a unique response cannot be determined. The confidence values permit identification of ambiguous digit classes and situations like "I don't know". If the difference between the highest and the second highest confidence value is large, it indicates that the class with the highest confidence value is almost certainly the true class of the input; if the first  $k$  ( $k > 1$ ) classes on the hypothesis list have very close confidence values, it indicates that the input is ambiguous; if the highest confidence value on the hypothesis list is very low, it is an indication of "I don't know". Obviously, if it is desired, a unique response can be derived from a hypothesis list by applying a decision process to the hypothesized classes and their associated confidence values. In a system of multiple digit classifiers, it is possible either to pass a hypothesis list generated by a digit classifier on to other digit classifiers for further processing, or to apply a decision process to all the hypothesis lists generated by the classifiers in the system to achieve a reliable and unique classification.

#### 3.2 A classification algorithm based on Fourier descriptors

The classification algorithm consists of the following processes:

- Boundary delineation

- Calculation of the Fourier descriptor vector from the boundary
- Construction of the reference feature vector set
- Generating hypotheses for the input digits

In our implementation, we apply a *border-following* algorithm proposed by Rosenfeld and Kak (1976) to extract the external boundary points of a digit. The sequence of the boundary points is taken in counterclockwise order. The starting point is defined as the first boundary point encountered as the boundary image is scanned from top to bottom, left to right. The Fourier descriptor vector of a digit is generated by the formula provided by the corresponding Fourier descriptor representation.

The feature vector set is generated by applying a hierarchical clustering technique to the Fourier descriptor vectors of the digits in the training set. The clustering algorithm was adapted from Duran and Odell (1974). A number of clusters are generated for each digit class, and each cluster represents a subclass of the digit class. The mean vector of each cluster defines a Fourier descriptor (reference) feature vector, and the class of the majority of the digits contained in the cluster provides the truth class of the feature vector.

The hypothesis list of an input digit is generated as follows. For every input digit, its Fourier descriptor vector is matched with all vectors in the feature vector set. The feature vectors whose Euclidean distance to the Fourier descriptor vector of the input digit is less than a threshold are collected as the matched feature vectors. The confidence value of a matched feature vector is derived as the reciprocal of the distance measure between the input Fourier descriptor vector and the matched feature vector. The distinct classes associated with the matched feature vectors and their corresponding confidence values form the hypothesis list for the input digit.

The algorithm was implemented and used to generate experimental results for the evaluation of the five Fourier descriptor representations introduced in the following section.

### 4 Five Fourier descriptor representations and their performance

In this section we examine five Fourier descriptor representations; each descriptor title has been abbreviated in an evident fashion to indicate its origin: SB, GCT, ERIM1, ERIM2 and DFT, the classic discrete Fourier transform. Their individual definitions are provided below. All five Fourier descriptor representations use boundary features of digits, but differ either in digit boundary functions and/or the transform from Fourier coefficients to Fourier descriptor vectors.

In order to convey the characteristics of the five Fourier descriptor representations, in addition to theoretical analyses, we implemented all five Fourier descriptor representations and computed the Fourier descriptor vectors of five images shown in Fig. 8. Figure 8a is the boundary image of a two, Fig. 8b is the same as Fig. 8a except that the starting point,

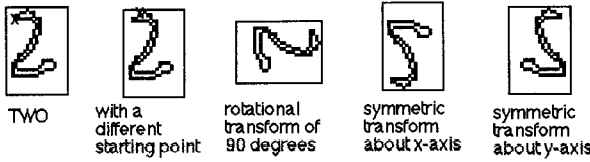


Fig. 8a-e. A set of test images

marked as  $x$ , is different, Fig. 8c is obtained by rotating the digit in Fig. 8a through 90 deg and Fig. 8d,e by symmetrical transformation of Fig. 8a about the  $x$ -axis and the  $y$ -axis, respectively. For each of the five Fourier descriptor representations, the Fourier descriptor vectors of the five images are displayed in the graphs in Fig. 9. In each graph, The Fourier descriptor vectors of the five images in Fig. 8 are represented as Vectors 1 to 5. The horizontal axis represents the indices of the elements in the Fourier descriptor vectors, and the vertical axis indicates the values of the elements of the vectors.

We also implemented the hypothesis-generating algorithm presented in the preceding section for each of the five Fourier descriptor representations. We will compare the five Fourier descriptor representations as an element of the classifier, based on our experiment results.

#### 4.1 SB representation

The SB representation was developed by Shridhar and Badreldin (1984). In this representation, the boundary function is defined by the Cartesian coordinates  $\{(x(m), y(m)) \ni m = 1, 2, \dots, L\}$ , where  $L$  is the length of the boundary curve of the digit. The Fourier series of  $x(m)$  and  $y(m)$  is

$$x(m) = \sum_{n=-\infty}^{\infty} a(n)e^{2j\pi nm/L}, \quad \text{and}$$

$$y(m) = \sum_{n=-\infty}^{\infty} b(n)e^{2j\pi nm/L}$$

where  $a(n)$  and  $b(n)$  are the Fourier coefficients:

$$a(n) = \sum_{m=1}^L x(m)e^{-2j\pi nm/L}, \quad \text{and}$$

$$b(n) = \sum_{m=1}^L y(m)e^{-2j\pi nm/L}.$$

The SB representation of a digit is a Fourier descriptor vector. The elements of the vector are derived as

$$S(n) = \frac{r(n)}{r(1)},$$

where

$$r(n) = \sqrt{|a(n)|^2 + |b(n)|^2}, \quad n = 1, 2, \dots$$

Based on an analysis of digits reconstructed from Fourier coefficients, Shridhar and Badreldin suggest the use of Fourier

descriptor vectors of dimension 10 to 15 for handwritten digit recognition.

In this representation, both the boundary function and the Fourier coefficients contain the location, orientation, and scale information about the original digit. Shridhar and Badreldin showed that their Fourier descriptor vectors are invariant to rotation, translation, and dilation. Figure 9a shows that the SB representations of all the images are identical. It illustrates that the SB representation is invariant to not only rotation, translation, and dilation but also the symmetric transforms and change of starting point. These properties of the SB representation can be derived, in the general case, from the mapping from Fourier coefficients to Fourier descriptor vector. The SB representation cannot discriminate any of the ambiguous digit-class pairs listed in Sect. 2.

#### 4.2 GCT representation

The GCT representation was proposed by Ghorbel, Cazuguel and Tocnaye (1988). They used a radius function  $\rho(l)$  to represent a closed boundary curve, where  $\rho(l)$  measures the length of the line connecting the boundary point  $l$  to the centroid of the boundary curve (see Fig. 10a). The GCT Fourier descriptor representation of a digit is a vector with elements  $\phi(k)$ ,

$$\phi(k) = \frac{1}{L} \int_0^L \log \rho(l) e^{-2j\pi lk/L} dl,$$

where  $L$  is the length of the curve,  $k = 1, 2, \dots$ . Because the mapping from the digit boundary to the boundary function  $\rho(l)$  is not one-to-one, the digit boundary cannot be reconstructed from the Fourier coefficients, although the GCT Fourier descriptor representation and the boundary function have a one-to-one relationship. The boundary function  $\rho(l)$  is invariant under symmetric, translational, and rotational transforms, but is sensitive to change in the size of the closed curve. Ghorbel et al. showed through an example that the GCT representation is slightly sensitive to dilation. Our experiments show (Fig. 9b) that the GCT representation is also slightly sensitive to symmetric transforms. Based on the above discussion, we conclude that the GCT representation cannot discriminate the ambiguous digit-class pairs listed in Sect. 2.

#### 4.3 ERIM1 representation

This Fourier descriptor representation is one of two used by ourselves. The motivation in using this descriptor is to recover orientation information from the SB representation. The ERIM1 representation is represented by a vector with elements derived as

$$S_0^0 = \left| \frac{a(1)}{r(1)} \right|, \quad \text{and} \quad S_1^0 = \left| \frac{b(1)}{r(1)} \right| \quad \text{with} \quad S_n^0 = S(n)$$

where  $a(1)$ ,  $b(1)$ ,  $S(n)$ ,  $n = 2, 3, \dots$  are the same as those defined in the SB representation. Our experiments (see the

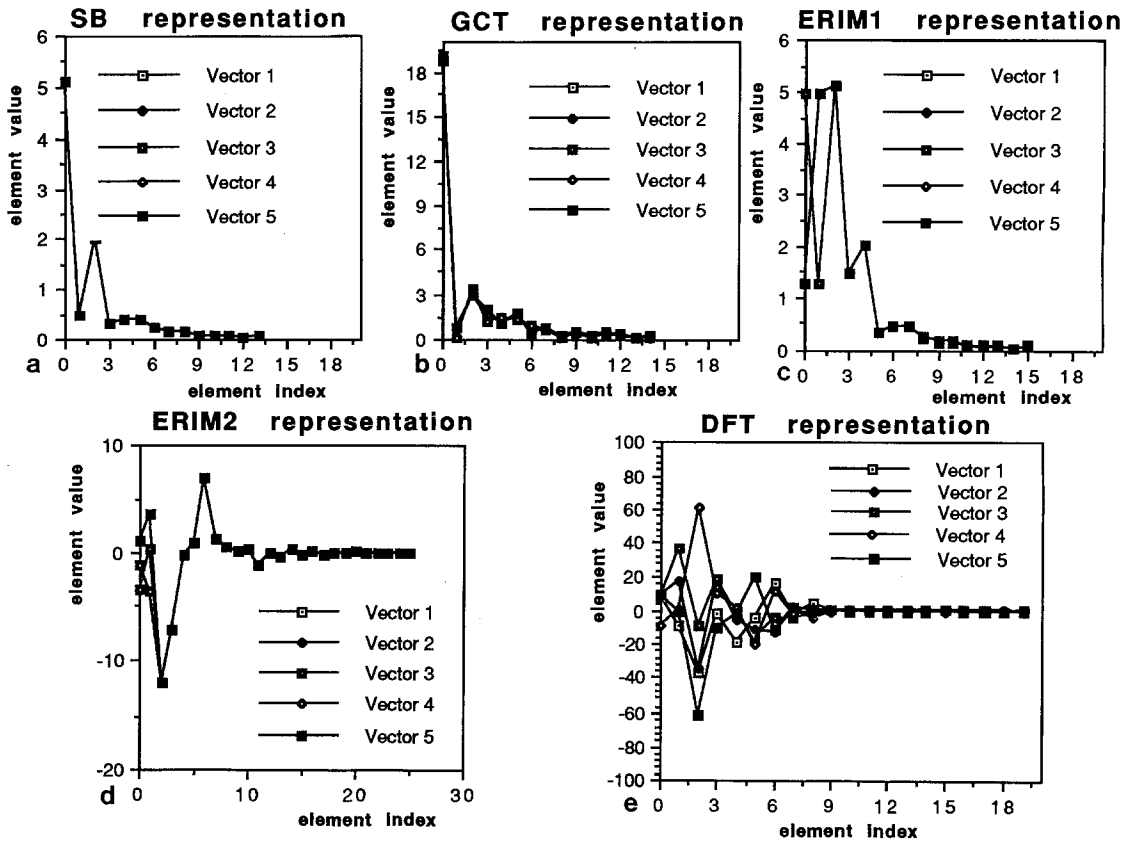


Fig. 9a-e. Graphical illustration of the Fourier descriptor vectors

transforms. This suggests that the ERIM1 Fourier descriptor representation may discriminate the digit-class pairs where ambiguity arises from rotational similarity.

#### 4.4 ERIM2 representation

This Fourier descriptor representation was also used by ourselves. In the ERIM2 representation, the boundary function of a digit describes the orientation of the vector connecting a boundary point to the centroid of the boundary curve. Let a boundary curve be  $\{(x_i, y_i) \ni i = 0, 1, \dots, L-1\}$ . Assume the  $xy$  coordinates have their origin at the centroid of the boundary curve,  $(x_0, y_0)$  is the starting point on the boundary curve, and the boundary points are continuous and taken in counterclockwise order. Mathematically, the boundary function  $\phi$  is defined as

$$\phi(x_i, y_i) = \begin{cases} \frac{\pi}{2}, & \text{when } x_i = 0 \\ \arctan\left(\frac{y_i}{x_i}\right), & \text{when } x_i > 0 \text{ and } y_i \geq 0 \\ \pi + \arctan\left(\frac{y_i}{x_i}\right), & \text{when } x_i < 0 \text{ and } y_i \geq 0 \\ \pi - \arctan\left(\frac{y_i}{x_i}\right), & \text{when } x_i < 0 \text{ and } y_i \leq 0 \\ -\arctan\left(\frac{y_i}{x_i}\right), & \text{when } x_i > 0 \text{ and } y_i \leq 0 \end{cases}$$

The Fourier expansion of  $\phi$  is

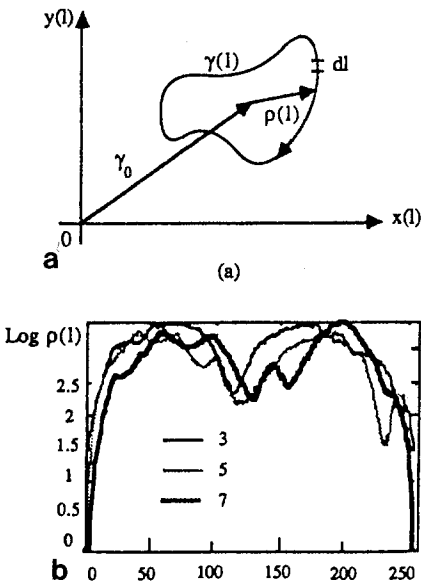


Fig. 10. a Boundary curve description; b developed digit contours [adapted from Ghorbel et al. (1988)]

graph in Fig. 9c) show that the Fourier descriptor vectors in this representation are invariant under symmetric transforms and change of starting point, and only the first two elements of the Fourier descriptor vectors are sensitive to rotational

$$\phi(x_m, y_m) = \sum_{n=-\infty}^{\infty} a(n)e^{2j\pi nm/L}, \quad m = 0, 1, \dots, L-1, \quad \text{and}$$

$$a(n) = \sum_{m=1}^L x(m)e^{-2j\pi nm/L}, \quad n = 0, \pm 1, \pm 2, \dots$$

The Fourier descriptor vector for a digit is obtained by using the following formula proposed by Granlund (1972):

$$Q_1 = \frac{a_2|a_1|}{a_1^2}, \quad \text{and} \quad Q_n = \frac{a_{1+n}a_{1-n}}{a_1^2}, \quad \text{for } n = 2, 3, \dots$$

Granlund showed that  $Q_n$ ,  $n > 1$ , is invariant under rotation, translation, dilation, and change of starting point on the boundary, while  $Q_1$  is invariant under translation and dilation but not rotation. Figure 9d shows the graphical display of this representation for the images in Fig. 8. Note the elements in a Fourier descriptor vector are complex numbers. Let  $\text{Real}Q_i$  and  $\text{Img}Q_i$  represent the real and imaginary part of the  $i$ th element in a vector. Each vector in Fig. 9d has the following form:  $(\text{Real}Q_1, \text{Img}Q_1, \dots, \text{Real}Q_{13}, \text{Img}Q_{13})$ . The graph in Fig. 9d indicates that the ERIM2 representation is sensitive to neither change in the boundary starting point nor symmetric transform about the  $x$ -axis, but ERIM2 is sensitive to both rotation and symmetric transformation about the  $y$ -axis.

#### 4.5 DFT representation

This representation is based on the classic discrete Fourier transform (Gonzalez and Wintz 1987). Our goal in using this representation has been to eliminate the ambiguous digit-class pairs caused by both rotational and symmetric invariant transforms. In this representation, we consider the boundary curve of a digit as being on the complex plane. The boundary function is defined as

$$F(m) = (x(m) - x_c) + j(y(m) - y_c)$$

where  $x(m)$  and  $y(m)$  are the Cartesian coordinates of the  $m$ th point on the boundary curve,  $(x_c, y_c)$  is the centroid of the digit boundary, and  $L$  is the length of the boundary curve. The Fourier series of  $F(m)$  is

$$F(m) = \sum_{n=-\infty}^{\infty} c(n)e^{2j\pi nm/L}$$

where  $c(n)$  are Fourier coefficients derived as

$$c(n) = \sum_{m=1}^L [(x(m) - x_c) + j(y(m) - y_c)] e^{-2j\pi nm/L}$$

The representation of a digit is a vector of Fourier coefficients,  $c(n)$ . Since both the mapping between the boundary and the boundary function  $F(m)$  and the mapping between the Fourier series and  $F(m)$  is one-to-one, this representation does not lose or gain any boundary feature information. The graphs in Fig. 9 illustrate that this representation is

sensitive to rotation transforms as well as symmetric transforms about the  $x$ -axis and the  $y$ -axis. Therefore it is capable of distinguishing all the ambiguous digit-class pairs arising from these invariances. However, the DFT is also very sensitive to variations of the boundaries of digits of the same classes. Consequently, this transform requires a particularly large number of reference vectors, or archetypes.

#### 4.6 Summary

We have presented five different digit representation schemes based on Fourier descriptors. Since all five representations are based on boundary curves of digits, they all may give rise to ambiguous digit classes. The SB representation and GCT representation are invariant to rotation transforms, symmetric transforms, and change of the boundary starting point. Neither of these two representations can differentiate any of the ambiguous digit-class pairs presented in Sect. 2. The ERIM1 representation is sensitive to rotational transforms. Therefore, it eliminates the confused digit-class pairs which could arise from rotational invariance, and those ambiguities which could arise from a combination of rotational and  $y$ -axis symmetric invariance. However, digit-class confusion pairs can arise with this representation as a result of  $x$ -axis symmetric invariance. ERIM2 permits  $x$ -axis symmetric invariance, as does ERIM1. But, because it is sensitive to both rotation and  $y$ -axis symmetric transformations, fewer confusion pairs may arise with ERIM2 than ERIM1. The classic DFT representation is sensitive to rotational and symmetric transforms. The digit-class confusion pairs arising under this representation are only those caused by the boundary features. The classic DFT representation gives good performance if the input digits are normalized in orientation and in size prior to subsequent processing.

#### 4.7 Performance of the five Fourier descriptor representations

For each of the five Fourier descriptor representations, reference feature vectors were developed from a training set containing 520 digits for each digit class. The test set contains more than 14 000 well-segmented digits. The digits in both the training and test sets were normalized in size, location, and orientation. The orientation normalization adjusts the digits skewed by less than  $\pm 45$  deg.

The performance of the five Fourier descriptor representations is shown in Table 1. The dimension of Fourier descriptor vectors for SB and GCT was 14, and for ERIM1, ERIM2, and DFT it was 16, 26, and 28, respectively. The digit classifier based on the SB representation had the highest percentage of correct hypothesis lists, the smallest set of reference feature vectors (that is, archetypes), but the longest hypothesis lists. It generated more than five classes, on average, for each input. This experimental result is consistent with our previous discussion. Since the SB representation is invariant under translation, rotation, dilation, and symmetric transformation, and is also insensitive to any change in the



boundary starting point, each feature vector is expected to match a variety of digit types. On the other hand, feature vectors derived from different digit classes can be close to each other under the euclidean metric. That is why the average number of matched classes, for each input, is large. Although the GCT representation has properties similar to the SB representation, the performance of the GCT representation was not as good. In our experiment, the GCT representation had a larger set of feature vectors and was significantly lower (in fact, the lowest of all the five representations) in number of correct hypothesis lists, although its average hypothesis list length was a little shorter (close to five classes for each input digit) than that of SB. The performance of the ERIM1 representation was encouraging. It generated significantly shorter hypothesis lists, and its percentage of correct hypothesis lists was close to the SB representation. We note that since only the first two elements of a vector in the ERIM1 representation are sensitive to rotation, better performance should be achieved if two-stage matching is performed. In this case, in the first stage it would be possible to match the last 14 elements of the vectors, and in the second stage, the first two elements. Since the ERIM2 representation is more sensitive to the variations of digits, it had fewer ambiguous digit classes, a larger set of feature vectors, and consequently it generated rather short hypothesis lists. Since the DFT representation is sensitive to rotation and symmetric transformation about both the  $x$ -axis and the  $y$ -axis, it had the fewest ambiguous digit classes, generated a high percentage of correct hypothesis lists and the shortest hypothesis lists. Since the DFT has a large set of feature vectors and the dimension of its Fourier descriptor vectors is high, it requires more computation.

**Table 1.** Experimental results using each of the five Fourier descriptor representations

Fourier descriptor representation	Correct hypothesis lists	Incorrect hypothesis lists	Hypothesis list mean length	Number of archetypes
SB	99.8%	0.2%	5.6	418
GCT	98.6%	1.4%	4.7	422
ERIM1	99.2%	0.8%	3.3	650
ERIM2	98.8%	1.2%	2.2	710
DFT	99.5%	0.5%	1.5	1235

## 5 Conclusion

We have discussed the properties of Fourier descriptors and their application to handwritten digit recognition. We have illustrated possible ambiguous digit classes which arise from (1) reliance upon the boundary features of the digits and (2) invariance of the boundary features to similarity and symmetric transformations. We observed that a Fourier descriptor representation should be sensitive to symmetric transforms. And, we concluded that it is not realistic to force a unique classification for every input from a Fourier descriptor-based

digit classifier. However, such a digit classifier can be used effectively to generate hypotheses for a digit recognition system. We presented a hypothesis-generating strategy to allow a classifier to produce more than one possible classification for each input. To permit an empirical test, we examined five different versions of Fourier descriptors, and compared their performance. Two of the five representations, the SB representation and the GCT representation, were invariant to rotation, translation, dilation, and symmetric transformation, and were insensitive to changes of the boundary starting point. The other three representations, ERIM1, ERIM2 and DFT, were more sensitive to the variations of digits. Based on our study, we conclude that a Fourier descriptor-based classifier is a well-behaved classification approach (for digits whose boundaries are intact). By analyzing the digit feature a Fourier descriptor represents, as well as the sensitivity of that representation to rotation and symmetric transforms, it is possible to identify the misclassifications that can be generated by the classifier. A Fourier descriptor-based digit classifier can either serve as a preclassifier in a digit recognition system, or it can be combined with other digit classifiers that are based on complementary digit features. In either case, the hypothesis-generating scheme described in this paper is effective. The choice among different types of Fourier descriptor representations is dependent upon the application. If the input digits to digit classifiers vary enormously in orientation and size, the SB representation is a good choice. However, it requires additional digit classifiers as part of a system to discriminate the digit-class pairs listed in Sect. 2. Shridhar and Badreldin (1984) obtained a good result by using the SB representation in a two-stage classification system. If the ERIM1 representation is used, the classifier can eliminate the ambiguous digit-class pairs arising from rotational invariance. If the input digits are normalized in size and orientation before the digit classification, the ERIM2 or the DFT representation is appropriate for use. Our experiments show that although a DFT-based representation has a large feature vector set (many archetypes) and long Fourier descriptor vectors, it is reliable and efficient.

*Acknowledgements.* This work was supported by the Office of Advanced Technology, United States Postal Service, Contract no. 104230-86-H-0042. We would like to thank Dr. Gary Herring and Mr. Carl O'Connor of the Office of Advanced Technology and Dr. John Tan of Arthur D. Little, Inc. for their support. We would also like to acknowledge the assistance of Dr. Brian Mitchell, with whom we had many fruitful discussions.

## References

- Arbter K, Snyder W (1990) Application of Affine-Invariant Fourier Descriptors to Recognition of 3-D Objects. *IEEE Trans on Pattern Analysis and Machine Intelligence* 12(7):640-647
- Ahmed P, Suen CY (1980) Computer Recognition of Totally Unconstrained Handwritten ZIP Codes. *International Joint Conference on Pattern Recognition and Artificial Intelligence*, pp 1-15
- Ballard D, Brown C (1982) *Computer Vision*. Prentice-Hall, Englewood Cliffs, New Jersey

- Casey RG (1970) Moment Normalization of Handprinted Characters. *IBM Journal of Research and Development* 14(5):548–557
- Crimmins T (1982) A Complete Set of Fourier Descriptors for Two-Dimensional Shapes. *IEEE Transactions on Systems, Man and Cybernetics* 2(6):848–855
- Duran BS, Odell P (1974) *Cluster Analysis*. Springer-Verlag, New York
- Fu KS (1980) Recent Developments in Pattern Recognition. *IEEE Transactions on Computers* 29(10):845–854
- Gader P, Forester B, Ganzberger M (1991) Recognition of Handwritten Digits Using Template and Model Matching. *Pattern Recognition* 24(5):421–432
- Ghorbel F, Cazuguel G, Tognaye JLB (1988) Similarity-Invariant Analysis of Handwritten ZIP Codes Using Fourier Descriptors. *International Journal of Reserach & Engineering in Postal Applications* 1(1):1–5
- Gonzalez C, Wintz P (1987) *Digital Image Processing*. Second Edition, Addison-Wesley Publishing Company, Reading, MA
- Granlund GH (1972) Fourier Preprocessing for Hand-Print Character Recognition. *IEEE Transactions on Computers* 21(2):195–201
- Impedovo S, Marangelli B, Fanelli A (1978) A Fourier Descriptor Set for Recognizing Nonstylized Numerals. *IEEE Transactions on Systems, Man and Cybernetic* 8(8):640–645
- Lai M, Suen C (1981) Automatic Recognition of Characters by Fourier Descriptors and Boundary Line Encodings. *Pattern Recognition* 14(6):383–393
- Le Cun Y, Boser B, Denker JS (1989) Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation* 1(4):541–551
- Lin CC, Chellappa R (1987) Classification of Partial 2-D Shapes Using Fourier Descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9(5):686–690
- Mitchell B, Gillies A (1989) A Model-Based Computer Vision System for Recognizing Handwritten ZIP Codes. *Machine Vision and Applications* 2(4):231–243
- Persoon E, Fu KS (1977) Shape Discrimination Using Fourier Descriptors. *IEEE Transactions on Systems, Man and Cybernetics* 7(3):170–179
- Rosenfeld A, Kak A (1976) *Digital Picture Processing*. Academic Press, New York
- Shridhar M, Badreldin A (1984) High-Accuracy Character Recognition Algorithm Using Fourier and Topological Descriptors. *Pattern Recognition* 17(5):515–524
- Schlosser S, Trenkle J, Vogt R (1991) Morphological Feature Set Optimization Using the Genetic Algorithm. In: *SPIE Image Algebra and Morphological Image Processing II*, San Diego, CA
- Wasserman PD (1989) *Neural Computing: Theory and Practice*. Van Nostrand-Reinhold, New York
- Zahn C, Roskies R (1972) Fourier Descriptors for Plane Closed Curves. *IEEE Transactions on Computers* 21(3):269–281