

AD-A132 842

SHAPE RECOGNITION AND DESCRIPTION: A COMPARATIVE STUDY

1/3

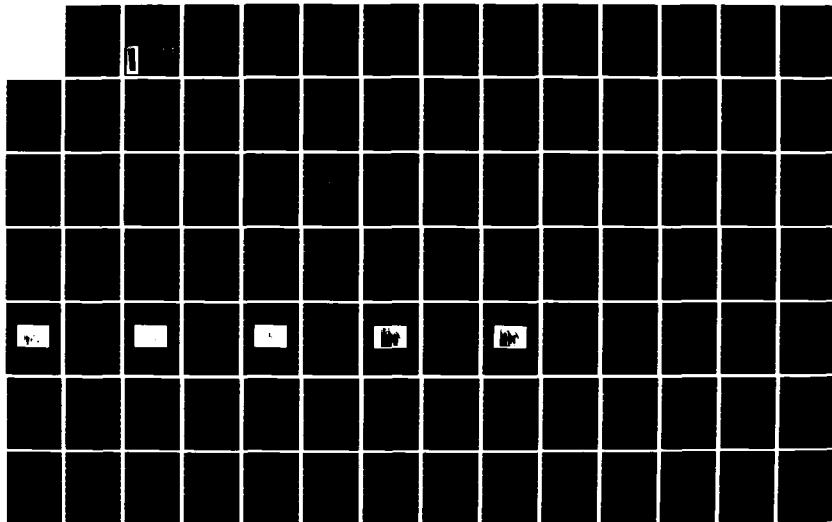
(U) PURDUE UNIV LAFAYETTE IN SCHOOL OF ELECTRICAL
ENGINEERING T A GROGAN ET AL. JUL 83 TR-EE83-22

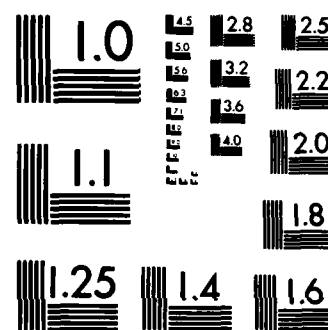
UNCLASSIFIED

ARO-18016. 11-EL-A DAAG29-81-K-0088

F/G 12/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

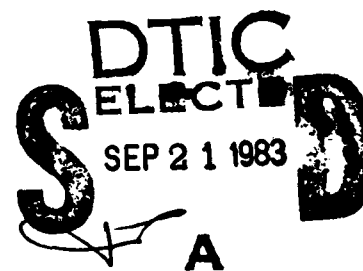
AD - A132842

DTIC FILE COPY

Shape Recognition and Description: A Comparative Study

Timothy A. Grogan
O. Robert Mitchell

TR-EE 83-22
July 1983



This document has been approved
for public release and sale; its
distribution is unlimited.

School of Electrical Engineering
Purdue University
West Lafayette, Indiana 47907

83 09 20 198

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
18016.11-EL-A	AD-A132842	
4. TITLE (and Subtitle) Shape Recognition and Description: A Comparative Study		5. TYPE OF REPORT & PERIOD COVERED Technical In house report
		6. PERFORMING ORG. REPORT NUMBER TR-EE 83-22
7. AUTHOR(s) Timothy A. Grogan and O. Robert Mitchell		8. CONTRACT OR GRANT NUMBER(s) DAA629-81-K-0088
9. PERFORMING ORGANIZATION NAME AND ADDRESS School of Electrical Engineering Purdue University West Lafayette, IN 47907		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE July 1983
		13. NUMBER OF PAGES 258
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) NA		
18. SUPPLEMENTARY NOTES The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Shape recognition, Fourier descriptors, Moments, partial shape recognition, curvature function, Walsh transform, aircraft recognition.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) An important problem in the extraction of information from images is shape recognition. Several methods of analyzing binary images using global shape methods based upon functional approximation have been reported in the literature. However, there has been a lack of information comparing the effectiveness of these methods in shape analysis. Five methods of global shape analysis are compared on two basis: The five methods compared are 1) Fourier descriptors of the boundary; 2) Walsh points of the boundary; 3) the cumulative angular deviant Fourier descriptors; 4) moments of the silhouette; and 5) moments of the boundary.		

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

6

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

First, the different methods are introduced, their geometric properties presented, and the formulae for some generic shapes are provided. Then the methods are compared on the basis of the empirical facts derived from a set of aircraft shape recognition experiments. The shapes are different views of six aircraft. The aircraft silhouettes are the two-dimensional projected images of three-dimensional rigid bodies. The five methods are ranked according to their performance from these experiments. A new method for the recognition of partial shapes based on the Fourier-Mellin transform is introduced. A shift and scale invariant correlation of the complete and partial shape's curvature functions is obtained by applying the Mellin transform to the magnitude of the Fourier transform of the curvature functions. The logarithm of the shift in the correlation function corresponds to the time scaling of the partial shape's curvature function. Then the shift in an ordinary correlation of the complete shape's curvature is the time shift necessary to complete the alignment. Then a pointwise comparison of the curvature functions can be made to determine matching and non-matching contour segments. Some initial recognition experiments for partial shapes are carried out and the results reported.

**SHAPE RECOGNITION AND DESCRIPTION:
A COMPARATIVE STUDY**

Timothy A. Grogan
O. Robert Mitchell

School of Electrical Engineering
Purdue University
West Lafayette, Indiana 47907

TR-EE 83-22

July 1983

Accession For

NINE SEAS	☒
D. J. P.	☐
U. S. A.	☐

A

ACKNOWLEDGEMENTS

We would like to thank Dr. Frank P. Kuhl of the U. S. Army ARRADCOM for his interest, suggestions, and consistent support in the pursuit of knowledge concerning shape recognition. We also appreciate the interest and support of Mr. Donald Bush of the Rome Air Development Center for his interest in the shape recognition research here.

A special thanks is due to present and past students here at Purdue who have built a strong program in shape recognition research on which this research is based. These include Tim Wallace, Rahim Rostompour, John Gifford, Marcus Glenn, and Didier Charpentier.

We want to thank Prof. Anthony Reeves of Cornell University who is a colleague and an expert on the use of moments for his helpful reports and software. We wish to thank George Gobel and the entire Engineering Computer Network software and hardware staffs for the computational capabilities which allowed this research. Thanks to Sharon Katz for preparing many of the illustrations.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vii
LIST OF FIGURES.....	x
ABSTRACT	xviii
CHAPTER 1 INTRODUCTION.....	1
1.1 Shape: An Introduction	1
1.2 Introduction to Shape Analysis Methods.....	2
1.3 Categories of Shape Analysis Methods	4
CHAPTER 2 THE RECOGNITION SYSTEM.....	6
2.1 Recognition System Overview	6
2.2 Recognition System Simulation	8
2.3 Scheme of Experiments	14
2.3.1 Image Resolution Experiment	21
2.3.2 Feature Vector Experiment	21
2.3.3 Library Projection Experiment.....	28
2.3.4 Imaging Noise Experiment	29
2.3.5 Partial Shape Experiment	32
CHAPTER 3 FOURIER DESCRIPTORS OF THE BOUNDARY	49
3.1 Introduction	49
3.2 Properties	52
3.3 Calculation	69
3.4 Normalization	73
3.5 Generic Shapes	78
3.6 Experimental Results	88
3.7 Conclusions.....	96

CHAPTER 4 WALSH POINTS OF THE BOUNDARY	99
4.1 Introduction	99
4.2 Truncation and ϵ -Net	104
4.3 Normalization	107
4.4 Experimental Results	117
4.5 Conclusions	124
 CHAPTER 5 CUMULATIVE ANGULAR DEVIANT FOURIER DESCRIPTOR	 125
5.1 Introduction	125
5.2 Properties	126
5.3 Calculation	128
5.4 Normalization	131
5.5 Generic Shapes	132
5.6 Experimental Results	134
5.7 Conclusions	141
 CHAPTER 6 MOMENTS OF THE SILHOUETTE	 144
6.1 Introduction	144
6.2 Properties	145
6.3 Other Moment Sets	148
6.4 Normalization	152
6.5 Calculation	154
6.6 Reconstruction	157
6.7 Generic Shapes	162
6.8 Experimental Results	164
6.9 Conclusions	168
 CHAPTER 7 MOMENTS OF THE BOUNDARY CURVE	 173
7.1 Introduction	173
7.2 Normalization	173
7.3 Generic Shapes	174
7.4 Calculation	176
7.5 Experimental Results	176
7.6 Conclusions	180
 CHAPTER 8 COMPARISON OF THE GLOBAL SHAPE METHODS	 184
8.1 Introduction	184
8.2 Analytic Comparisons	184

8.3 Empirical Comparisons	189
8.4 Conclusions	200
CHAPTER 9 FOURIER-MELLIN TRANSFORM TECHNIQUE FOR PARTIAL SHAPE RECOGNITION	208
9.1 Introduction	208
9.2 Curvature Function	209
9.3 Mellin Transform	211
9.4 Fourier-Mellin Technique	212
9.5 Experimental Results	214
9.6 Conclusions	249
CHAPTER 10 CONCLUSIONS AND RECOMMENDATIONS	250
LIST OF REFERENCES	253
VITA	259

LIST OF TABLES

Table	Page
3.1 FDS image resolution experimental results: classification accuracy (%).....	92
3.2 FDS image resolution experimental results: median angle error $\Delta\phi_x, \Delta\phi_y$ (°).....	93
3.3 FDS feature vector experimental results.	94
3.4 FDS imaging noise experimental results.....	95
3.5 FDS library sampling experimental results.....	97
3.6 FDS partial shape experimental results.....	98
4.1 WAL image resolution experimental results: classification accuracy (%).....	118
4.2 WAL image resolution experimental results: median angle error $\Delta\phi_x, \Delta\phi_y$ (°).....	119
4.3 WAL feature vector experimental results.....	120
4.4 WAL imaging noise experimental results.	121
4.5 WAL library sampling experimental results.	122
4.6 WAL partial shape experimental results.	123
5.1 CAD image resolution experimental results: classification accuracy (%).....	137

5.2	CAD image resolution experimental results: median angle error $\Delta\phi_x, \Delta\phi_y$ (°).....	138
5.3	CAD feature vector experimental results.....	139
5.4	CAD imaging noise experimental results.	140
5.5	CAD library sampling experimental results.	142
5.6	CAD partial shape experimental results.	143
6.1	Table of the function $D_Y(a_{i-1}, a_i)$	155
6.2	Table of the function $C_Y(a_{i-1}, a_i)$	156
6.3	MOM image resolution experimental results: classification accuracy (%).....	165
6.4	MOM image resolution experimental results: median angle error $\Delta\phi_x, \Delta\phi_y$ (°).....	166
6.5	MOM feature vector experimental results.	167
6.6	MOM imaging noise experimental results.....	169
6.7	MOM library sampling experimental results.	170
6.8	MOM partial shape experimental results.....	171
7.1	MOMB image resolution experimental results: classification accuracy (%).....	177
7.2	MOMB image resolution experimental results: median angle error $\Delta\phi_x, \Delta\phi_y$ (°).....	178
7.3	MOMB feature vector experimental results.	179
7.4	MOMB imaging noise experimental results.	181
7.5	MOMB library sampling experimental results.	182
7.6	MOMB partial shape experimental results.	183

9.1 Fourier-Mellin partial shape method classification results for overhead views.	248
--	-----

LIST OF FIGURES

Figure	Page
2.1 Recognition system block diagram.	7
2.2 Coordinate system for imaging three-dimensional object.	10
2.3 Direction code for chain codes.	12
2.4 Mig aircraft shape library of 143 (13x11) views, 256x256 image resolution.	15
2.5 B57 aircraft shape library of 143 (13x11) views, 256x256 image resolution.	16
2.6 Phantom aircraft shape library of 143 (13x11) views, 256x256 image resolution.	17
2.7 F104 aircraft shape library of 143 (13x11) views, 256x256 image resolution.	18
2.8 F105 aircraft shape library of 143 (13x11) views, 256x256 image resolution.	19
2.9 Mirage aircraft shape library of 143 (13x11) views, 256x256 image resolution.	20
2.10 Samples from unknown shape test sets, 16x16 image resolution.	22
2.11 Samples from unknown shape test sets, 32x32 image resolution.	23

2.12 Samples from unknown shape test sets, 64x64 image resolution.	24
2.13 Samples from unknown shape test sets, 128x128 image resolution.	25
2.14 Samples from unknown shape test sets, 256x256 image resolution.	26
2.15 Samples from unknown shape test sets, 512x512 image resolution.	27
2.16 F104 library of 9 (3x3) views, 256x256 image resolution.	30
2.17 F104 library of 49 (7x7) views, 256x256 image resolution.	31
2.18 Noise free F104 silhouette, 128x128 image resolution.	33
2.19 Noise free F104 contour, 128x128 image resolution.	34
2.20 F104 image with noise added (SNR = 3 dB), 128x128 image resolution.	35
2.21 Noisy F104 contour (SNR = 3 dB), 128x128 image resolution.	36
2.22 F104 image with noise added (SNR = 6 dB), 128x128 image resolution.	37
2.23 Noisy F104 contour (SNR = 6 dB), 128x128 image resolution.	38
2.24 F104 image with noise added (SNR = 10 dB), 128x128 image resolution.	39
2.25 Noisy F104 contour (SNR = 10 dB), 128x128 image resolution.	40
2.26 F104 image with noise added (SNR = 20 dB), 128x128 image resolution.	41

2.27 Noisy F104 contour (SNR = 20 dB), 128x128 image resolution.	42
2.28 Sample unknown shape contours chopped 0 %.	43
2.29 Sample unknown shape contours chopped 10 %.	44
2.30 Sample unknown shape contours chopped 20 %.	45
2.31 Sample unknown shape contours chopped 30 %.	46
2.32 Sample unknown shape contours chopped 40 %.	47
2.33 Sample unknown shape contours chopped 50 %.	48
3.1 Boundary function in the complex plane.	50
3.2 Translation.	53
3.3 Rotation.	53
3.4 Dilation (scale).	55
3.5 Starting point shift.	55
3.6 N-fold symmetry ($N = 8$).	57
3.7 Axial symmetry.	57
3.8 Bilateral symmetry.	61
3.9 Reflection.	63
3.10 Line shape.	67
3.11 Polygon parameter definitions for Fourier descriptors of the boundary.	67
3.12 Fundamental shape library [MOEL82].	79
3.13 Fundamental shape library showing sample points.	80
3.14 Rectangle.	83

3.15	Isosceles triangle.....	83
3.16	Simple line shape.	87
3.17	Thin cross.	87
3.18	F104 contours reconstructed from the Fourier descriptors of the boundary.....	89
3.19	Fourier descriptors of the boundary (magnitude) for F104 contour.....	90
4.1	Sketch of an ϵ -neighborhood.....	106
4.2	Original F104 contour.	109
4.3	Walsh points with corresponding ϵ -neighborhood after normalization, $2^m = 16$ points.....	110
4.4	Walsh points with corresponding ϵ -neighborhood after normalization, $2^m = 32$ points.....	111
4.5	Walsh points with corresponding ϵ -neighborhood after normalization, $2^m = 64$ points.....	112
4.6	Walsh points with corresponding ϵ -neighborhood after normalization, $2^m = 128$ points.....	113
4.7	Original F104 contour projections (624 points).	114
4.8	Normalized Walsh points projections ($2^m = 16$).....	114
4.9	Normalized Walsh points projections ($2^m = 32$).....	115
4.10	Normalized Walsh points projections ($2^m = 64$).....	115
4.11	Normalized Walsh points projections ($2^m = 128$).	116
5.1	Polygon parameter definition for cumulative angular deviant Fourier descriptors.....	130
5.2	Cumulative angular deviant Fourier descriptors magnitude) for F104 contour.....	135

5.3	F104 contours reconstructed from the cumulative angular deviant Fourier descriptors.....	136
6.1	Letter 'E' silhouette.....	160
6.2	The letter 'E' reconstructed from the 15 th order moments using the discrete inverse method.	161
8.1	Combined results for the image resolution experiment: classification accuracy.....	191
8.2	Combined results for the image resolution experiment: median x angle error.	192
8.3	Combined results for the image resolution experiment: median y angle error.....	193
8.4	Combined results for the feature vector experiment: classification accuracy.....	194
8.5	Combined results for the feature vector experiment: median x angle error.	195
8.6	Combined results for the feature vector experiment: median y angle error.....	196
8.7	Combined results for the library projection experiment: classification accuracy.....	197
8.8	Combined results for the library projection experiment: median x angle error.	198
8.9	Combined results for the library projection experiment: median y angle error.....	199
8.10	Combined results for the imaging noise experiment: classification accuracy.....	201
8.11	Combined results for the imaging noise experiment: median x angle error.	202
8.12	Combined results for the imaging noise experiment: median y angle error.....	203

8.13 Combined results for the partial shape experiment: classification accuracy.....	204
8.14 Combined results for the partial shape experiment: median x angle error.	205
8.15 Combined results for the partial shape experiment: median y angle error.....	206
9.1 Complete F104 contour prototype (top) and 10% chopped unknown (bottom) contours.	215
9.2 Original F104 prototype curvature function, no filtering.	217
9.3 F104 prototype curvature function, 30 sample filter.	218
9.4 F104 prototype curvature function filtered and resampled to 256 samples.	219
9.5 10% chopped F104 shifted curvature function filtered and resampled).	220
9.6 Correlation function for prototype and unknown curvature functions.....	221
9.7 Prototype (solid) and shifted unknown (dotted) curvature functions.....	222
9.8 Difference function for prototype and shifted unknown curvature functions.....	223
9.9 Magnitude of the FFT for the prototype curvature function.	224
9.10 Magnitude of the FFT for the unknown curvature function.	225
9.11 Mellin transform (DMT) magnitude of the FFT magnitude, prototype (solid) and unknown (dotted).....	226
9.12 Fourier-Mellin cross correlation function.	227

9.13 Correlation function for prototype and scaled unknown curvature functions.....	228
9.14 Prototype (solid) and scaled & shifted unknown (dotted) curvature functions.....	230
9.15 Difference function for prototype and scaled & shifted unknown curvature functions.....	231
9.16 Weighted difference function.....	232
9.17 Weighted difference function (dotted) with segmentation (solid).....	233
9.18 Side view of F104 chopped 0, 10, 20, 30, 40, 50, 60, 70, 80, and 90%.....	234
9.19 Contours reconstructed from filtered & resampled curvature functions (0 - 90 % chopped).	235
9.20 Segmented reconstructed contour, 10% chopped.....	236
9.21 Segmented reconstructed contour, 20% chopped.....	236
9.22 Segmented reconstructed contour, 30% chopped.....	237
9.23 Segmented reconstructed contour, 40% chopped.....	237
9.24 Segmented reconstructed contour, 50% chopped.....	238
9.25 Segmented reconstructed contour, 60% chopped.....	238
9.26 Segmented reconstructed contour, 70% chopped.....	239
9.27 Segmented reconstructed contour, 80% chopped.....	239
9.28 Segmented reconstructed contour, 90% chopped.....	240
9.29 Overhead view of six aircraft.....	241
9.30 MIG overhead view chopped 0% to 90%.....	242
9.31 B57 overhead view chopped 0% to 90%.....	243

9.32 Phantom overhead view chopped 0% to 90%.....	244
9.33 F104 overhead view chopped 0% to 90%.....	245
9.34 F105 overhead view chopped 0% to 90%.....	246
9.35 Mirage overhead view chopped 0% to 90%.....	247

ABSTRACT

Grogan, Timothy Alan. Ph.D., Purdue University, August 1983. SHAPE RECOGNITION AND DESCRIPTION: A COMPARATIVE STUDY. Major Professor: O. Robert Mitchell.

An important problem in the extraction of information from images is shape recognition. Several methods of analyzing binary images using global shape methods based upon functional approximation have been reported in the literature. However, there has been a lack of information comparing the effectiveness of these methods in shape analysis.

Five methods of global shape analysis are compared on two basis. The five methods compared are 1) Fourier descriptors of the boundary, 2) Walsh points of the boundary, 3) the cumulative angular deviant Fourier descriptors, 4) moments of the silhouette, and 5) moments of the boundary. First, the different methods are introduced, their geometric properties presented, and the formulae for some generic shapes are provided. Then the methods are compared on the basis of the empirical facts derived from a set of aircraft shape recognition experiments. The shapes are different views of six aircraft. The aircraft silhouettes are the two-dimensional projected images of three-dimensional rigid bodies. The five methods are ranked according to their performance from these experiments.

A new method for the recognition of partial shapes based on the Fourier-Mellin transform is introduced. A shift and scale invariant correlation of the complete and partial shape's curvature functions is obtained by applying the

Mellin transform to the magnitude of the Fourier transform of the curvature functions. The logarithm of the shift in the correlation function corresponds to the time scaling of the partial shape's curvature function. Then the shift in an ordinary correlation of the complete shape's curvature and the scaled partial shape's curvature is the time shift necessary to complete the alignment. Then a pointwise comparison of the curvature functions can be made to determine matching and non-matching contour segments. Some initial recognition experiments for partial shapes are carried out and the results reported.

CHAPTER 1

INTRODUCTION

1.1 Shape: An Introduction

The function of many information processing systems is to either extract information from a signal or to generate a signal from the realization of algorithms which contain encoded information. This shape information can then be used to describe a known shape or to recognize (detect or classify) an unknown shape. This report compares several methods of representing shape information. In particular, the shapes dealt with herein will be two-dimensional regions in the plane that are simply connected. The representation of a three-dimensional object (a simply connected region) will also be discussed by way of its projection onto the plane.

The objective then is to characterize the shape by extracting measured quantities. This might appear to be very straight forward, but shape is an allusive quality to quantify. Shape can be contrasted with texture. Texture is also difficult to define. Texture is the random or chaotic non-form of a signal and shape is the structure or form of the signal. These definitions however have a fuzzy boundary between them.

1.2 Introduction to Shape Analysis Methods

In the study of the architecture of neurons in the brain and sensory nerves, it has been suggested [PITT47] that humans perceive shape by performing comparisons to forms already stored in the brain. This comparison is performed invariant of various transformations on the shape such as scaling of size, translation, and motion. The structure of the neurons implement operations that are invariant over these groups of transformations [HOFF78]. Others have suggested that a shape is perceived when neural signals are transformed until they reach a stable signal or the neurons reach a stable condition (move to a strange attractor[THOM75]). It is at that moment that the shape is recognized. To actually implement this kind of algorithm on a computing machine will require massive parallelism. Some preliminary investigations into parallel implementation of shape algorithms have been shown to have increased effectiveness [TUOM83].

Some of the early methods in shape analysis were based on extracting gross parameters of the shape. For instance, topological invariant properties such as the number of connected components, the number of holes, Euler number, etc. were tabulated for a binary image [ROSE70]. Methods first used in biology, for instance, the statistics of random cord intersections and tangents, have their theoretical basis in integral geometry and the theory of random sets [SANT76]. Also, aspect ratio, area, and perimeter have been used [PAVL78]. Although robust in nature, these methods lack the ability to characterize fine detail in and among shapes. The most recent methods have attempted to provide this necessary discriminating ability.

Others investigators [FU82] have used formal languages with sets of rules (syntax) to describe a structure of a signal. Syntactic methods represent local

characteristics of a shape well. However, this approach begins having difficulty in generating the necessary rules in an environment containing many shapes or where one of many views of a three-dimensional shape are possible. Also, these methods still rely on some other operation to first extract an ordered list of shape primitives. These shape primitives or lists of primitives are the words and sentences processed to determine if an allowed sentence (a possible shape) has the proper syntax (i.e. has a desired shape.)

Some stochastic models have been developed for shape analysis [KASH81]. These methods seem to be especially useful for modeling broad classes of shapes.

A very different approach is to instead measure how much of the unknown signal is contained in each of several basis signals. A list of numbers is used to characterize the shape of the signal(region). If shape recognition (detection or classification) is desired, then instead of matching them directly, a characterization is used that is convenient for the operations needed in lining them up properly prior to the matching process. Then it can be determined if they match by simply measuring the distance between the two characterizations. The characterization is obtained by projecting the signals onto the basis signals. This is done for each basis signal and the list of values obtained constitutes the characterization of the shape. It is desirable if the basis signals chosen are sufficiently different in order to extract the greatest possible distinctive or independent shape quantities.

The theory of functional approximation and linear vector spaces [KOLM75] provides exactly the formalism necessary to accomplish such a task. The methods of shape analysis compared in this report differ mainly in what sort of basis signal are used to characterize a shape. The measurement of

quantities used for signal recognition (detection or classification) and description is termed feature extraction.

1.3 Categories of Shape Analysis Methods

Shape analysis methods can be assigned to one of several categories. One possible distinction can be made on the basis of whether local or global context is used in extracting the shape information. With a *local* shape analysis method only a subset of the elements in the characterization is affected by information in a local neighborhood on the shape. However, with a *global* method almost every element in the characterization is affected by information in a local neighborhood.

Another partition can be made based on how much of the shape must be available for the method to still perform adequately. *Complete* shape analysis methods need the entire shape present to perform well. *Partial* shape methods are designed to perform well even when some of the shape is either missing or severely corrupted by noise. All but one of the methods discussed and compared in this report are global methods. However, the performance of these methods using partial shapes is also included. In Chapter 8, a new method for partial shape recognition based on the Fourier-Mellin transform is presented[GROG83].

Shape methods can also be categorized on the basis of whether the interior of the object or only the exterior boundary of the object is used in forming the characterization[PAVL78]. An *external scalar* shape method analyzes the essentially one-dimensional function describing the boundary of the object. *Internal spatial* analysis methods operate on the spatial distribution of the object's interior elements.

The main goal of this report is to compare five methods of complete shape analysis. Each method will be discussed in detail in a separate chapter. The methods will be compared by first providing their defining relationships and listing many of the properties that pertain to shape recognition and description. Also provided are generic shapes with their corresponding representations. Lastly, empirical results from several large recognition experiments are presented. The combined results from all five of the shape methods are then discussed in a following chapter. In the chapter immediately following is a discussion of the recognition system used and experiments carried out to obtain the empirical recognition results.

CHAPTER 2

THE RECOGNITION SYSTEM

2.1 Recognition System Overview

Information processing systems are often designed to sense the environment. Signals are collected, information contained in patterns of the signal are extracted, and then an appropriate response is issued based on that information. A general system for pattern recognition is depicted in Figure 2.1. Let's briefly discuss the operations that take place as patterns are processed by such a system.

Information about the environment is acquired by collecting energy. This collection is accomplished by a device (*collector*) such as an antenna, a lens, microphone, etc., and is ordinarily distributed in space and/or time. A transducer converts the gathered energy into a form suitable for processing by the recognition system. The collector and transducer together are often called the *sensor*. The pattern is then presented to the *preprocessing* functional block. This block usually contains filters which either enhance or restore the signal. Some noise reduction is often attempted. The *segmentation* block separates or labels portions of the incoming signal. This block is often a simple thresholding operation, but can be much more sophisticated such as a pixel classifier or labeling of the signal components as a result of a feature clustering algorithm. The basic idea is to discriminate among portions of the signal. The segmentation functional block is often the most difficult block to design. This

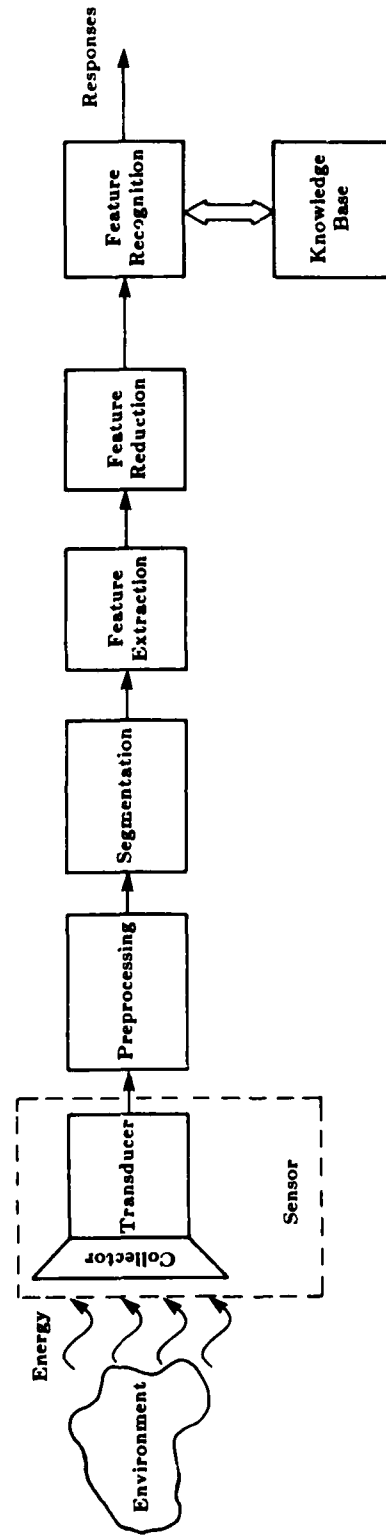


Figure 2.1 Recognition system block diagram.

functional component is very much like the pattern recognizer to follow. A tradeoff between the complexity of the segmentation and the remaining portion of the pattern recognition system is often the main source of engineering difficulties in implementing a system which can process signals from a real world environment. Next is *feature extraction*. This is the functional component where the shape features are extracted when using a shape analysis method. Comparing the different shape methods mainly means changing this functional component with another employing a different shape method. This component can also be the functional unit that provides the primitives needed when using a syntactic method. Often following feature extraction is *feature reduction*. This is employed to eliminate redundancy in the extracted features or to simply prune down the number of features to a manageable size. The methods often used are linear transformations including the principal components transformation. The *feature recognition* functional element following can be simple template matching using an error norm followed by a ranking of the errors (K weighted nearest neighbors), a syntactic analyzer(parser), or another decision function based on a probabilistic model. The feature classifier often calls upon the *knowledge base* for information such as templates, productions rules(grammars), statistical parameters, or a model with its associated parameters.

2.2 Recognition System Simulation

Now the simulation of the recognition system used to obtain the results contained in this report will be discussed. Shapes for investigating the recognition system's performance are simulated rather than using actual sensor data in order to obtain more control over the experiments. Silhouettes of

three-dimensional objects are obtained using a computer graphics program. Models of military aircraft are formed by the union of cylindric and rhombic prisms whose size and orientation have been specified. The faces of these prisms are specified by the list of their vertices. The vertices are first rotated in order to obtain the silhouette of the object corresponding to viewing from a particular aspect angle, (ϕ_x, ϕ_y, ϕ_z) . (See Figure 2.2.) Then the object, $O(x,y,z)$, is parallel projected (P operator) onto the (x', y') image plane. So, corresponding to a particular aspect angle (ϕ_x, ϕ_y, ϕ_z) is the image $I(x', y') = P\{O(x, y, z) : (\phi_x, \phi_y, \phi_z)\}$. To simulate the projection operation an array of individual parallel rays are "fired" at the model. Where a ray intersects the object an image picture element (pixel) is set to one and where there is no intersection the pixel is assigned a value of zero. So, in this case we have a discrete image,

$$I_{ij} = I(i\Delta\theta, j\Delta\theta) = P\{O(x, y, z) : (\phi_x, \phi_y, \phi_z), \Delta\theta\},$$

where $\Delta\theta$ is the spacing between rays.

No attempt is made to perform any noise removal or any other preprocessing. In other words, the preprocessing component is not simulated.

In one set of experiments, noise is added to the two level image in order to simulate a real world situation. The discrete noise image, n_{ij} , is added to the discrete projected silhouette image to obtain $\tilde{I}_{ij} = I_{ij} + n_{ij}$. This corrupted image is then thresholded, i.e.

$$\tilde{I}_{ij}^t = \begin{cases} 1, & \text{if } \tilde{I}_{ij} \geq t \\ 0, & \text{if } \tilde{I}_{ij} < t \end{cases}$$

where t = threshold.

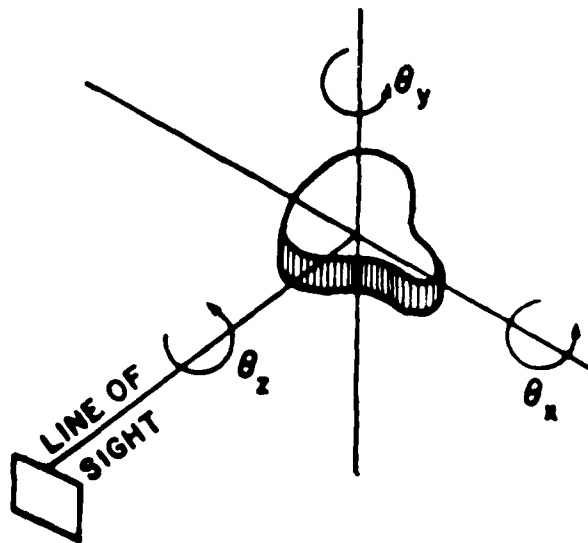


Figure 2.2 Coordinate system for imaging three-dimensional object.

In order to reduce storage requirements, instead of retaining the object silhouette image, only the boundary is stored. The boundary information is stored in the form of an eight direction chain code. The chain code is a list of eight valued numbers that describe the direction to take to get from the present position to the next as the object is traced counterclockwise. The chain code is the ordered list, $V: a_0 a_1 \dots a_{K-1}$, where a_i = the direction code specifying the direction to the next image point along the contour, and K = length of the chain. The directions of the chain code are shown in Figure 2.3.

Each different shape method generates a feature vector $\underline{f} = [f_0 \dots f_n]^T$ which is a characterization of either the boundary function or the image itself. Let $f(\underline{x})$ be the function to be approximated, then $f(\underline{x}) = \sum_i f_i \phi_i(\underline{x})$, and $f_i = \int f(\underline{x}) \theta_i(\underline{x}) d\underline{x}$, where $\phi_i(\underline{x})$ are the basis functions of the method and $\{\theta_i(\underline{x})\}$ is the reciprocal basis for $\{\phi_i(\underline{x})\}$. For later processing, the set $\{f_i\}$ is usually truncated to a finite set of elements K , and reordered to form the feature vector $\underline{f} = [f_0 \dots f_{K-1}]^T$. The coefficients, f_i , are used in forming the characterization.

Further reduction of the feature vector is performed by applying a principle components transformation followed by a truncation of the elements in the feature vector. The principle components transformation is derived by combining the feature vectors from all the libraries and calculating the sample covariance matrix. That is

$$\Sigma_{ij} = \frac{1}{N'} \sum_{\ell=1}^{N'} (f_i^\ell - \bar{f}_i) (f_j^\ell - \bar{f}_j),$$

where

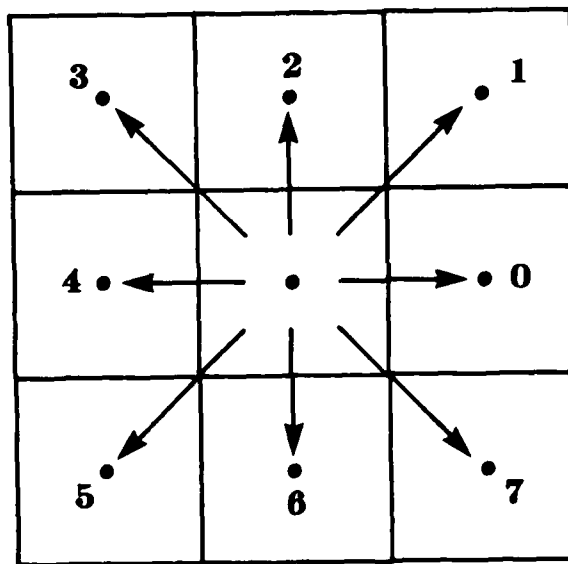


Figure 2.3 Direction code for chain codes.

$\Sigma_{ij} = (i, j)^{\text{th}}$ element of the covariance matrix

$$\bar{f}_i = \frac{1}{N'} \sum_{\theta=1}^{N'} f_{i\theta} = \text{mean value of the } i^{\text{th}} \text{ feature.}$$

$N' =$ number of feature vectors in the combined libraries.

$f_{i\theta} = i^{\text{th}}$ feature of the θ^{th} feature vector in the combined library.

Σ is a symmetric matrix. The principle components transformation then is the matrix of eigenvectors of the covariance ordered by eigenvalue from the largest to the smallest, i.e. $[P]^T[\Sigma][P] = \Lambda$, where Λ is a diagonal matrix with the eigenvalues along the diagonal and P is the matrix of eigenvectors. Letting P' be the matrix P with its rows reordered by eigenvalue (largest to smallest) and zero rows for the j^{th} rows, $j > K'$, then the reduced feature vector is $\hat{f} = P' f$. In general, the reduced feature vectors in the experiments consist of twelve (12) elements.

Each individual airplane shape feature library is transformed into a reduced feature vector library. These six libraries for the six different military aircraft make up the knowledge base or templates used by the classifier.

Feature vectors from unknown shapes are then classified using a nearest neighbor template matching. The unknown feature is compared to every feature vector in the shape libraries. The unknown shape is then labeled as belonging to the aircraft containing the library feature vector, $\hat{f}^{L_j}$, having the smallest distance to the unknown feature vector, $\hat{f}^u$. The distance is measured as the sum of the square of the difference between each feature vector element in the feature vector, i.e.

$$d_{L_j} = \sum_{i=1}^{K'} [\hat{f}_i^u - \hat{f}_{ij}^{L_j}]^2,$$

where $\hat{f}_{ij}^{L_j} =$ denotes the i^{th} element in the j^{th} feature vector of the L^{th} library.

2.3 Scheme of Experiments

Five basic classification experiments were designed to test the performance of the shape analysis methods under varying conditions. Following is a description of each experiment. Each of these experiments was carried out for each for the five complete shape analysis methods. The results of these experiments are reported in each of the chapters describing that particular method. In chapter 8 is a comparison of the results of these experiments for the differing shape methods.

The shapes used for data in the experiments are two-dimensional projections of three-dimensional models of six different military aircraft. They are the Mig, B57, Phantom, F104, F105, and the Mirage. The shape libraries are the sets of feature vectors derived from multiple views of each of these six aircraft. Figures 2.4 to 2.9 show each of the six military aircraft shape libraries consisting of 143 (13x11) views at an image resolution of 256x256 pixels. The viewing angles (in degrees) used to generate the libraries are pairs of ϕ_x and ϕ_y where

ϕ_x :	2.0	4.6	8.0	17.2	31.5	51.6	90.0
	128.5	148.4	162.7	171.9	175.3	177.6	

and

ϕ_y :	-90.0	-72.2	-53.9	-36.1	-17.8	0.0
	17.8	36.1	53.9	72.2	90.	

Fifty (50) random views not already contained in the libraries are generated for each of the six aircraft (6x50=300 unknowns). These unknown shapes are used as test patterns to be classified.

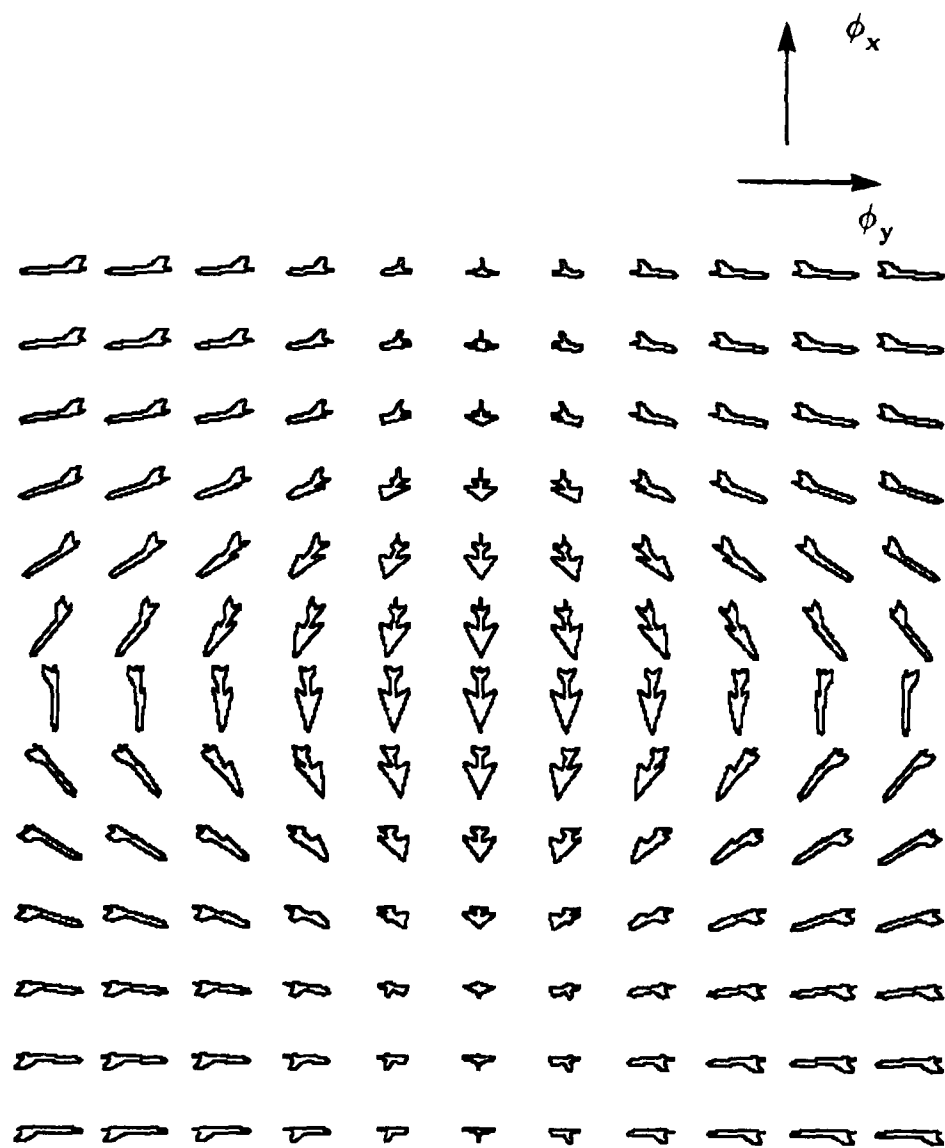


Figure 2.4 Mig aircraft shape library of 143 (13x11) views, 256x256 image resolution.

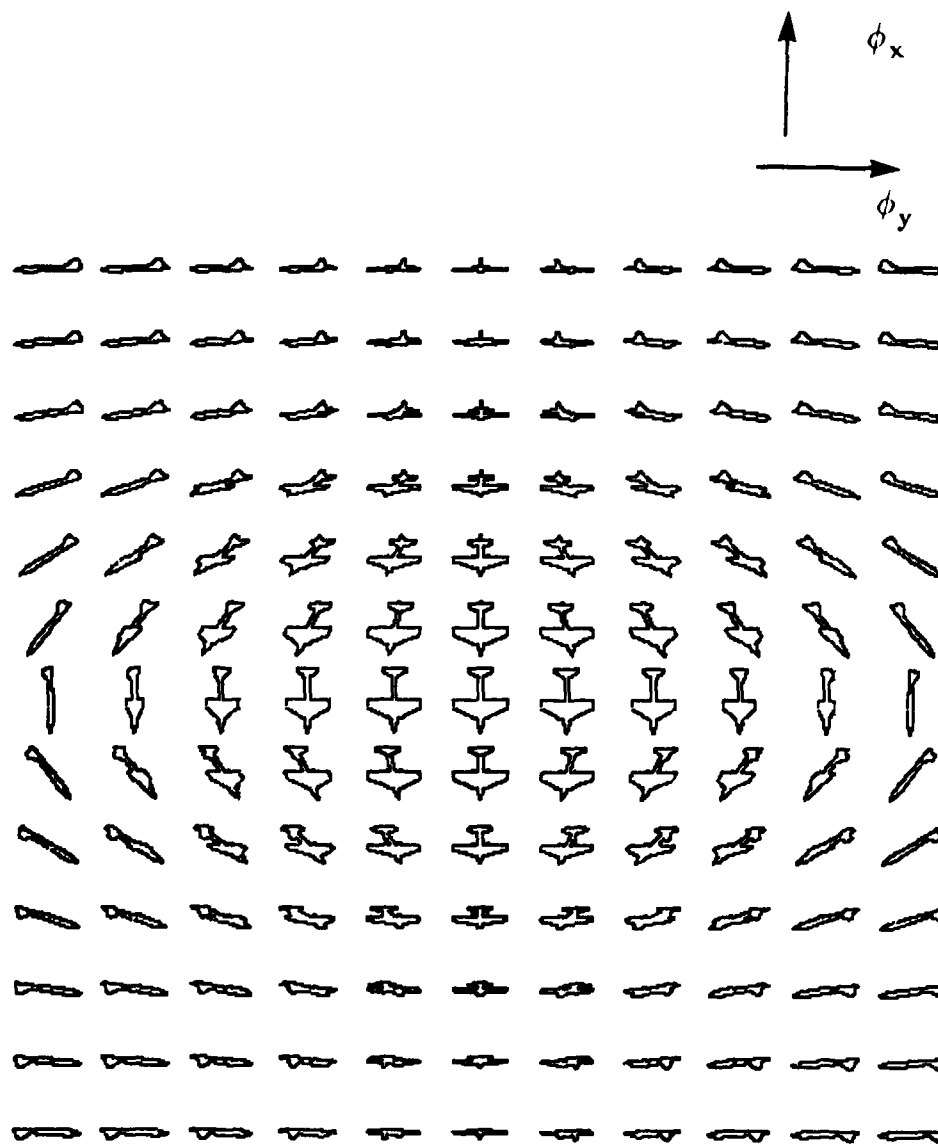


Figure 2.5 B57 aircraft shape library of 143 (13x11) views, 256x256 image resolution.

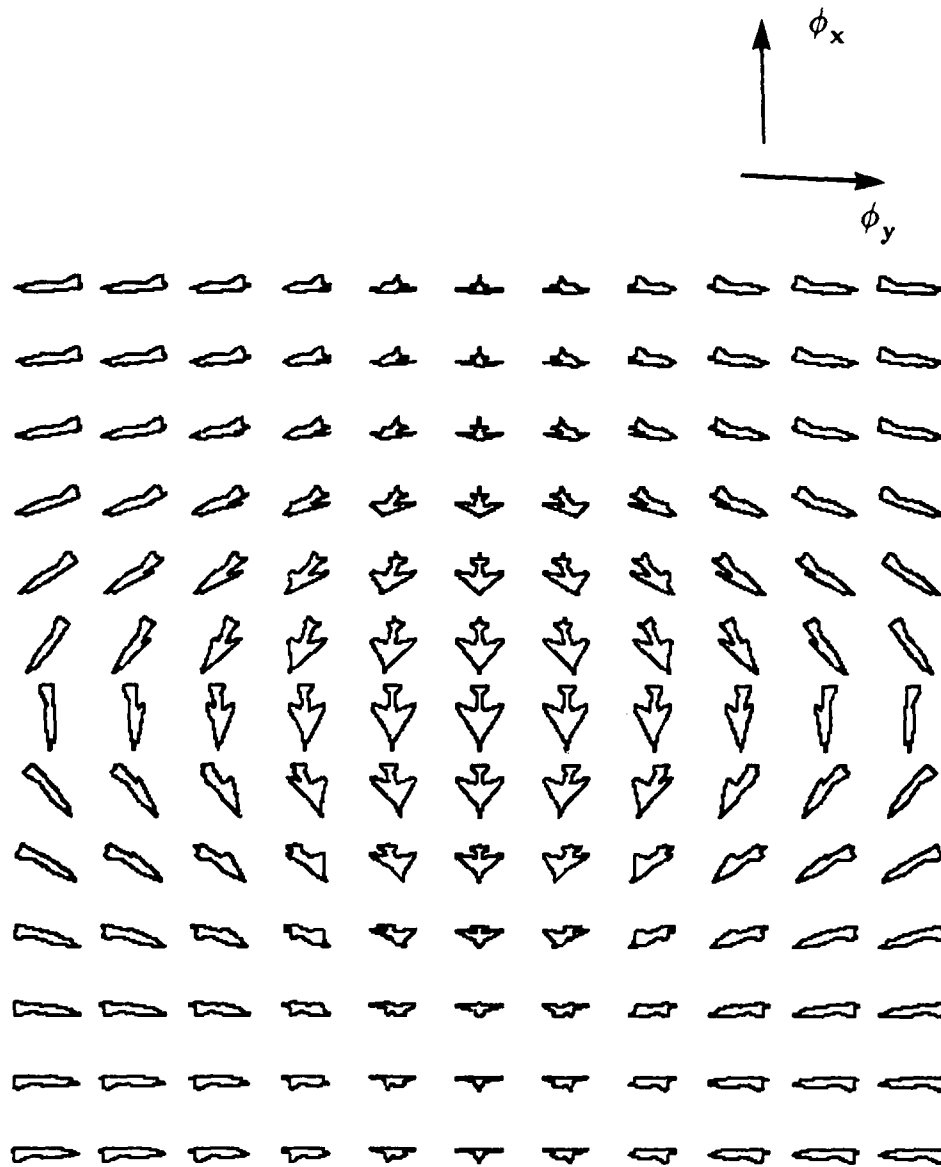


Figure 2.6 Phantom aircraft shape library of 143 (13x11) views, 256x256 image resolution.

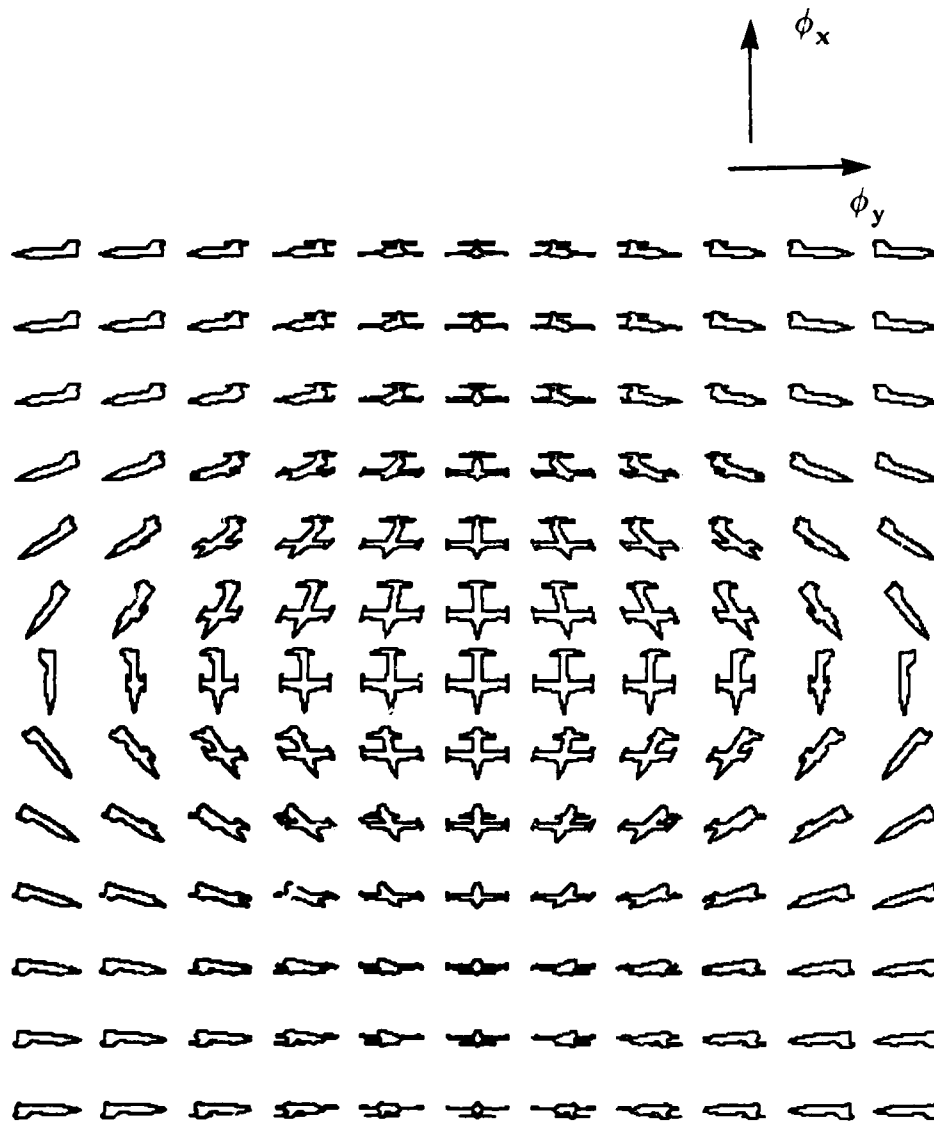


Figure 2.7 F104 aircraft shape library of 143 (13x11) views, 256x256 image resolution.

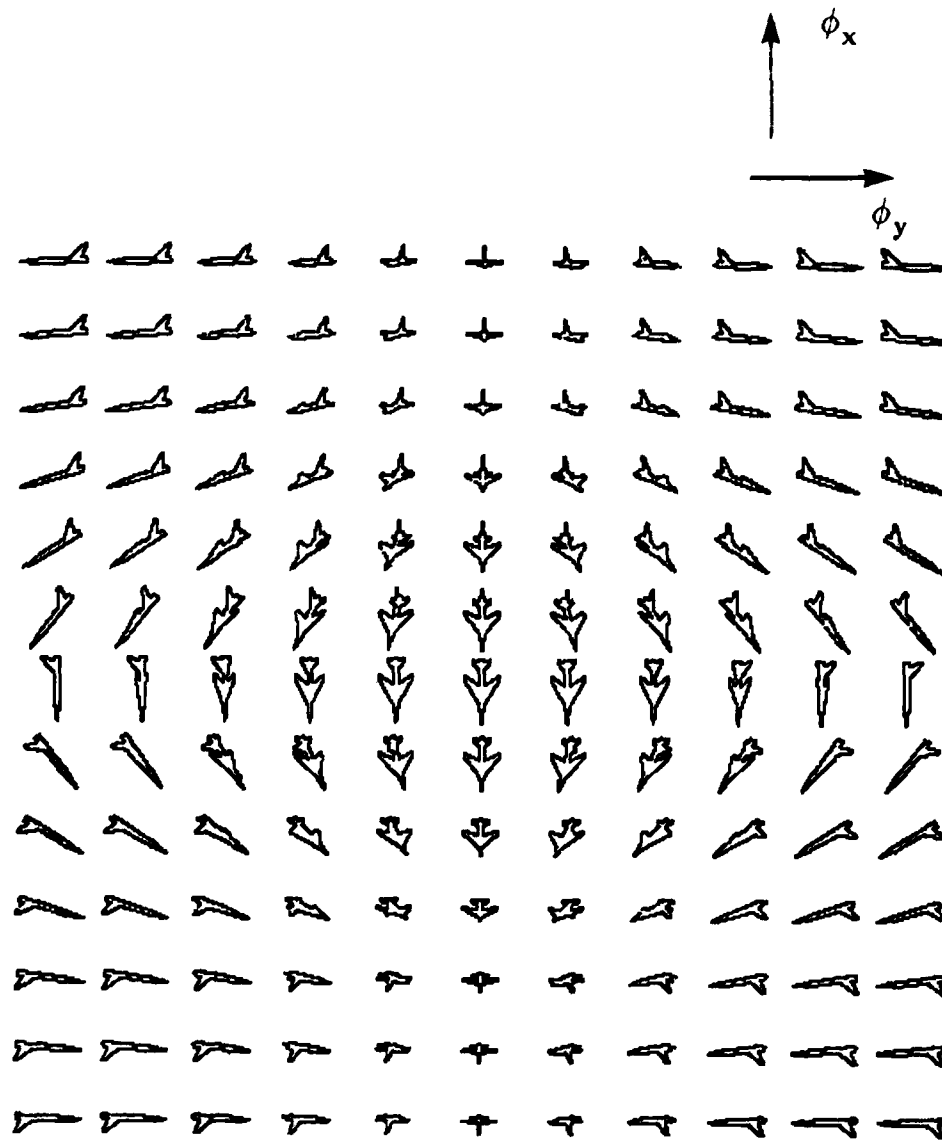


Figure 2.8 F105 aircraft shape library of 143 (13x11) views, 256x256 image resolution.

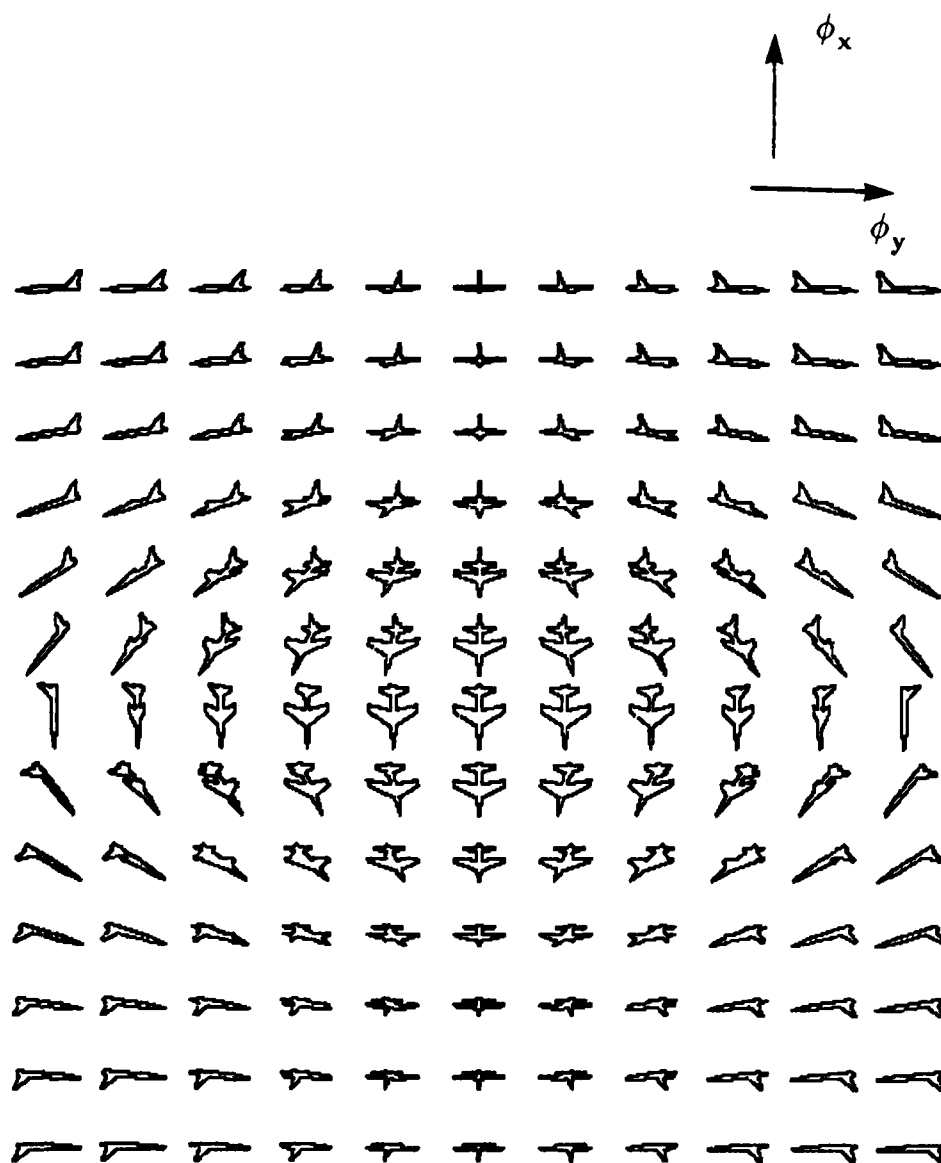


Figure 2.9 Mirage aircraft shape library of 143 (13x11) views, 256x256 image resolution.

2.3.1 Image Resolution Experiment

This experiment was designed to test the shape methods as noise was introduced as a result of varying image spatial resolution. The spacing of the rays (Δl) varies the image quality and thereby the quality of the projected shape. Both libraries and the unknowns were generated with pixel arrays resulting in images with 16x16 to 512x512 pixels per view. The same number of elements in the feature vector was used throughout this experiment. Each aircraft library consisted of 143 views corresponding to an 13x11 array of aspect angles covering a hemisphere. Figures 2.10 to 2.15 show samples from the unknown shape test sets at each of the six different image resolutions.

2.3.2 Feature Vector Experiment

This experiment was designed to provide insight into how compactly the shape information was represented by the feature vector. With a fixed number of views in each of the aircraft libraries(143) and with the libraries generated at a fixed image resolution (256x256), the feature vector length was gradually increased for both the libraries and the unknowns. The unknown image resolution was also varied to provide additional information on how well the feature vectors characterize the shape. The experiment was carried out using the full feature set (in its varying lengths) and also when the feature extraction is followed by the principle components feature reduction resulting in a reduced length feature vector (12 real numbers.)

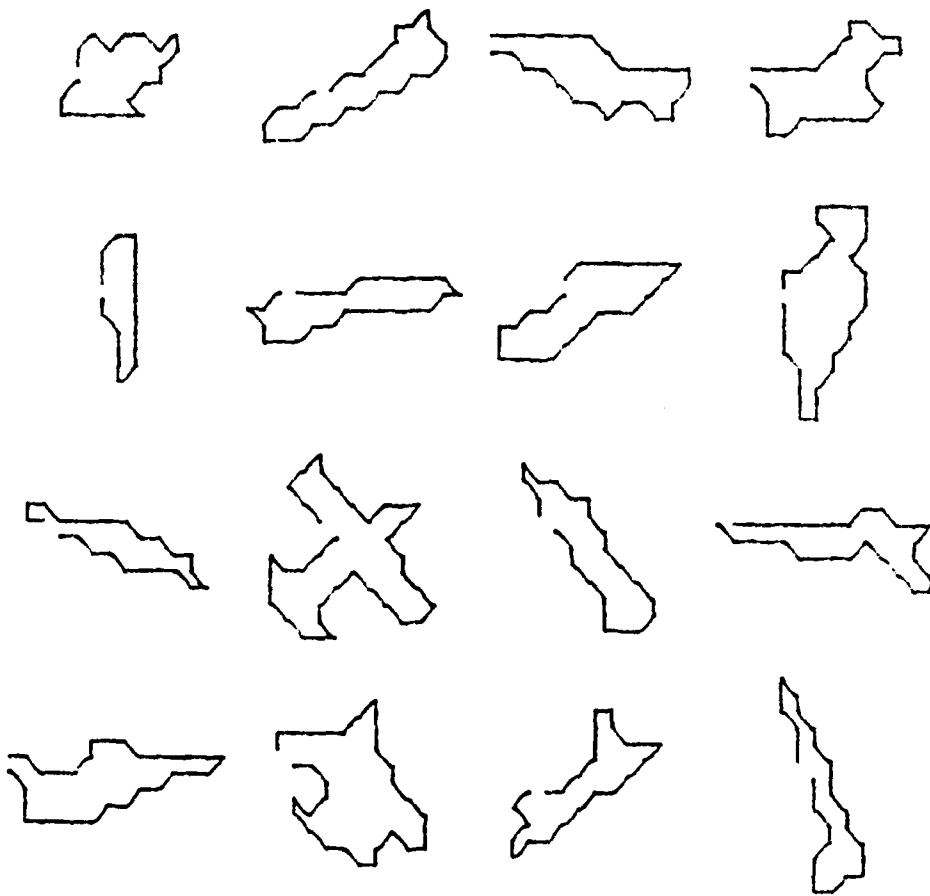


Figure 2.10 Samples from unknown shape test sets, 16x16 image resolution.

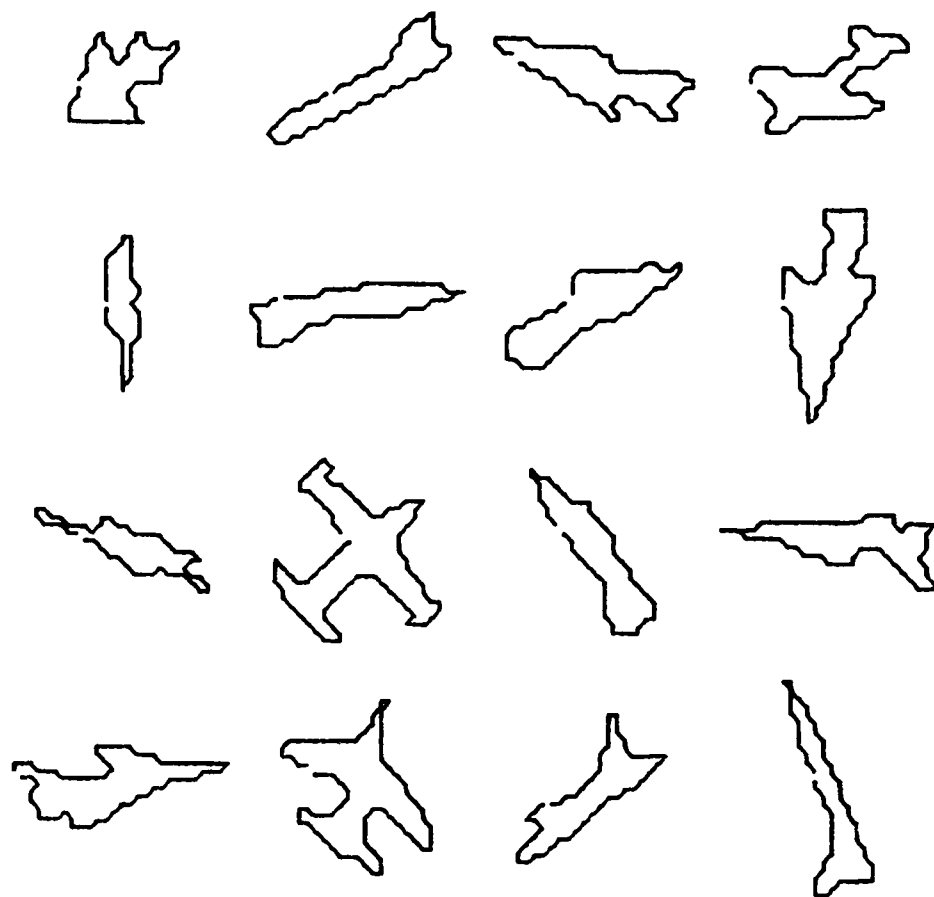


Figure 2.11 Samples from unknown shape test sets, 32x32 image resolution.

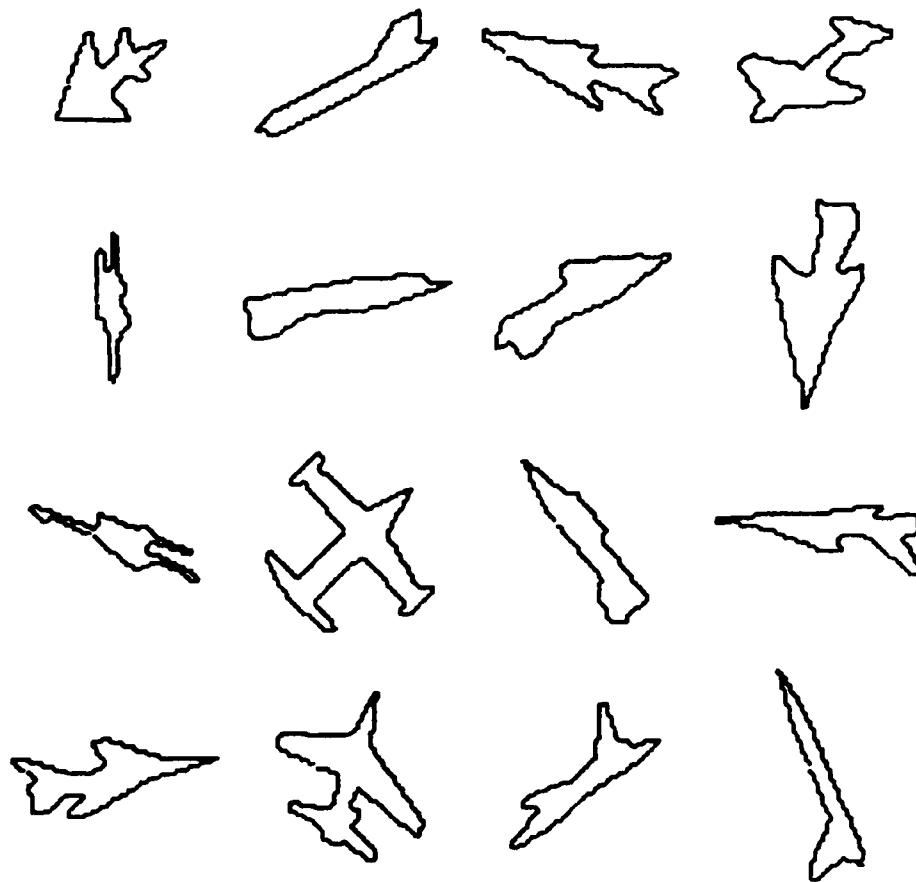


Figure 2.12 Samples from unknown shape test sets, 64x64 image resolution.

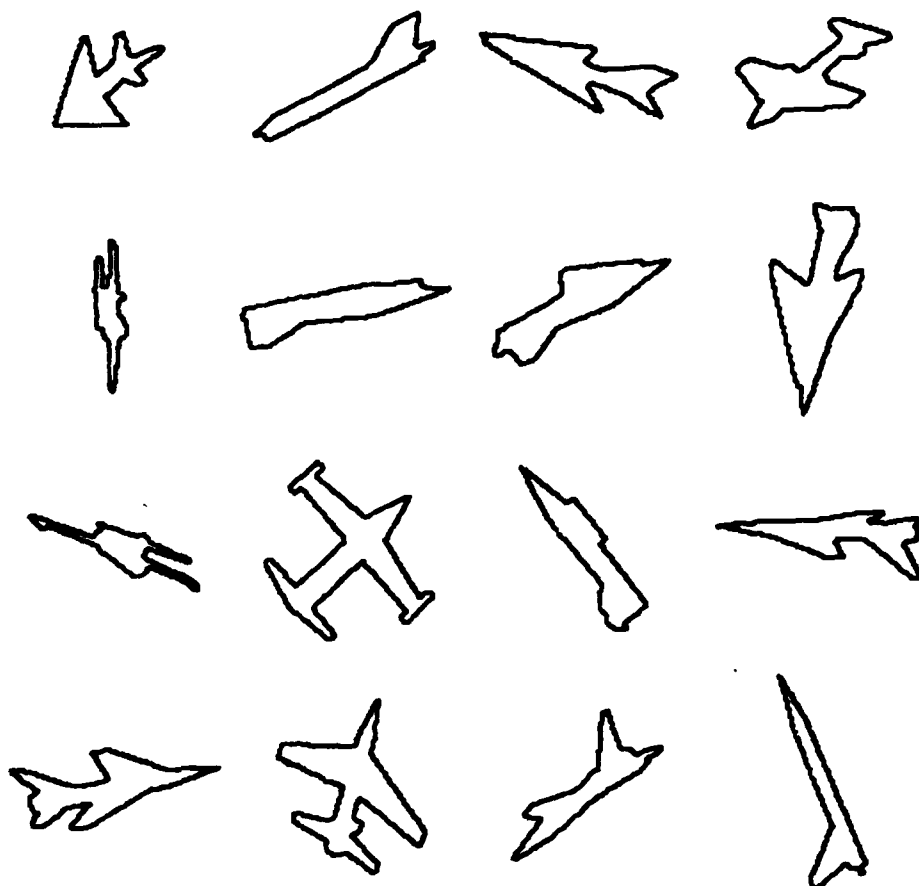


Figure 2.13 Samples from unknown shape test sets, 128x128 image resolution.

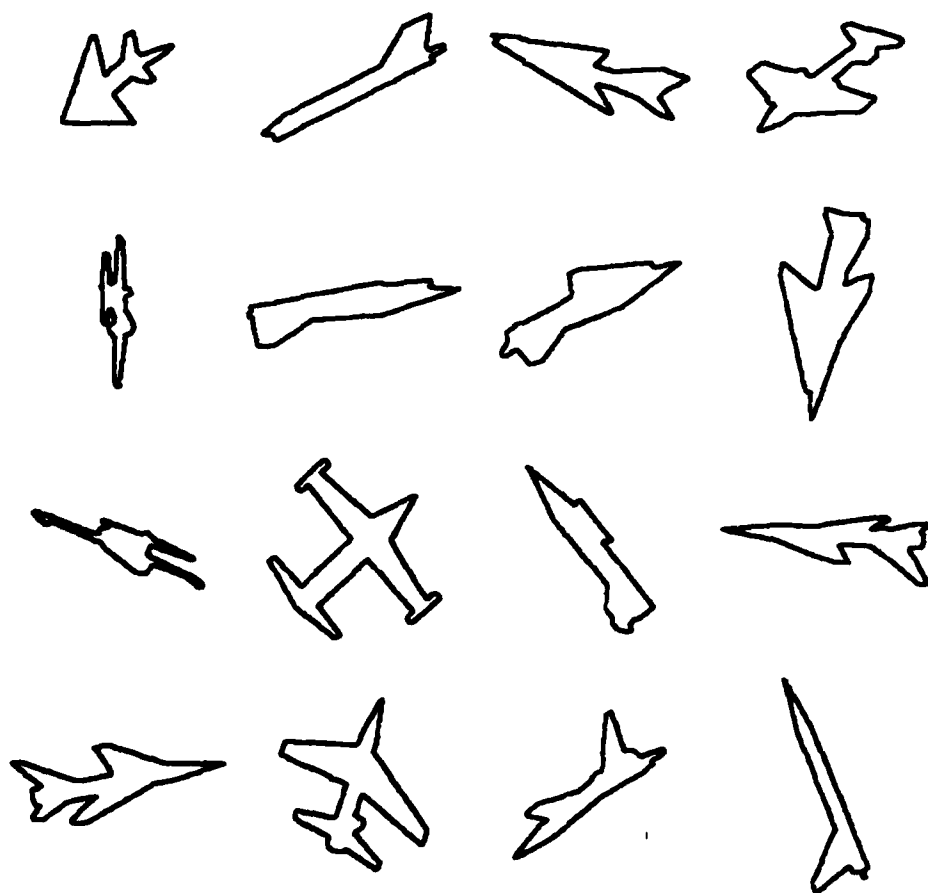


Figure 2.14 Samples from unknown shape test sets, 256x256 image resolution.

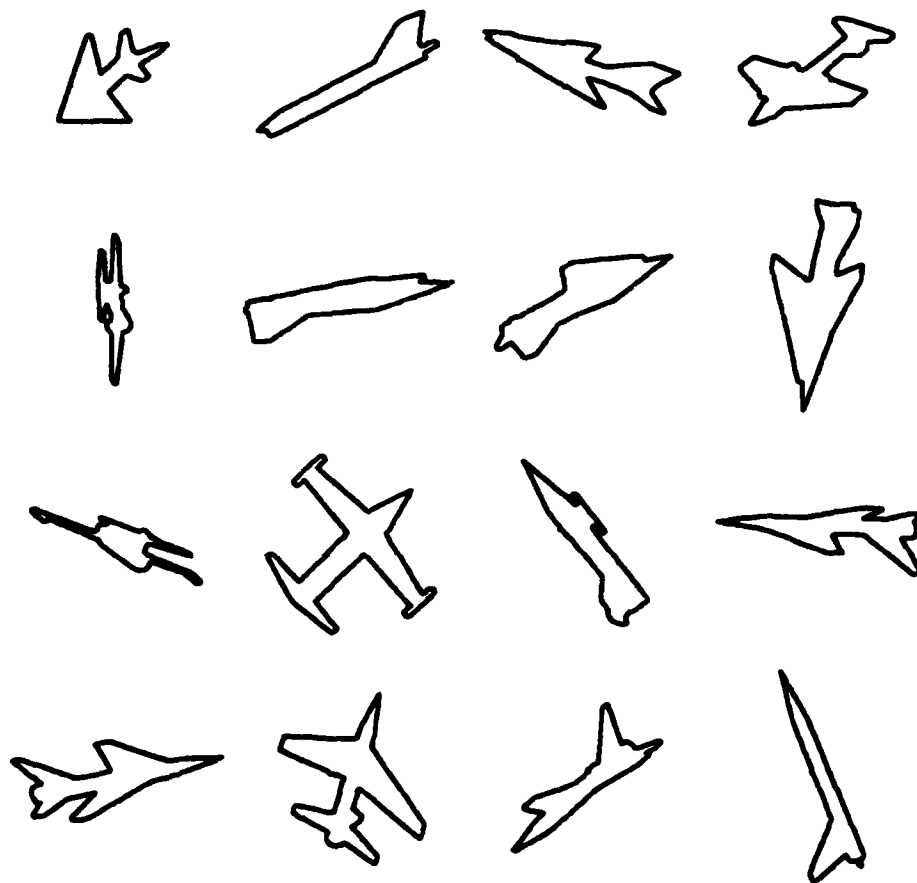


Figure 2.15 Samples from unknown shape test sets, 512x512 image resolution.

2.3.3 Library Projection Experiment

This experiment was designed to show variations in classification accuracy as the number of views in the shape libraries was decreased. This tested the different methods ability to properly classify a three-dimensional object from a limited number of views. It also tested how well the method could discriminate between very similar shapes that belong in different aircraft libraries. This occurs when there are many views from each aircraft. These are opposing effects however. So, to obtain more definitive results about these two effects would require an independent method of measuring the similarity among shapes. This however was not available. Others have investigated methods of reducing the library storage requirements [CHAR81, GLEN82]. Results from these investigations imply that only a small number of views are necessary to represent the shape adequately for the purposes of shape recognition. Figures 2.16 and 2.17 show the F104 libraries at an image resolution of 256x256 with 9 (3x3) and 49 (7x7) views per library, respectively. The viewing angles (in degrees) used to generate the 9 view (3x3) libraries are pairs of ϕ_x and ϕ_y where

$$\phi_x: \quad 8.0 \quad 90.0 \quad 171.9$$

and

$$\phi_y: \quad -72.2 \quad 0.0 \quad 72.2$$

The viewing angles (in degrees) used to generate the 49 view (7x7) libraries are

pairs of ϕ_x and ϕ_y where

$$\phi_x: \quad 2.0 \quad 8.0 \quad 31.5 \quad 90.0 \quad 148.4 \quad 171.9$$

and

$$\phi_y: \quad -90.0 \quad -53.9 \quad -17.8 \quad 0.0 \quad 17.8 \quad 53.9 \quad 90.$$

2.3.4 Imaging Noise Experiment

This experiment was designed to test the performance of the shape methods when the shapes have been perturbed by imaging noise. To simulate this effect, 50 unknown views of each of the six aircraft were generated. The images consisted of a background level, g_b , of 96 and a level, g_0 , of 160 when the pixel was contained in the view of the three-dimensional aircraft two-dimensional projection. Then zero mean, white, Gaussian noise, n_{ij} , was added to the image, I_{ij} . The image was then thresholded at 128. So, $\hat{I}_{ij} = I_{ij} + n_{ij}$,

where $E[n_{ij}] = 0$, $E[n_{ij} n_{km}] = \sigma^2 \delta_{(i-k)(j-m)}$ and

$$\text{Prob}[n_{ij} < n] = \int_{-\infty}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx. \quad \text{The image was then thresholded at 128.}$$

So,

$$\hat{I}_{ij}^t = \begin{cases} 1, & \text{if } \hat{I}_{ij} \geq 128 \\ 0, & \text{if } \hat{I}_{ij} < 128. \end{cases}$$

The contours of the resulting regions were traced and the chain code for the region having the longest chain code was retained to represent that unknown. These unknowns were then classified and the results tabulated for each method. The same operation was carried out on the six sets of 50 unknowns for four different noise levels corresponding to signal-to-noise ratios of 3, 6, 10,

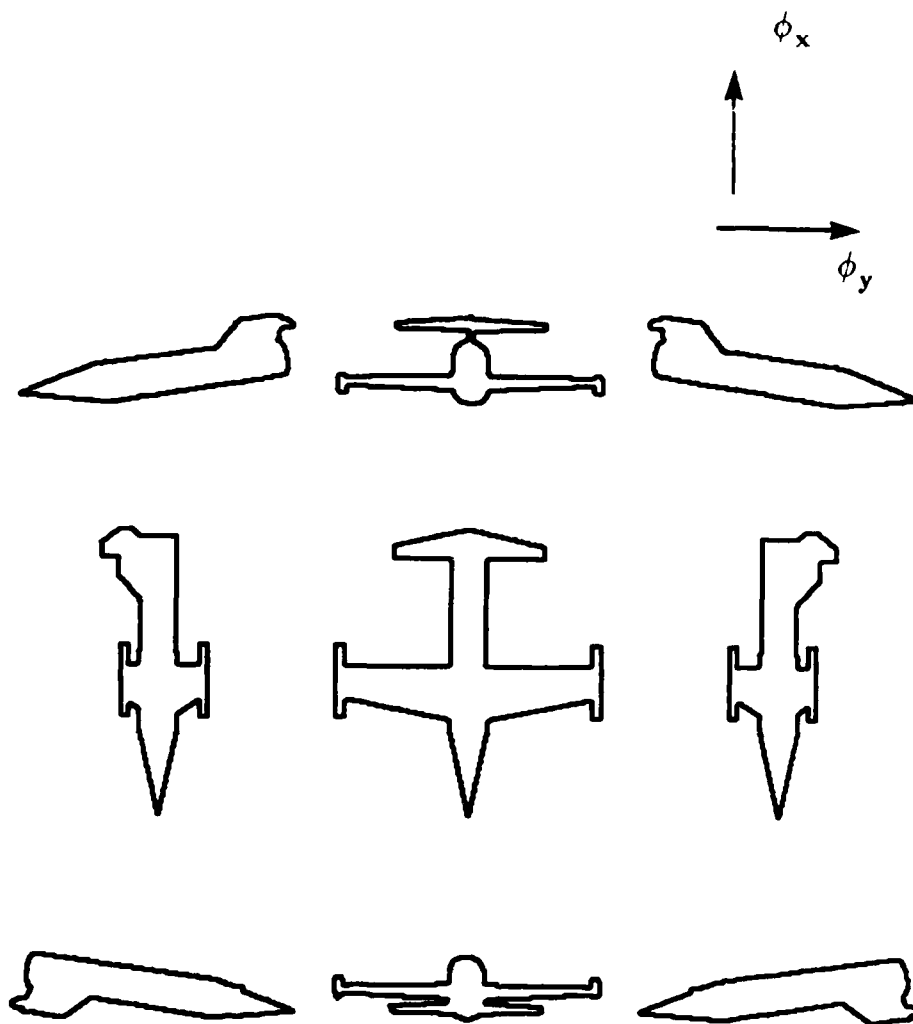


Figure 2.16 F104 library of 9 (3x3) views, 256x256 image resolution.

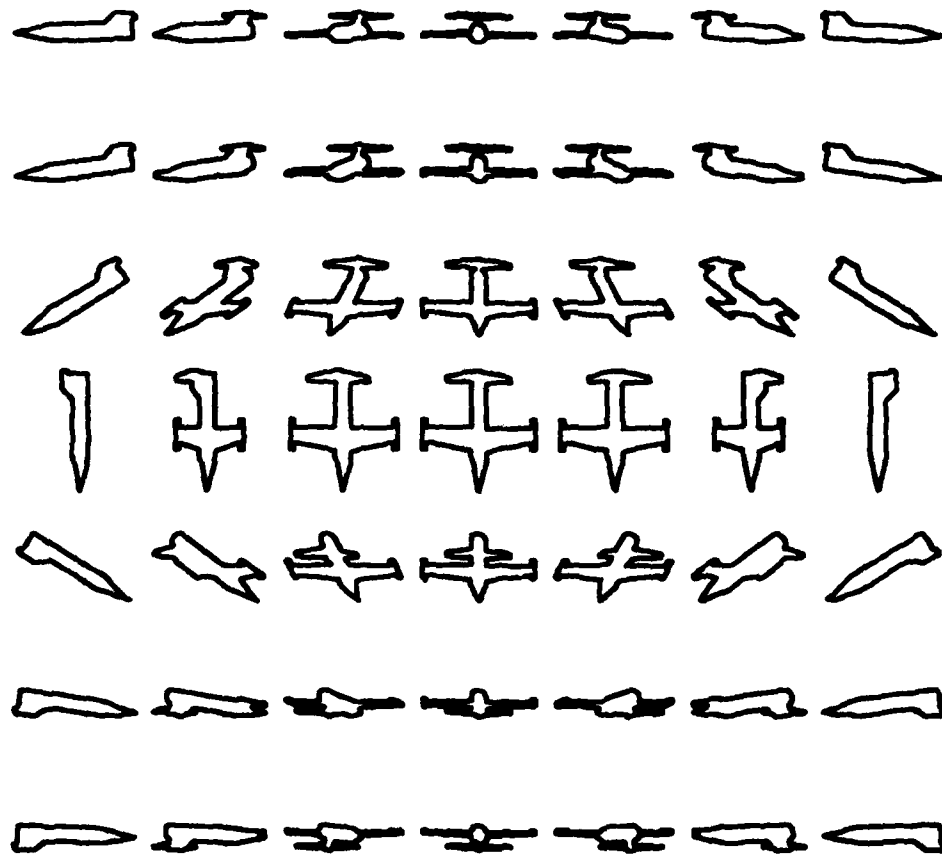


Figure 2.17 F104 library of 49 (7x7) views, 256x256 image resolution.

and 20 dB. The signal-to-noise ratio here is defined as

$$\text{SNR} = 20 \log \left(\frac{\Delta}{\sigma} \right),$$

where $\Delta = g_o - g_b$, and σ = standard deviation of the white Gaussian noise process. ($\Delta = 64$ and $\sigma = 40.3, 32.08, 20.02$, and 6.4 .) This same procedure was repeated so that noisy unknown shapes were generated at image resolutions of 64×64 and 128×128 image resolutions. Figures 2.18 to 2.27 depict images of an F104 at each of the five signal-to-noise ratios and their corresponding contours. The libraries were at an image resolution of 256×256 and contained 143 views per library.

2.3.5 Partial Shape Experiment

This experiment was designed to test the performance of the shape methods in classifying partial shapes. Partial shapes were generated by taking the 50 unknowns for each aircraft and chopping from 10% to 50% of the contour. The chopped out portion was replaced with a straight edge, a 90° turn, and another straight edge to close the contour. These unknowns were processed to extract their shape feature vectors. Each feature vector was reduced and classified. This experiment was performed with a constant feature vector length and with the libraries at a fixed image resolution (256×256) and number of views in each library (143). The unknowns were all at an image resolution of 128×128 . Figures 2.28 to 2.33 show examples of partial shapes at each of 0%, 10%, 20%, 30%, 40% and 50% chopped.

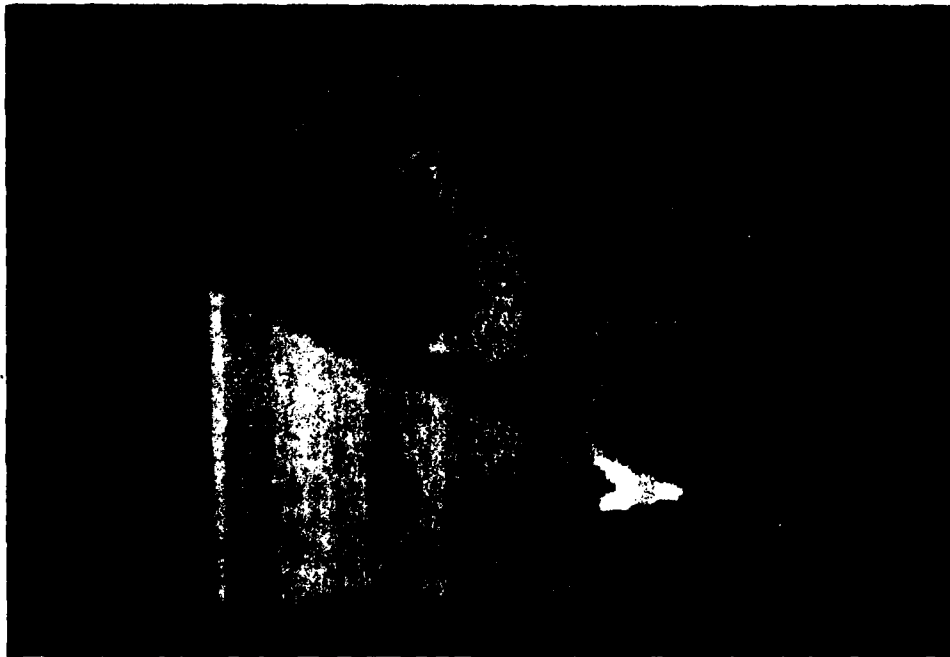


Figure 2.18 Noise free F104 silhouette, 128x128 image resolution.

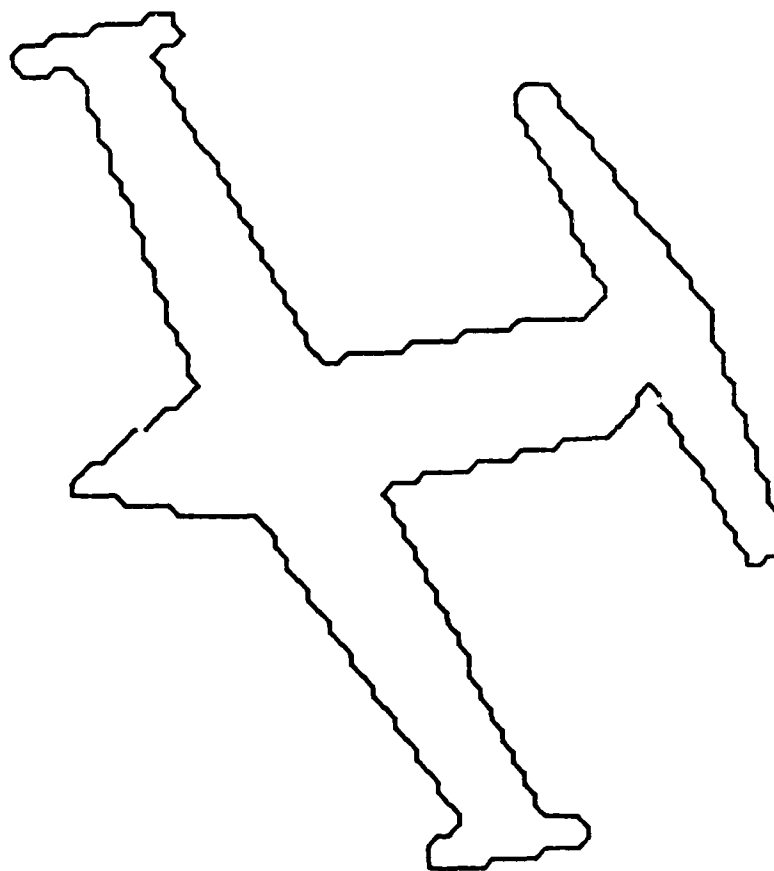


Figure 2.19 Noise free F104 contour, 128x128 image resolution.



Figure 2.20 F104 image with noise added ($\text{SNR} = 3 \text{ dB}$), 128×128 image resolution.

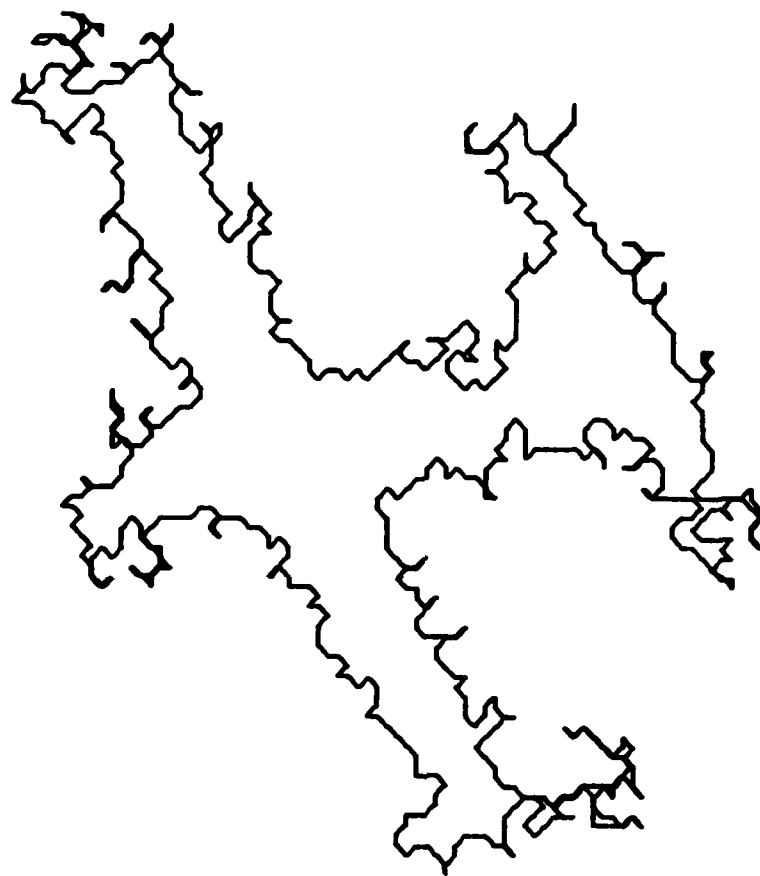


Figure 2.21 Noisy F104 contour ($\text{SNR} = 3 \text{ dB}$), 128×128 image resolution.

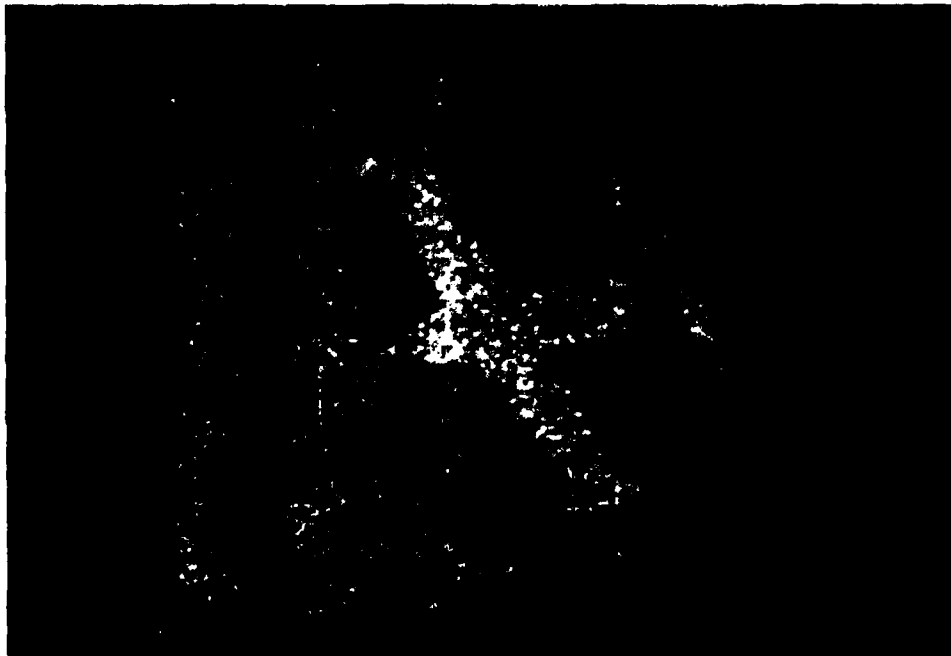


Figure 2.22 F104 image with noise added (SNR = 6 dB), 128x128 image resolution.

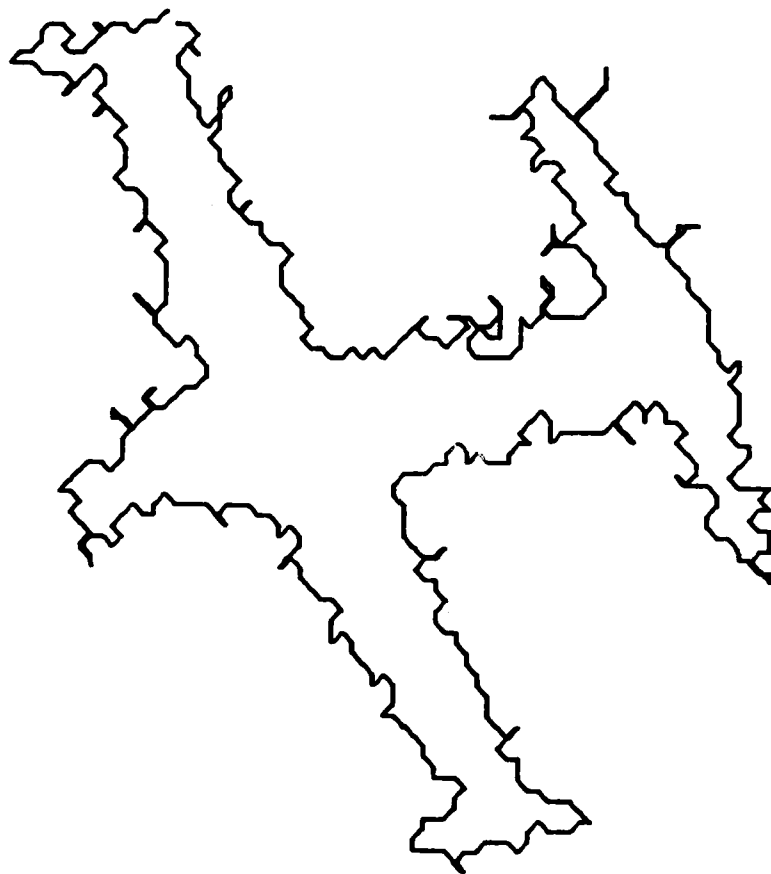


Figure 2.23 Noisy F104 contour ($\text{SNR} = 6 \text{ dB}$), 128x128 image resolution.



Figure 2.24 F104 image with noise added ($\text{SNR} = 10 \text{ dB}$), 128×128 image resolution.

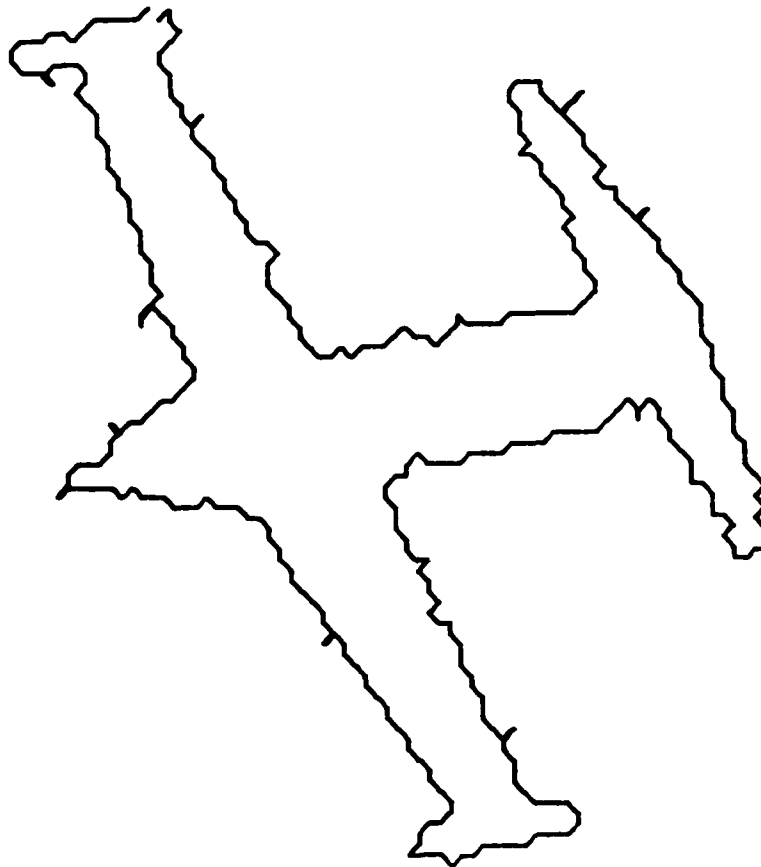


Figure 2.25 Noisy F104 contour ($\text{SNR} = 10 \text{ dB}$), 128x128 image resolution.

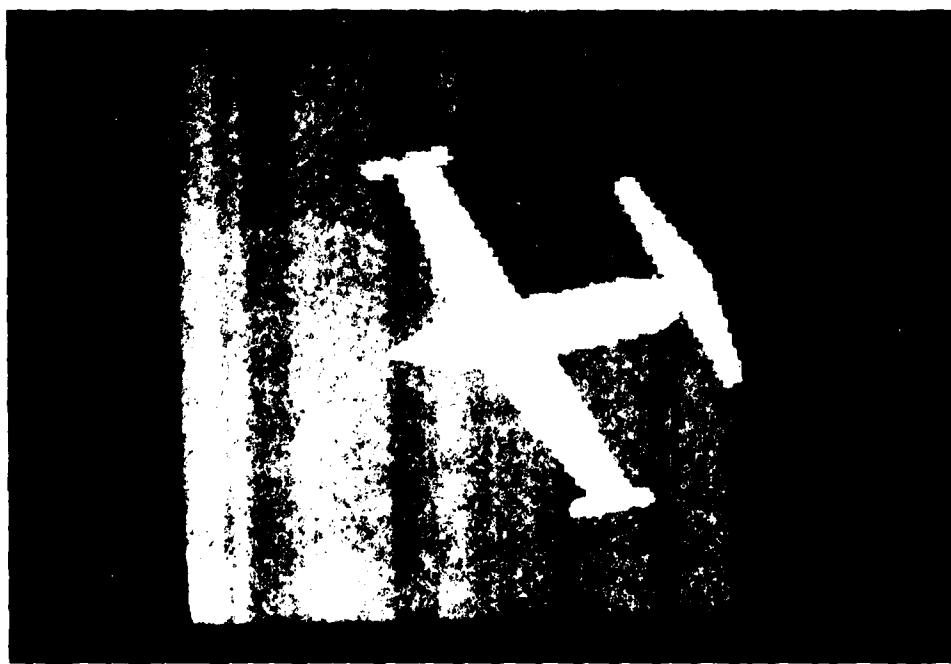


Figure 2.26 F104 image with noise added (SNR = 20 dB), 128x128 image resolution.

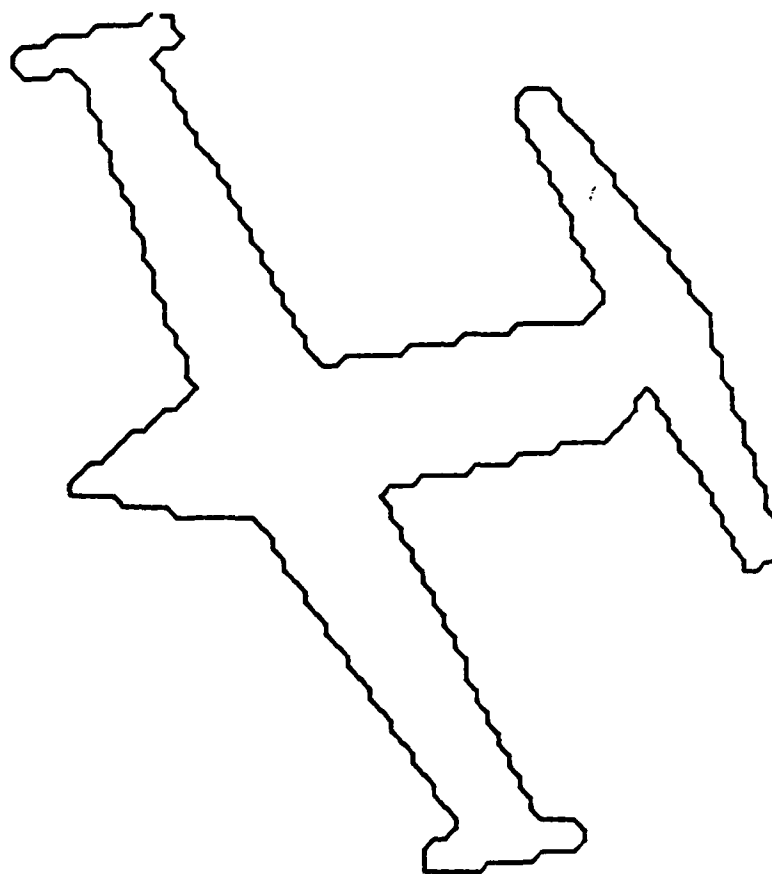


Figure 2.27 Noisy F104 contour (SNR = 20 dB), 128x128 image resolution.

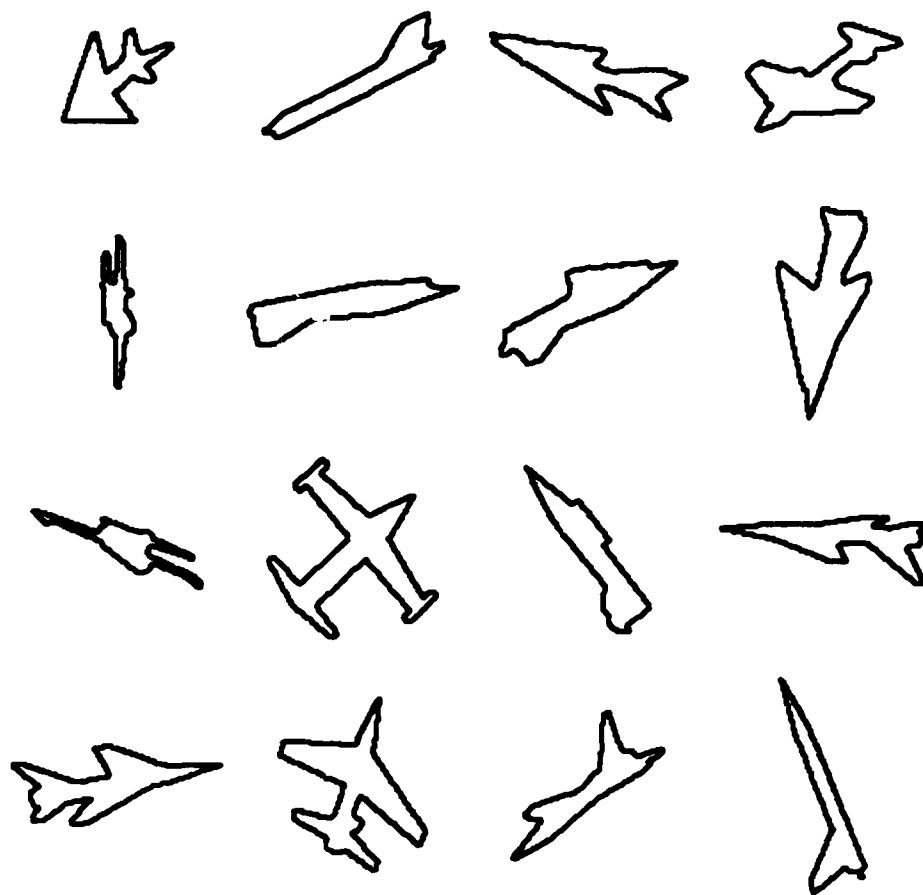


Figure 2.28 Sample unknown shape contours chopped 0 %.

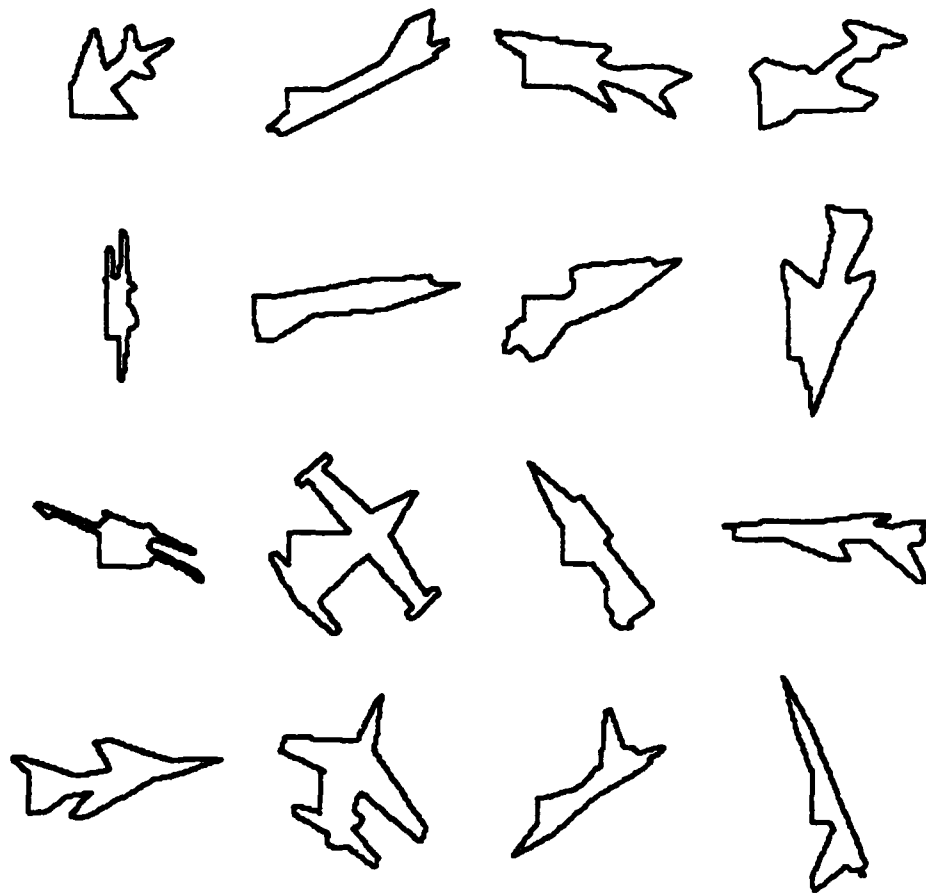


Figure 2.29 Sample unknown shape contours chopped 10 %.

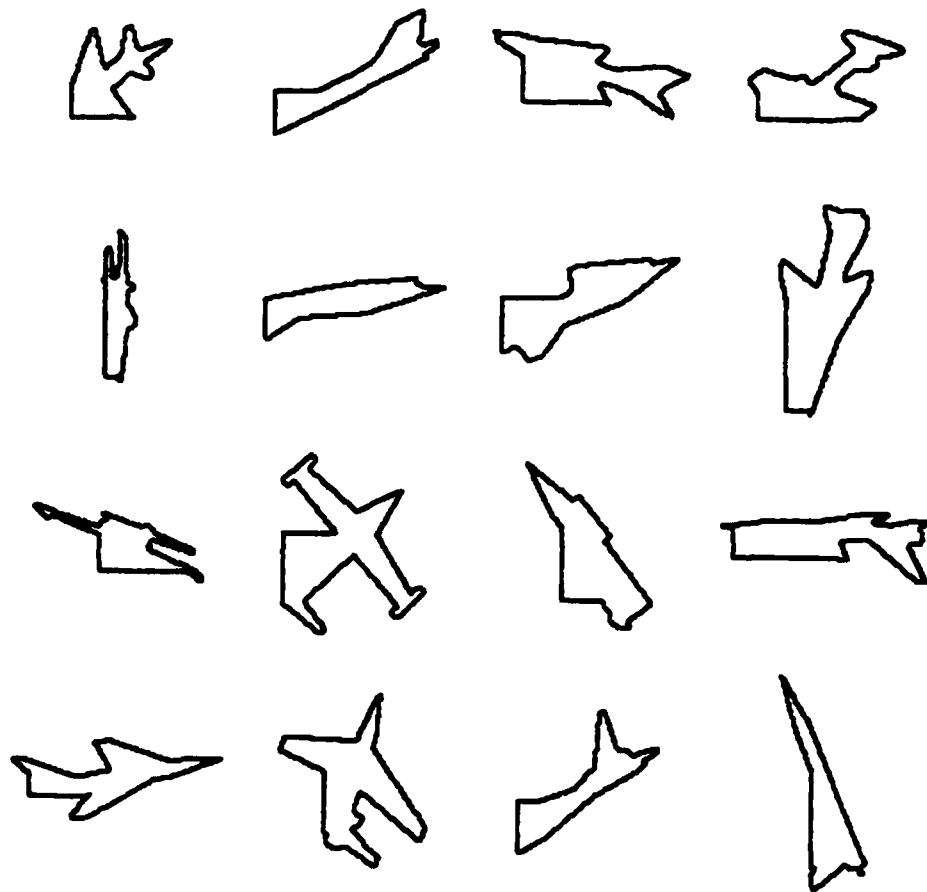


Figure 2.30 Sample unknown shape contours chopped 20 %.

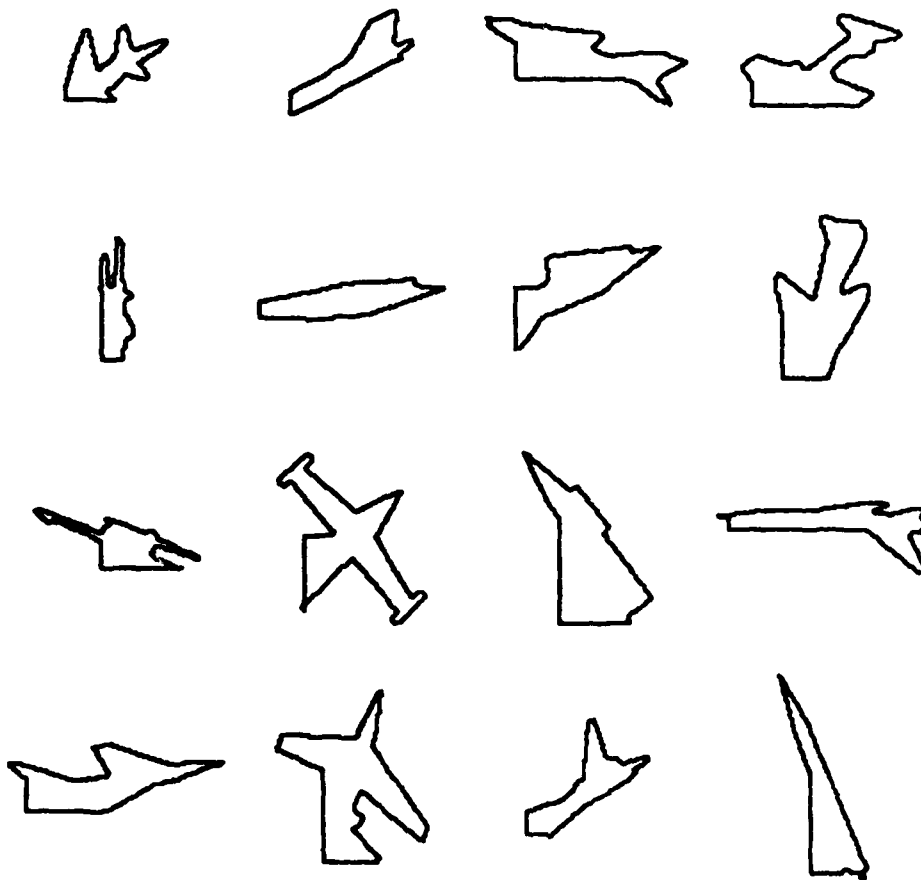


Figure 2.31 Sample unknown shape contours chopped 30 %.

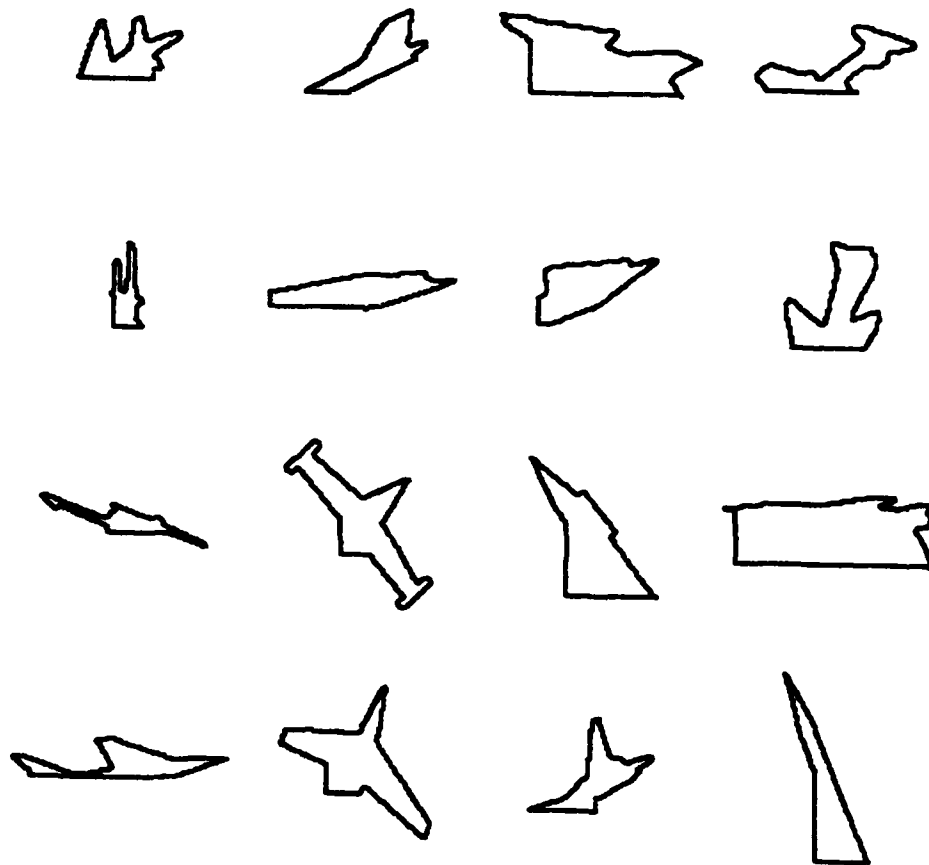


Figure 2.32 Sample unknown shape contours chopped 40 %.

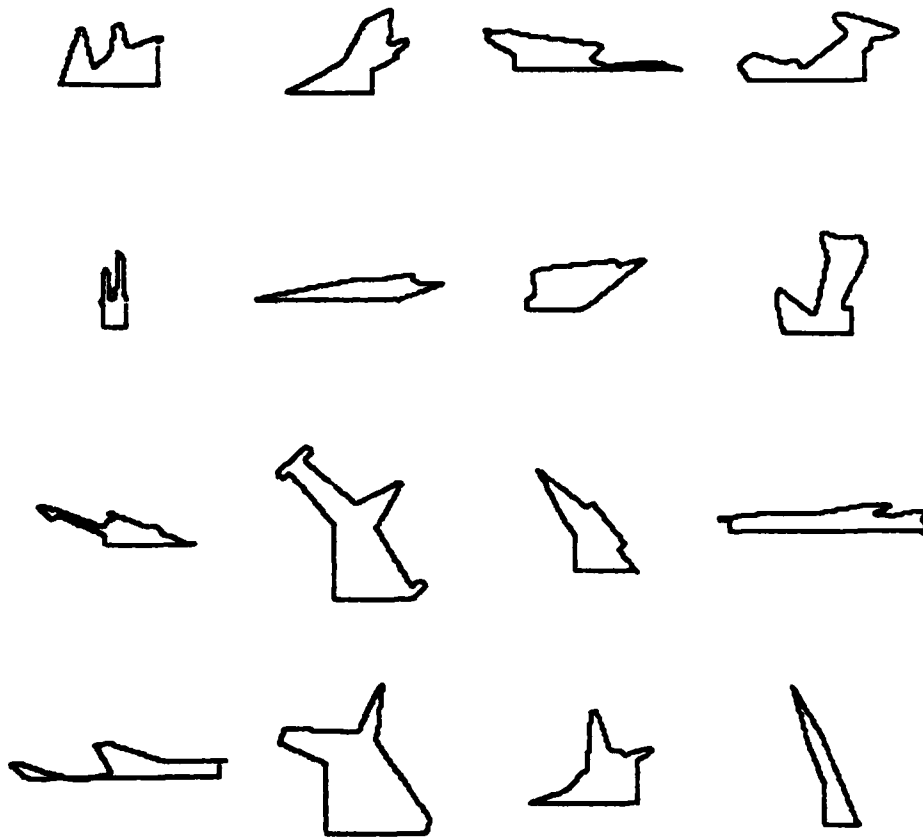


Figure 2.33 Sample unknown shape contours chopped 50 %.

CHAPTER 3

FOURIER DESCRIPTORS OF THE BOUNDARY

3.1 Introduction

The first method to be discussed is based on the well known method of Fourier series analysis for functional approximation. Having been provided with the silhouette of an object, the contour or boundary completely specifies the two-dimensional shape. The contour can be parameterized as a function of time by tracing around the boundary in counterclockwise direction. (See Figure 3.1.) As the tracing continues the function begins to repeat. Since this function is periodic, it can be expanded into a Fourier series. Each basis function of the complex exponential Fourier series is non-zero almost everywhere over each period. So, this is a global method of shape analysis.

The boundary function, γ , maps the real number line, $\mathbb{R}$, into the complex plane, $\mathbb{C}$. The projection on the real axis is the x component and the orthogonal projection on the imaginary axis is the y component.

$$\gamma : \mathbb{R} \rightarrow \mathbb{C}$$

$$\gamma(t) = (x(t), y(t)) \in \mathbb{R} \times \mathbb{R}$$

$$\gamma(t) = x(t) + iy(t) \in \mathbb{C}.$$

The velocity of the tracing is $v(t) = \frac{d\gamma(t)}{dt}$. So, the speed of tracing is equal to the magnitude of the velocity, i.e.

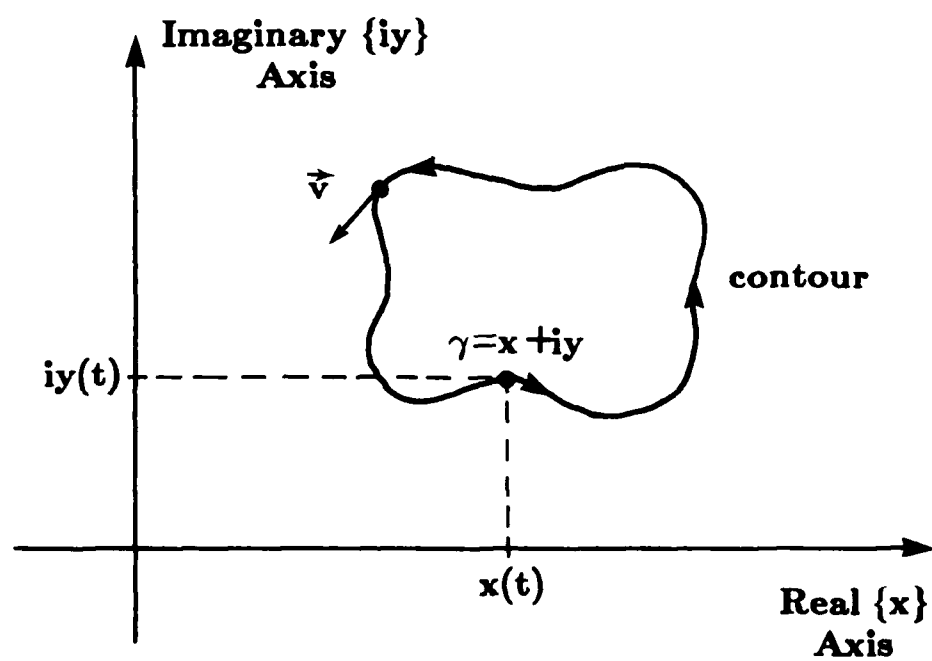


Figure 3.1 Boundary function in the complex plane.

$$|v(t)|^2 = \frac{d\gamma(t)}{dt} \frac{d\gamma^*(t)}{dt}.$$

If we assume that the tracing takes place at a constant speed (*uniform tracing*), then

$$v = |\dot{\gamma}(t)| = \left[\left(\frac{dx(t)}{dt} \right)^2 + \left(\frac{dy(t)}{dt} \right)^2 \right]^{1/2} = \text{constant}.$$

Since the speed is constant and the period of the trace is T , then $T = vL$, where L is the total arc length once around the contour. By definition, the arc length is

$$L = \int_0^T |\dot{\gamma}(t)| dt < \infty$$

Assuming $\gamma(t)$ is a continuous, bounded, periodic function with finite arc length (i.e. rectifiable), $\gamma(t)$ can be expanded in the Fourier series

$$\gamma(t) = \sum_{n=-\infty}^{+\infty} c_n e^{i \frac{2\pi n}{T} t}.$$

where

$$c_n = \frac{1}{T} \int_0^T \gamma(t) e^{-i \frac{2\pi n}{T} t} dt.$$

3.2 Properties

A great deal is known about the properties of the Fourier series. These properties can be exploited to provide some insight into the manner in which the Fourier series coefficients represent shape. Following is a list of many properties of the Fourier coefficients [GRAN72,RICH74,PERS77]. The properties will be used later to facilitate the manipulation of the Fourier coefficients for shape recognition. Several of the properties also apply to the shape description problem.

- 1) *Completeness:* The Fourier basis functions $e^{i \frac{2\pi n}{T} t}$ form an orthogonal basis, complete in $C[a, b]$ (or $L_2[a, b]$).
- 2) *Convergence:* Since the contour function γ is a bounded rectifiable Jordan simply connected closed contour (SCC), the Fourier series converges.
- 3) *Continuity.* Since $\gamma(t)$ is everywhere continuous (SCC) and its first is continuous almost everywhere, the modulus of the coefficients are of order $1/n^2$, i.e. $|c_n| = o(1/n^2)$.
- 4) *Translation:* If an object is translated in the plane by $Z_0 = x_0 + iy_0$, then $\gamma' = \gamma(t) + Z_0$ which implies that

$$c'_0 = c_0 + Z_0$$

$$c'_n = c_n, n \neq 0.$$

(See Figure 3.2.)

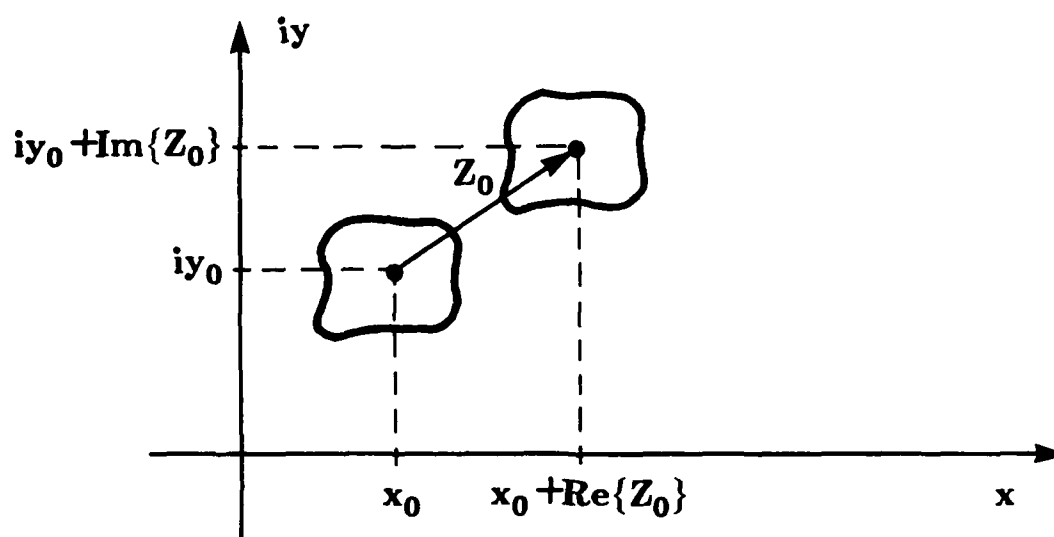


Figure 3.2 Translation.

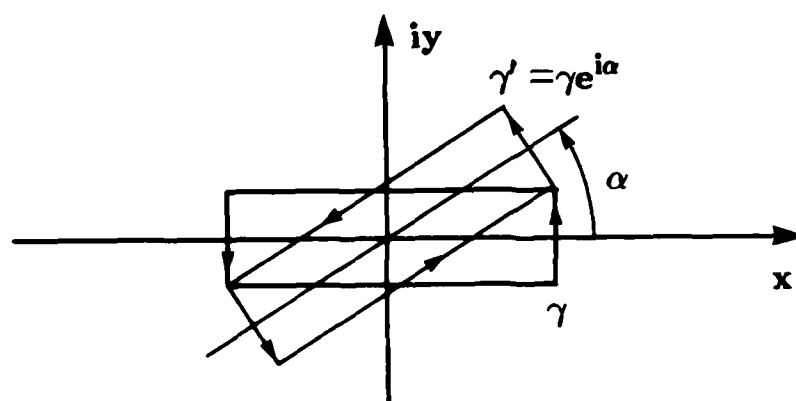


Figure 3.3 Rotation.

- 5) *Rotation*: Rotation of the shape through an angle α about the origin implies that c'_n , the coefficients of the rotated object have a constant phase added. That is

$$\gamma'(t) = \gamma(t) e^{i\alpha} \Rightarrow c'_n = c_n e^{i\alpha}, \forall n.$$

(See Figure 3.3.)

Proof:

$$x' = x \cos \alpha - y \sin \alpha$$

$$y' = x \sin \alpha + y \cos \alpha$$

So,

$$\gamma' = x' + iy' = (x + iy) e^{i\alpha}$$

$$= \gamma e^{i\alpha}$$

Hence,

$$c'_n = \frac{1}{T} \int_0^T \gamma'(t) e^{-i \frac{2\pi n}{T} t} dt = \frac{1}{T} \int_0^T \gamma(t) e^{i\alpha} e^{-i \frac{2\pi n}{T} t} dt$$

$$c'_n = c_n e^{i\alpha}.$$

Ω

- 6) *Dilation (scale)*: If the object γ is dilated by a factor λ , then

$$\gamma'(t) = \lambda \gamma(t) \Rightarrow c'_n = \lambda c_n, \forall n.$$

(See Figure 3.4.)

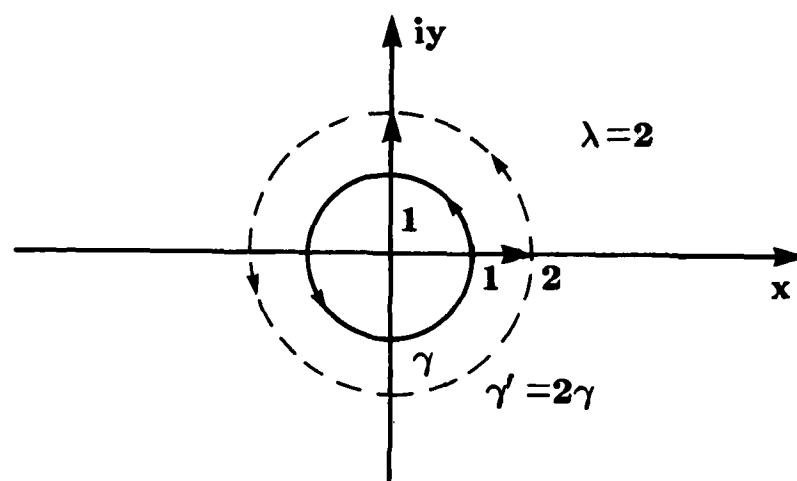


Figure 3.4 Dilation (scale).

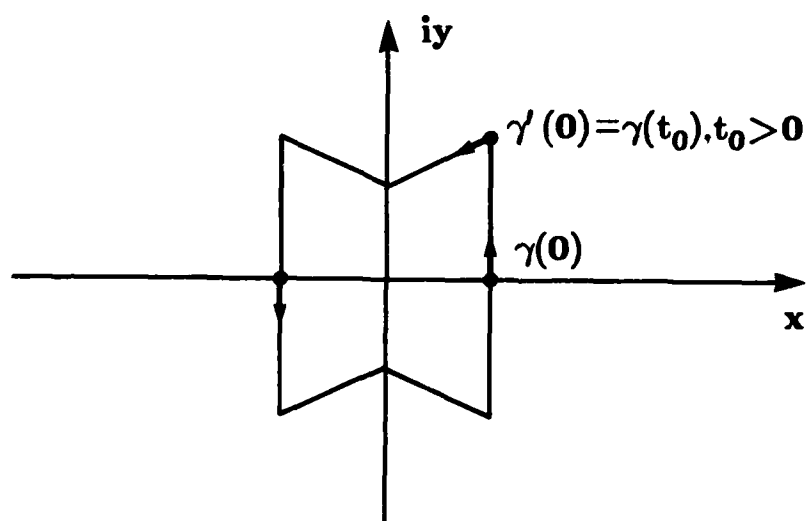


Figure 3.5 Starting point shift.

- 7) *Starting point shift*: Shifting the starting point counterclockwise along the contour, corresponds to beginning at the point in time, " t_0 ", on the original trace. (See Figure 3.5.) So,

$$\gamma'(t) = \gamma(t + t_0) \rightarrow c'_n = c_n e^{i \frac{2\pi n}{T} t_0}, \forall n.$$

Proof:

$$c'_n = \frac{1}{T} \int_0^T \gamma(t + t_0) e^{-i \frac{2\pi n}{T} t} dt.$$

Let $u = t + t_0$,

$$c'_n = \frac{1}{T} \int_{t_0}^{t_0+T} \gamma(u) e^{-i \frac{2\pi n}{T} (u-t_0)} du = \frac{1}{T} \int_{t_0}^{t_0+T} \gamma(u) e^{-i \frac{2\pi n}{T} u} du e^{i \frac{2\pi n}{T} t_0}$$

$$c'_n = c_n e^{i \frac{2\pi n}{T} t_0}, \forall n.$$

Ω

- 8) *N-fold Rotational Symmetry*: If a shape exhibits N-fold rotational symmetry, then rotating the object by an integral multiple of the angle $\alpha = \frac{2\pi}{N}$ about its center, c_0 , and moving the starting point clockwise by $\frac{kT}{N}$, we have the same trace. (See Figure 3.6.) Hence,

$$\gamma(t) - c_0 = (\gamma(t - \frac{kT}{N}) - c_0) e^{i \frac{2\pi k}{N}}$$

This implies that

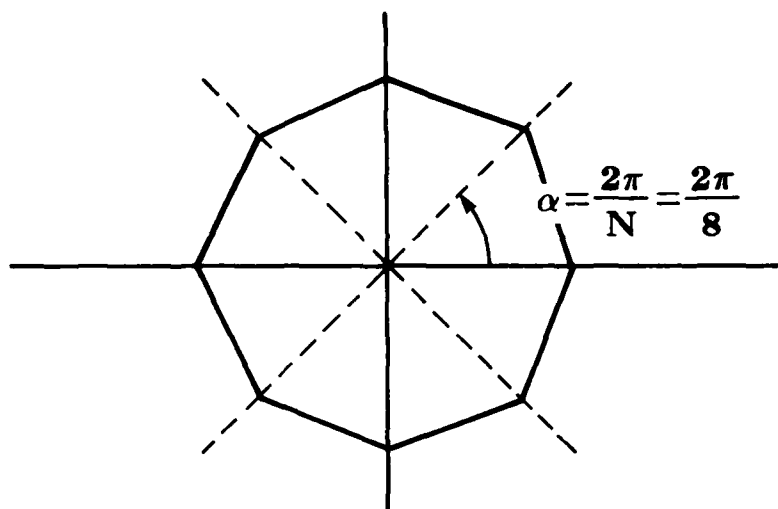


Figure 3.6 N-fold symmetry ($N=8$).

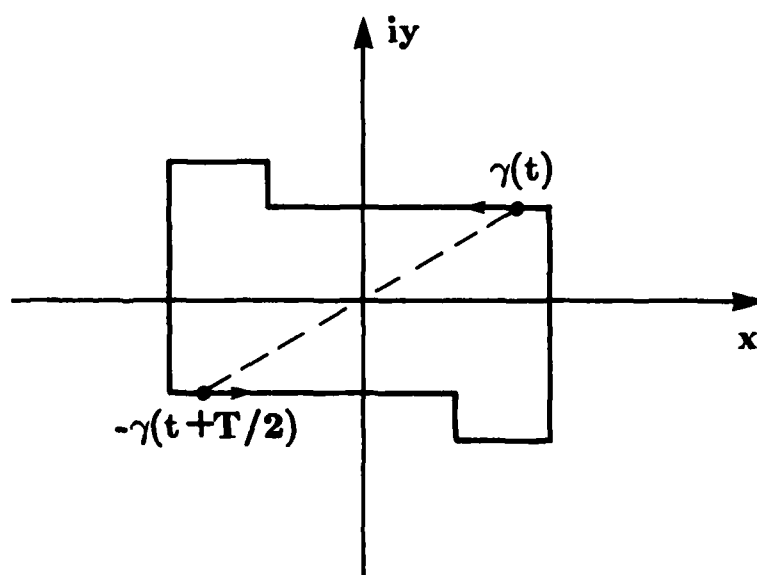


Figure 3.7 Axial symmetry.

$c_n \neq 0$ only if $(n-1) \bmod N = 0$,

c_0 can be anything,

$c_n = 0$, for all other n .

Proof:

$$c_n = \frac{1}{T} \int_0^T \gamma(t) e^{-i \frac{2\pi n}{T} t} dt$$

$$c_n = \frac{1}{T} \int_0^T \left(\gamma\left(t - \frac{kT}{N}\right) - c_0 \right) e^{i \frac{2\pi k}{N}} + c_0 e^{-i \frac{2\pi n}{T} t} dt$$

$$= e^{-i \frac{2\pi k}{N} n} \frac{1}{T} \int_{-\frac{kT}{N}}^{T - \frac{kT}{N}} ((\gamma(u) - c_0) e^{i \frac{2\pi k}{N}} + c_0) e^{-i \frac{2\pi n}{T} u} du$$

$$c_n = e^{-i \frac{2\pi k}{N} n} \frac{1}{T} \int_{-\frac{kT}{N}}^{T - \frac{kT}{N}} \gamma(u) e^{-i \frac{2\pi n}{T} u} du e^{i \frac{2\pi k}{N}}, \quad n \neq 0$$

and

$$c_0 = e^{-i \frac{2\pi k}{N}} c_0 - c_0 e^{-i \frac{2\pi k}{N}} + c_0.$$

So,

$$c_n = \begin{cases} e^{-i \frac{2\pi k}{N} (n-1)} c_0, & n \neq 0 \\ c_0. & \end{cases}$$

$$(1 - e^{-i\frac{2\pi}{N}k(n-1)})c_n = 0, \quad \forall n, n \neq 0, k = 0, 1, \dots, N-1.$$

which implies that

$$c_n \neq 0, \text{ only if } (n-1) \bmod N = 0.$$

and

$$c_n = 0, \text{ otherwise, except } c_0 \text{ can be anything.}$$

Ω

9) *Axial Symmetry*: For a shape which is symmetric about its center (Figure

$$3.7), c_0, \text{ then } \gamma(t) - c_0 = -\left[\gamma\left(t + \frac{T}{2}\right) - c_0\right]. \text{ Which implies that}$$

$$c_n \neq 0, \text{ only if } n \text{ is odd}$$

$$c_n = 0, \text{ otherwise, except}$$

$$c_0 = \text{anything.}$$

Proof:

$$c_n = \frac{1}{T} \int_0^{\frac{T}{2}} \gamma(t) e^{-i\frac{2\pi n}{T}t} dt + \frac{1}{T} \int_{\frac{T}{2}}^T \gamma(t) e^{-i\frac{2\pi n}{T}t} dt.$$

$$\text{Letting } \lambda = t - \frac{T}{2},$$

$$c_n = \frac{1}{T} \int_0^{\frac{T}{2}} \gamma(t) e^{-i \frac{2\pi n}{T} t} dt + \frac{1}{T} \int_0^{\frac{T}{2}} \gamma(\lambda + \frac{T}{2}) e^{-i \frac{2\pi n}{T} (\lambda + \frac{T}{2})} d\lambda \\ + \frac{2}{T} \int_0^{\frac{T}{2}} c_0 e^{-i \frac{2\pi n}{T} (\lambda + \frac{T}{2})} d\lambda$$

$$c_n = \frac{1}{T} \int_0^{\frac{T}{2}} \gamma(t) e^{-i \frac{2\pi n}{T} t} dt (1 - e^{-in\pi}), n \neq 0$$

$$c_n = c_0, n = 0.$$

Ω

- 10) *Bilateral Symmetry*: If γ is symmetric about the real axis (see Figure 3.8), then for some $t_0 \in [0, T]$, $\tilde{\gamma}(t) = \gamma(t + t_0) = \tilde{\gamma}^*(-t)$ implies that $\tilde{c}_n$ are purely real.

Proof:

Since γ is continuous and symmetric about the real axis, then it must cross the real axis. Let t_0 be the shift necessary so that $\tilde{\gamma}(t) = \gamma^*(-t) = \gamma(t + t_0)$.

$$\tilde{c}_n = \frac{1}{T} \int_0^T \tilde{\gamma}(t) e^{-i \frac{2\pi n}{T} t} dt = \frac{1}{T} \int_0^T \tilde{\gamma}^*(-t) e^{-i \frac{2\pi n}{T} t} dt.$$

$$= \left[\frac{1}{T} \int_0^T \tilde{\gamma}(-t) e^{i \frac{2\pi n}{T} t} dt \right]^* = \left[\frac{1}{T} \int_{-T}^0 \tilde{\gamma}(t) e^{-i \frac{2\pi n}{T} t} dt \right]^*$$

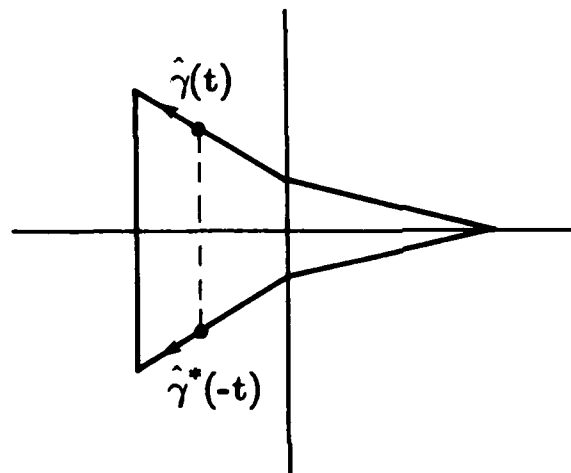


Figure 3.8 Bilateral symmetry.

$$\tilde{c}_n = \tilde{c}_n^*.$$

Ω

- 11) *Reflections*: Let γ' be an object which is the reflection of γ about the real (x) axis (see Figure 3.9.) Then

$$\gamma'(t) = \gamma^*(-t) \Rightarrow c'_n = c_n^*.$$

Proof:

$$\begin{aligned} c'_n &= \frac{1}{T} \int_0^T \gamma'(t) e^{-i \frac{2\pi n}{T} t} dt = \frac{1}{T} \int_0^T \gamma^*(-t) e^{-i \frac{2\pi n}{T} t} dt. \\ &= -\frac{1}{T} \int_0^{-T} \gamma^*(\lambda) e^{i \frac{2\pi n}{T} \lambda} d\lambda = \frac{1}{T} \int_{-T}^0 \gamma^*(\lambda) e^{i \frac{2\pi n}{T} \lambda} d\lambda \\ &= \frac{1}{T} \int_0^T \gamma^* e^{i \frac{2\pi n}{T} t} dt = \left[\frac{1}{T} \int_0^T \gamma(t) e^{-i \frac{2\pi n}{T} t} dt \right]^* \end{aligned}$$

$$c'_n = c_n^*.$$

Ω

- 12) *Area*: For a simply connected closed curve γ , bounding the region, R ,

$$\text{Area} = \iint_R dx dy.$$

Using Green's Theorem,

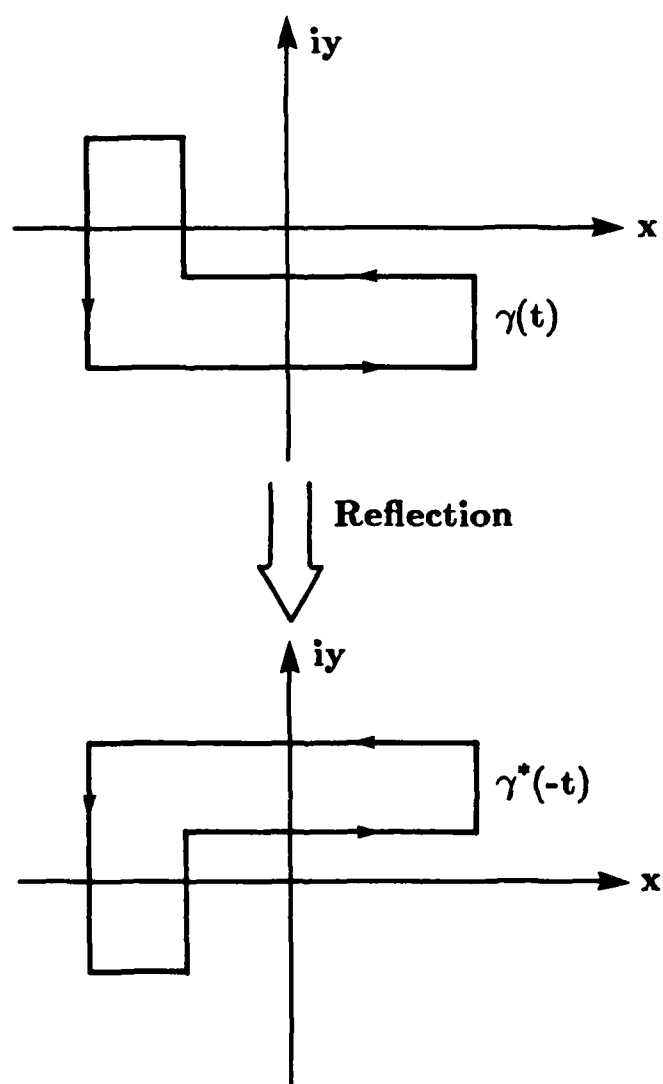


Figure 3.9 Reflection.

$$\begin{aligned}\text{Area} &= \frac{1}{2} \int_{\gamma} x dy - \frac{1}{2} \int_{\gamma} y dx \\ &= \frac{1}{2} \int_0^T \left[x \frac{dy}{dt} - y \frac{dx}{dt} \right] dt\end{aligned}$$

$$\text{Area} = \sum_{n=-\infty}^{+\infty} \pi n |c_n|^2.$$

Proof:

$$x = \text{Re}\{\gamma\} = \frac{\gamma + \gamma^*}{2}$$

$$y = \text{Im}\{\gamma\} = \frac{\gamma - \gamma^*}{2i}$$

$$x = \frac{1}{2} \sum_n (c_n e^{i \frac{2\pi n}{T} t} + c_n^* e^{-i \frac{2\pi n}{T} t})$$

$$y = \frac{1}{2i} \sum_n (c_n e^{i \frac{2\pi n}{T} t} - c_n^* e^{-i \frac{2\pi n}{T} t})$$

$$\frac{dx}{dt} = \frac{i\pi}{T} \sum_n n (c_n e^{i \frac{2\pi n}{T} t} - c_n^* e^{-i \frac{2\pi n}{T} t}); \quad \frac{dy}{dt} = \frac{\pi}{T} \sum_n n (c_n e^{i \frac{2\pi n}{T} t} + c_n^* e^{-i \frac{2\pi n}{T} t})$$

$$\text{Area} = \frac{1}{2} \int_0^T x \frac{dy}{dt} - \frac{1}{2} \int_0^T \frac{dx}{dt} y dt$$

$$= \frac{1}{2} \int_0^T \frac{1}{2} \sum_n (c_n e^{i \frac{2\pi n}{T} t} + c_n^* e^{-i \frac{2\pi n}{T} t}) \frac{\pi}{T} \sum_m (c_m e^{i \frac{2\pi m}{T} t} + c_m^* e^{-i \frac{2\pi m}{T} t}) dt$$

$$- \frac{1}{2} \int_0^T \frac{i\pi}{T} \sum_n n (c_n e^{i \frac{2\pi n}{T} t} - c_n^* e^{-i \frac{2\pi n}{T} t}) \frac{1}{2i} \sum_m (c_m e^{i \frac{2\pi m}{T} t} - c_m^* e^{-i \frac{2\pi m}{T} t}) dt$$

$$= \frac{\pi}{4T} \int_0^T \sum_n \sum_m [m c_n c_m e^{i \frac{2\pi}{T} (n+m)t} + m c_n c_m^* e^{i \frac{2\pi}{T} (n-m)t}]$$

$$\begin{aligned}
& + m c_n^* c_m e^{i \frac{2\pi}{T}(m-n)t} + m c_n^* c_m^* e^{-i \frac{2\pi}{T}(m+n)t}] dt \\
& - \frac{\pi}{4T} \int_0^T \sum_n \sum_m [n c_n c_m e^{i \frac{2\pi}{T}(n+m)t} - n c_n c_m^* e^{i \frac{2\pi}{T}(n-m)t} \\
& \quad - n c_n^* c_m e^{i \frac{2\pi}{T}(m-n)t} + n c_n^* c_m^* e^{-i \frac{2\pi}{T}(m+n)t}] dt \\
& = \frac{\pi}{4} \sum_n [-n c_n c_{-n} + n c_n c_n^* - c_n^* c_{-n}^* - n c_n c_{-n} + n c_n c_n^* - c_n^* c_{-n}^*] \\
\text{Area} & = \pi \sum_n n [c_n c_n^* - \frac{1}{2} c_n c_{-n} - \frac{1}{2} c_n^* c_{-n}^*] \\
& = -\frac{\pi}{2} \sum_{n=-\infty}^{+\infty} n [c_n c_{-n} - 2 c_n c_n^* + c_n^* c_{-n}^*] \\
& = -\frac{\pi}{2} \sum_{n=1}^{\infty} [n c_n c_{-n} - 2 n c_n c_n^* + n c_n^* c_{-n}^* - \\
& \quad n c_{-n} c_n + 2 n c_{-n} c_{-n}^* - n c_{-n}^* c_n^*] \\
& = -\frac{\pi}{2} \sum_{n=1}^{\infty} [2 n c_{-n} c_{-n}^*] \\
& = \pi \sum_{n=-\infty}^{+\infty} n c_n c_n^* = \pi \sum_{n=-\infty}^{+\infty} n |c_n|^2
\end{aligned}$$

- 13) *Line shape*: A line shape (see Figure 3.10) is one with no interior such that after some time $\frac{T}{2} - \alpha$ ($0 \leq \alpha \leq \frac{T}{2}$), the contour begins to exactly retrace but going in the opposite direction. That is

$$\gamma(t + (\frac{T}{2} - \alpha)) = \gamma(-t + \frac{T}{2}).$$

which implies that

$$c_{-n} = c_n e^{-i \frac{n\pi\alpha}{T}}.$$

And if the starting point is an endpoint of the line shape, $\alpha = 0$. In that case

$$c_{-n} = c_n.$$

Note that a line shape can have only support at most 2-fold symmetry.

Also since $|c_{-n}| = |c_n|$, then

$$\text{Area} = \sum_{n=1}^{\infty} n [|c_n|^2 - |c_{-n}|^2] \equiv 0.$$

Proof:

Given $\gamma(t) = \gamma(T - \alpha - t)$,

$$\begin{aligned} c_n &= \frac{1}{T} \int_{-\frac{\alpha}{2}}^{\frac{T}{2} - \frac{\alpha}{2}} \gamma(t) e^{-i \frac{2\pi n}{T} t} dt + \frac{1}{T} \int_{\frac{T}{2} - \frac{\alpha}{2}}^{T - \frac{\alpha}{2}} \gamma(t) e^{-i \frac{2\pi n}{T} t} dt \\ &= \frac{1}{T} \int_{-\frac{\alpha}{2}}^{\frac{T}{2} - \frac{\alpha}{2}} \gamma(t) e^{-i \frac{2\pi n}{T} t} dt + \frac{1}{T} \int_{\frac{T}{2} - \frac{\alpha}{2}}^{T - \frac{\alpha}{2}} \gamma(T - \alpha - t) e^{-i \frac{2\pi n}{T} t} dt \end{aligned}$$

Let $\lambda = T - \alpha - t$,

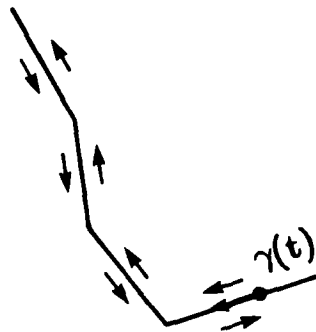


Figure 3.10 Line shape.

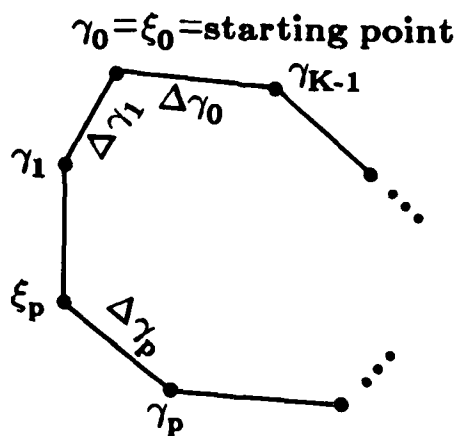


Figure 3.11 Polygon parameter definitions for Fourier descriptors of the boundary.

$$\begin{aligned}
c_n &= \frac{1}{T} \int_{-\frac{\alpha}{2}}^{\frac{T}{2} - \frac{\alpha}{2}} \gamma(t) e^{-i \frac{2\pi n}{T} t} dt + \frac{1}{T} \int_{-\frac{\alpha}{2}}^{\frac{T}{2} - \frac{\alpha}{2}} \gamma(\lambda) e^{i \frac{2\pi n \lambda}{T}} e^{-i \frac{2\pi n}{T} \alpha} d\lambda \\
&= \frac{1}{T} \int_{-\frac{\alpha}{2}}^{\frac{T}{2} - \frac{\alpha}{2}} \gamma(t) [e^{-i \frac{2\pi n}{T} t} + e^{i \frac{2\pi n}{T} t} e^{-i \frac{2\pi n \alpha}{T}}] dt \\
&= \frac{2}{T} e^{-i \frac{n\pi\alpha}{T}} \int_{-\frac{\alpha}{2}}^{\frac{T}{2} - \frac{\alpha}{2}} \gamma(t) \cos \frac{2\pi n}{T} (t - \frac{\alpha}{2}) dt.
\end{aligned}$$

So,

$$c_{-n} = c_n e^{-i \frac{n\pi\alpha}{T}}$$

Ω

- 14) *Error bound:* An error bound for the truncated Fourier series has been derived by Giardina [GIAR77]. If γ is written as $\vec{\gamma} = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$, then given $\epsilon > 0$ there exists a K such that if $n > K$ then

$$\|\vec{\gamma} - \vec{\gamma}_n\| < \epsilon, \quad \text{where } \vec{\gamma}_n = \begin{pmatrix} x_n(t) \\ y_n(t) \end{pmatrix}.$$

where

$$K = \frac{T}{2\pi^2\epsilon} \max[V_0^T(\dot{x}(t)), V_0^T(\dot{y}(t))].$$

$V_0^T(\dot{x}(t))$ and $V_0^T(\dot{y}(t))$ are the total variation of $\dot{x}(t)$ and $\dot{y}(t)$ respectively.

The Fourier series of $x(t)$ and $y(t)$ are truncated to n terms to form $x_n(t)$

and $y_n(t)$. The norm is defined as

$$\|\vec{\gamma} - \vec{\gamma}_n\| = \max \left[\sup_{0 \leq t \leq T} |x - x_n|, \sup_{0 \leq t \leq T} |y - y_n| \right].$$

3.3 Calculation

Since the images are discrete, connecting the contours of the pixels around the boundary produces a polygon. The Fourier coefficients can be computed directly from the increments in time (arc length) and position as the polygon is traced. (See Figure 3.11.) This is called the direct Fourier transform (DFT). The following derivation follows along the lines of [KUHL82], except here the complex exponential series is used. Also, a much more simple expression for the DC term is obtained.

Let the increment in position be the complex number $\Delta\gamma_i$, and the increment in time will be $\Delta t_i = |\Delta\gamma_i| = \sqrt{\Delta x_i^2 + \Delta y_i^2}$. Since γ is continuous, its derivative, $\dot{\gamma}$, exists almost everywhere. So, let

$$\dot{\gamma} = \sum_{n=-\infty}^{+\infty} \beta_n e^{i \frac{2\pi n}{T} t}$$

be the Fourier series for $\dot{\gamma}$. But the series for γ can be differentiated term by term so that

$$\dot{\gamma} = \sum_{n=-\infty}^{+\infty} i \frac{2\pi n}{T} c_n e^{i \frac{2\pi n}{T} t}$$

So,

$$\beta_n = i \frac{2\pi n}{T} c_n \rightarrow c_n = -i \frac{T}{2\pi n} \beta_n, \quad n \neq 0.$$

Now,

$$\begin{aligned}\beta_n &= \frac{1}{T} \int_0^T \gamma(t) e^{-i \frac{2\pi n}{T} t} dt = \frac{1}{T} \sum_{p=1}^K \frac{\Delta\gamma_p}{\Delta t_p} \int_{t_{p-1}}^{t_p} e^{-i \frac{2\pi n}{T} t} dt \\ &= \frac{1}{T} \sum_{p=1}^K \frac{\Delta\gamma_p}{\Delta t_p} (e^{-i \frac{2\pi n}{T} t_p} - e^{-i \frac{2\pi n}{T} t_{p-1}}) (i \frac{T}{2\pi n})\end{aligned}$$

We can then write

$$c_n = \frac{T}{4\pi^2 n^2} \sum_{p=1}^K \frac{\Delta\gamma_p}{\Delta t_p} \left(e^{-i \frac{2\pi n}{T} t_p} - e^{-i \frac{2\pi n}{T} t_{p-1}} \right), \quad n \neq 0,$$

where

$$t_p = \sum_{i=1}^p \Delta t_i, \quad t_0 = 0,$$

K is the number of sides on the polygon, and $T = t_K =$ period. For the c_0 term,

$$\begin{aligned}c_0 &= \frac{1}{T} \int_0^T \gamma(t) dt = \frac{1}{T} \sum_{p=1}^K \int_{t_{p-1}}^{t_p} \gamma(t) dt \\ &= \frac{1}{T} \sum_{p=1}^K \int_{t_{p-1}}^{t_p} \left(\frac{\Delta\gamma_p}{\Delta t_p} t + \zeta_p \right) dt \\ &= \frac{1}{T} \sum_{p=1}^K [1/2 \Delta\gamma_p (\Delta t_p + 2 t_{p-1}) + \zeta_p \Delta t_p],\end{aligned}$$

where $\zeta_p = \gamma_{p-1} - \frac{\Delta\gamma_p}{\Delta t_p} t_{p-1}$, $\gamma_p = \gamma_{p-1} + \Delta\gamma_p$, $\zeta_1 = \gamma_0 =$ starting point.

Hence,

$$c_0 = \frac{1}{T} \sum_{p=1}^K (1/2 \Delta\gamma_p + \gamma_{p-1}) \Delta t_p.$$

For the discrete images, the contours are represented by the eight-direction chain code. So, the $\Delta\gamma_i$ and Δt_i can be determined using a lookup table indexed by the chain code.

$a_i:$	0	1	2	3	4	5	6	7
$\Delta\gamma:$	1	$1+i$	i	$-1+i$	-1	$-1-i$	$-i$	$1-i$
$\Delta t:$	1	$\sqrt{2}$	1	$\sqrt{2}$	1	$\sqrt{2}$	1	$\sqrt{2}$

The computational complexity is of order $N \cdot K$, where N is the number of coefficients calculated and K is the number of links in the chain coded contour. Using the chain codes can be inefficient, however, when there are runs of identical chain codes. By requiring additional storage for two vectors (one complex and one real), the computation necessary for the coefficients can be reduced. It is necessary to compute the vectors $\{\Delta t'_i\}$ and $\{\Delta\gamma'_i\}$ for $i = 0, 1, \dots, K'$ $K' < K$, where $\Delta\gamma'_i$ is the sum of the $\Delta\gamma_j$ for the i^{th} run and $\Delta t'_i$ is the sum of the Δt_j for the i^{th} run.

An alternative to calculating a direct transform of the polygon is to use the Fast Fourier Transform (FFT), whose computational complexity is of order $N \log_2 N$, $N = 2^m$. The FFT can be used if the discrete (complex) values representing the contour are resampled so that there are $N = 2^m$ discrete values. If the resampling can be done efficiently and with minimal distortion of the signal (shape), then the FFT can be used to compute $N = 2^m$ coefficients quickly.

However, there is distortion introduced by the resampling and also by assuming that the signal is now a periodic impulse train instead of a continuous complex valued function. If the contour is resampled to $N = 2^m \gg K$, then the distortion is minimal, but now the sequence of data values is much longer incurring a greater computational burden. The resampling operation also must be included in determining the computational complexity of this approach.

Since the contours used in the experiments are about 200 to 2000 chain code links and 32 complex Fourier coefficients are usually calculated, the resampling followed by the FFT was used. The resampling method used linear interpolation between the complex data values. If the sequence γ_p above is the input to the resampling, and $\tilde{\gamma}_j$ is the sequence resampled to $N = 2^m \geq K$ values, then the n^{th} coefficient using the FFT is

$$\tilde{c}_n = \sum_{j=0}^{N-1} \tilde{\gamma}_j e^{-\frac{2\pi}{N}nj}, \quad n = 0, 1, \dots, N-1$$

and where

$$\tilde{c}_{-n} = \sum_{j=0}^{N-1} \tilde{\gamma}_j e^{-i\frac{2\pi}{N}(N-n)j} = \sum_{j=0}^{N-1} \tilde{\gamma}_j e^{i\frac{2\pi n}{N}j} = \tilde{c}_{N-n}.$$

So, the FFT gives the $(-N/2 + 1)^{\text{th}}$ to $(N/2)^{\text{th}}$ coefficient.

As previously mentioned, this assumes that

$$\tilde{\gamma}(t) = \sum_{j=-\infty}^{\infty} \tilde{\gamma}_j \delta(t - \frac{jT}{N}), \quad \text{where } \tilde{\gamma}_{j+Nk} = \tilde{\gamma}_j, \quad \forall k.$$

Also notice that the $|\tilde{\gamma}_j|$'s are not necessarily equal. So, while the DFT gives the exact results for a continuous uniformly traced polygon, the FFT implementation is based on a sequence which is not uniform sampling along the contour. Sometimes the sequence γ_p is very much over sampled in order to

cause $\tilde{\gamma}_j$ to be approximately uniform sampling. But if a large amount of over sampling is performed, then the computational advantage of the FFT is lost.

One advantage of the DFT method is that only the coefficients actually desired need be calculated. On the other hand with the FFT method, you obtain 2^m coefficients whether or not that many are actually desired.

3.4 Normalization

It is usually desirable to compare shapes independent of size, orientation, and starting point. To compare two shape boundaries using their Fourier descriptors, it is necessary to scale, rotate, and shift their shapes in order to allow the "best" fit possible. This operation *normalizes* the Fourier descriptors for the unknown shape to an optimum orientation. It has been shown [PERS77] that the optimum scale, rotation angle, and relative starting point shift can be obtained to minimize the mean-squared error as the criterion.

Given two shapes γ and γ' which have coefficients c_n and c'_n , respectively, minimize the distance function

$$d(\gamma, \gamma') = \left[\sum_n^{N/2} |c_n - s e^{i(n\phi + \alpha)} c'_n|^2 \right]^{1/2},$$

where s , ϕ , and α are the scale, relative starting point, and rotation, respectively. In practice, only N coefficients are used to represent the shape. In that case, the sum is taken over the finite set of coefficients. So, it is necessary to minimize the following:

$$\sum_{n=N/2, n \neq 0} |c_n - s e^{i(n\phi + \alpha)} c'_n|^2$$

Next the partial derivatives are taken with respect to s , ϕ , and α and they are

set to zero. Then letting $c_n^* c'_n = \rho_n e^{i\psi_n}$, the following three coupled equations are obtained :

$$s = \frac{\sum_n \rho_n \cos(\psi_n + n\phi + \alpha)}{\sum_n c'_n c_n^*} \quad (1)$$

$$\tan(\alpha) = -\frac{\sum_n \rho_n \sin(\psi_n + n\phi)}{\sum_n \rho_n \cos(\psi_n + n\phi)} \quad (2)$$

$$\tan(\alpha) = -\frac{\sum_n n \rho_n \sin(\psi_n + n\phi)}{\sum_n n \rho_n \cos(\psi_n + n\phi)} \quad (3)$$

Combining (2) and (3) into one equation in ϕ

$$f(\phi) = \sum_n \rho_n \sin(\psi_n + n\phi) \sum_n n \rho_n \cos(\psi_n + n\phi) \\ - \sum_n \rho_n \cos(\psi_n + n\phi) \sum_n n \rho_n \sin(\psi_n + n\phi)$$

The optimum value for ϕ is obtained by finding the roots of $f(\phi)$. Then the corresponding values of s and α can be obtained from equations (1) and (2). Since usually only the $-N/2^{\text{th}}$ to the $N/2^{\text{th}}$ components are retained, the roots of $f(\phi)$ can be obtained. However, this is very expensive to compute, especially when the unknown has to be compared to many templates.

Instead of finding the zeroes of the transcendental equations above, several investigators [PROF82,RICH74] have used simple correlation methods to determine a good starting point and rotation normalization. The correlation between two contours is usually implemented by multiplying the Fourier transforms, taking the inverse transform, and then finding the peak.

Still another [GRAN72] has used algebraic combinations of the Fourier

AD-A132 842

SHAPE RECOGNITION AND DESCRIPTION: A COMPARATIVE STUDY

2/3

(U) PURDUE UNIV LAFAYETTE IN SCHOOL OF ELECTRICAL

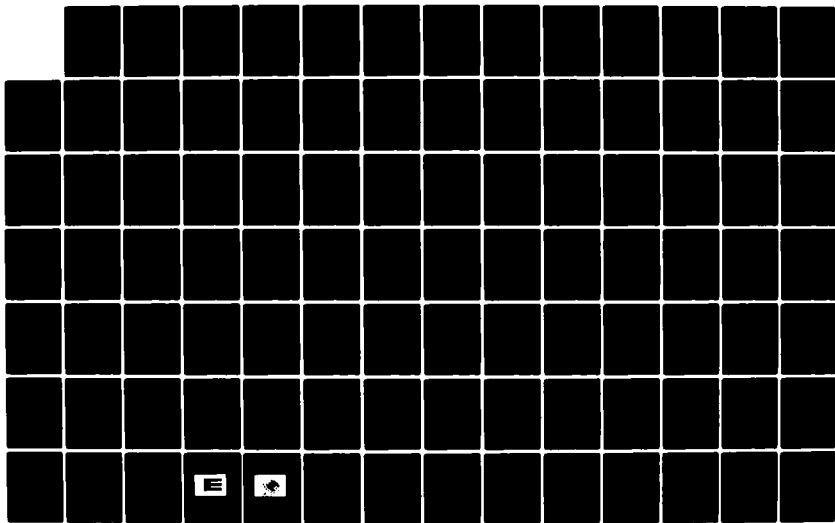
ENGINEERING T A GROGAN ET AL. JUL 83 TR-EE83-22

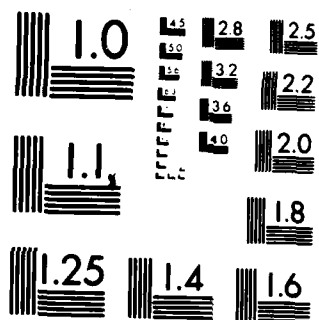
UNCLASSIFIED

ARO-18016.11-EL-A DAAG29-81-K-0088

F/G 12/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

coefficients to obtain starting point and rotation invariance.

To reduce the computational burden to normalize the coefficients, many others have suggested that one of several suboptimum methods be used. Instead of normalizing the coefficients of the unknown differently for each template, the coefficients are normalized to a "standard" orientation independent of the template.

First the object is translated to the origin by setting $c_0 = 0$. Because the fundamental or c_1 has always been observed to be the largest ($n \neq 0$), all the coefficients are scaled by dividing by $|c_1|$. Next, the shape is rotated and the starting point shifted. Most of the variations among the suboptimal methods described in the literature occur in how this rotation and starting point shift is obtained.

The simplest method is to rotate and shift so that the transformed coefficients c'_1 and c'_{-1} have zero phase [CRIM82,KUHL82]. So, the normalized coefficients c'_n are as follows:

$$c'_n = \frac{c_n}{|c_1|} e^{i(n t_0 + \alpha)}, \quad c'_0 = 0,$$

where

$$c'_1 = |c'_1| e^{i L c'_1} \quad \text{and} \quad c'_{-1} = |c'_{-1}| e^{i L c'_{-1}}.$$

This requires that t_0 and α to be

$$t_0 = -\frac{L c_1 - L c_{-1}}{2} \quad \text{and} \quad \alpha = -\frac{L c_1 + L c_{-1}}{2}.$$

This is equivalent to rotating and shifting the starting point so the the c_1 and c_{-1} components describing a fundamental ellipse has its major axis along the

x-axis.

There are two problems with this approach. The magnitude of c_{-1} may be zero. So, the phase angle is indeterminate. Second, there are two possible orientations which satisfy the zero phase condition. An additional 180° rotation and $-T/2$ starting point shift would also satisfy the zero phase condition.

Alternatives to this approach can help to mitigate these difficulties [WALL80a,WALL80b,MITC82a,MITC82b]. The following algorithm is used in this report to normalize the shape coefficients:

- 1) Set $c_0' = 0$.
- 2) Scale so that $|c_1'| = 1$.
- 3) Find the coefficient c_k that is next largest ($|k-1| \leq 5, k \neq 0, 1$). If $|k-1| > 5$, then let $k = 2$.
- 4) Rotate and shift the starting point so that $\angle c_1' = 0$ and $\angle c_k' = 0$, i.e.

$$c_n' = \frac{c_n}{|c_1|} e^{i(nt_0 + \alpha)},$$

$$\text{where } t_0 = \frac{\angle c_1 - \angle c_k}{k-1}, \alpha = -\frac{(k \angle c_1 - \angle c_k)}{k-1}.$$

The object $\gamma(t) = c_1 e^{i\frac{2\pi}{T}t} + c_k e^{i\frac{2\pi k}{T}t}$ has $|k-1|$ -fold symmetry. So, there are $|k-1|$ rotations and relative starting point shifts, multiples of $\frac{2\pi}{k-1}$ that will satisfy the zero phase condition. So, rotate and shift the starting point so that

$$\tilde{c}_n = c_n' e^{i \frac{2\pi m}{k-1} - i \frac{2\pi m}{k-1}}, \quad m = 0, 1, \dots, (|k-1| - 2).$$

to maximize the following criterion:

$$\sum_n \text{Re}\{\tilde{c}_n\} |\text{Re}\{\tilde{c}_n\}|.$$

This criterion effectively chooses the normalization that orients the contour so that the axis of one of the main lobes of the $|k-1|$ -fold fundamental shape $c_1 e^{it} + c_k e^{ikt}$ is along the positive x-axis and the starting point on the contour corresponding to that lobe is the one farthest from the origin. In order to reduce the number of rotations to perform and calculate the criterion only $|k-1| \leq 5$ is allowed. If $|k-1| > 5$, then $k = -1$ as in the first normalization method.

The correct recognition of a shape is very sensitive to this rotation and starting point normalization. In order to improve the classification accuracy, multiple sets of Fourier coefficients are used in classifying the shape. The descriptor for the best normalization above is used. In addition, if one of the multiple rotations has a value of its criterion that is within 95% of the best, this normalization is also used. To reduce the possibilities that the wrong coefficient is used in the normalization, the third largest coefficient is also used if its magnitude is within 95% of the second largest coefficient. Again, the next best normalization based on this third largest coefficient is also used if the criterion is 95% of the best normalization based on this new coefficient. So in all there can be as many as four sets of coefficients for each unknown shape. For the library features only the best normalization is used. One of the four possible normalizations is likely to match the proper template even when some noise is present.

3.5 Generic Shapes

In order to recognize a generic shape (one of a class of shapes), it is useful to know the analytic expression for the characterization. In this section are the analytic formulae for several basic shapes. Each generic shape is represented by a set of parameters and the formula for the Fourier coefficients. In order to recognize a shape, these parameters are estimated. The exact coefficients can be calculated. Then a comparison of the coefficients of the unknown and the model estimate can be made in order to make a decision [MITC81].

Other properties of the shape such as bilateral and axial symmetry can also be investigated. For instance, the next largest coefficient, $|c_k|$, after $|c_1|$ can be used to determine the fundamental shape of the object.

If $|c_k|$ is the next largest, the shape has $|k-1|$ -fold symmetry. Figure 3.12 displays a library of fundamental shapes indexed by the value of k and the ratio of the modulus of the coefficients [MOEL82]. The contours in this figure correspond to $\gamma(t) = c_1 e^{it} + c_k e^{ikt}$, i.e. inverting the Fourier descriptor having the two components c_1 and c_k . The frequency ratio is $1:k$ and the modulus ratio is $|\frac{c_1}{c_k}|$. Figure 3.13 shows four rows of the fundamental shape library displaying equal time interval sample points. The non-uniform sample spacing which occurs when there are only a finite number of coefficients can be observed.

If most of the energy is in the odd harmonic components, then the shape has (almost) axial symmetry. The test

$$t \times \sum_{n \text{ odd}} |c_n|^2 > \sum_{n \text{ even}} |c_n|^2,$$

where $0 < t < 1$ can determine if the shape has (almost) axial symmetry.

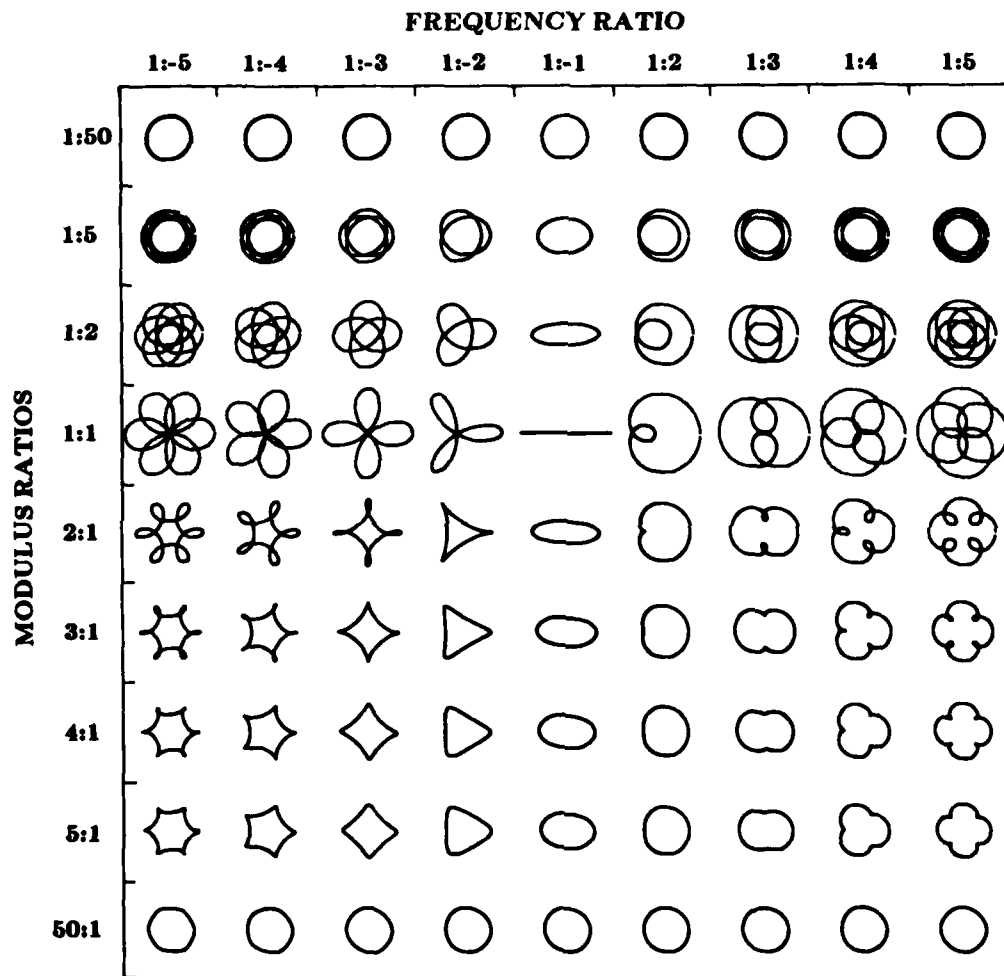


Figure 3.12 Fundamental shape library [MOEL82].

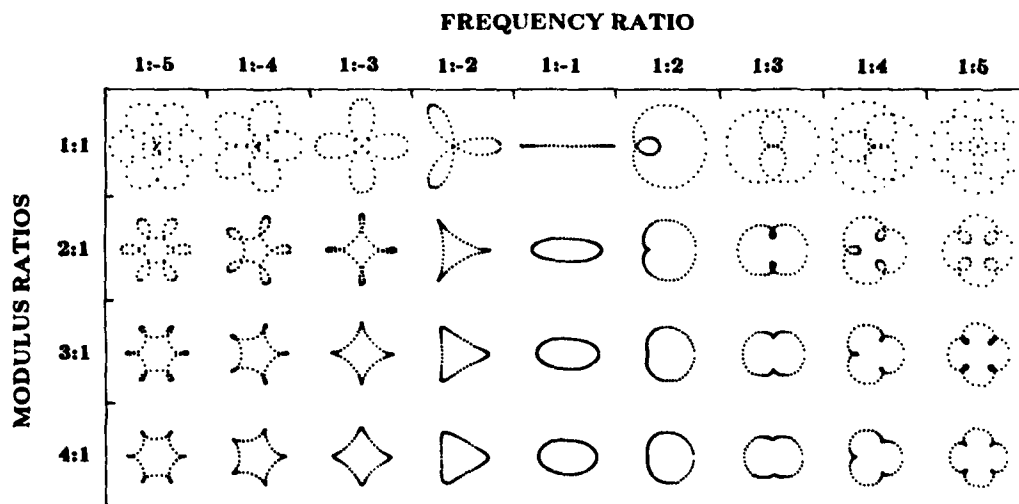


Figure 3.13 Fundamental shape library showing sample points.

The shape can be tested for (almost) bilateral symmetry (with respect to the x-axis) by determining if the components are nearly all purely real after normalization. If

$$t \times \sum_n |\operatorname{Re}\{\tilde{c}_n\}|^2 > \sum_n |\operatorname{Im}\{\tilde{c}_n\}|^2,$$

then the shape has (almost) bilateral symmetry.

Following is a collection of important generic shapes [LAWR72,YATE47] with the corresponding Fourier descriptors.

- 1) *Circle*: The algebraic equation for a circle centered at the origin is $x^2 + y^2 = r^2$. The substitutions $x = r \cos t$ and $y = r \sin t$ also satisfy the equation. So,

$$\gamma(t) = r \cos t + r \sin t = r e^{it}, \quad t \in [0, 2\pi]$$

describes a circle of radius r . It has only the component, $c_1 = r$. Unknown shapes can be tested for their "circularity" by determining the proportion of the energy in the c_1 component relative to the sum of the rest. If

$$t \times |c_1| > \sum_{n, n \neq 0, 1} |c_n|^2,$$

then the shape is approximately circular.

- 2) *Ellipse*: The algebraic expression for an ellipse centered at the origin is $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Substituting $x = a \cos t$, and $y = b \sin t$ satisfies the equation as $t \rightarrow 0$ to 2π tracing out an ellipse. The ellipse has a major axis of length "a" along the x-axis and a minor axis of length "b" ($a > b$.) This gives

$$\gamma(t) = a \cos t + i b \sin t = 1/2(a+b)e^{it} + 1/2(a-b)e^{-it}.$$

So, $c_1 = \frac{a+b}{2}$ and $c_{-1} = \frac{a-b}{2}$. This is why it was stated earlier that normalizing on c_1 and c_{-1} is the same as placing the fundamental ellipse best fitting the shape in its standard position. However, this shape is not uniformly traced since

$$v(t) = \left| \frac{d\gamma(t)}{dt} \right| = \left(a^2 \sin^2 t + b^2 \cos^2 t \right)^{1/2} \neq \text{constant}.$$

The elongation of a shape can be estimated by the ratio

$$\text{elongation} = \epsilon = \frac{|c_{-1}|}{|c_1|} = \frac{a-b}{a+b}.$$

If the ϵ is zero the object is nearly circular. If ϵ is greater than 0 the object is elongated. The value of ϵ will always be less than 1 for a simple closed contour that does not cross (traced counterclockwise).

- 3) *Rectangle*: Shown in Figure 3.14 is a rectangle in standard position having a length of a and a height of b . It's starting point is located on the positive x -axis. The rectangle can be described by the list of $\{\Delta\gamma_i\}$, $\{\Delta t_i\}$, γ_0 , and T obtained when tracing the rectangle counterclockwise.

$$\{\Delta\gamma_i\} = \{ b/2, -a, -ib, a, ib/2 \}$$

$$\{\Delta t\} = \{ b/2, a, b, a, b/2 \} \quad T = 2(a+b), \quad \gamma_0 = a/2 + i0.$$

This data is used to calculate the Fourier coefficients from the contour using the DFT given the following formula

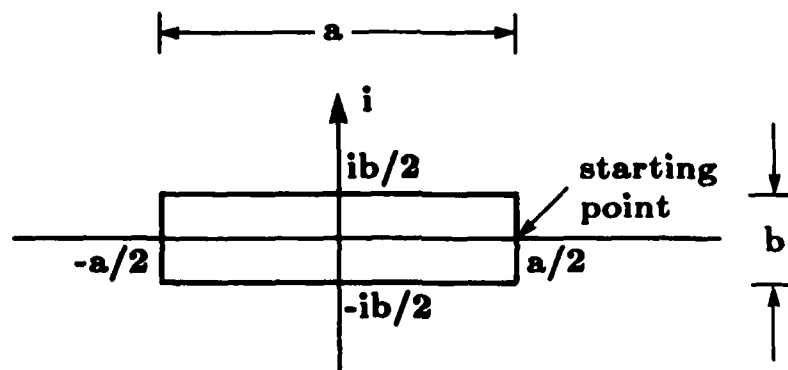


Figure 3.14 Rectangle.

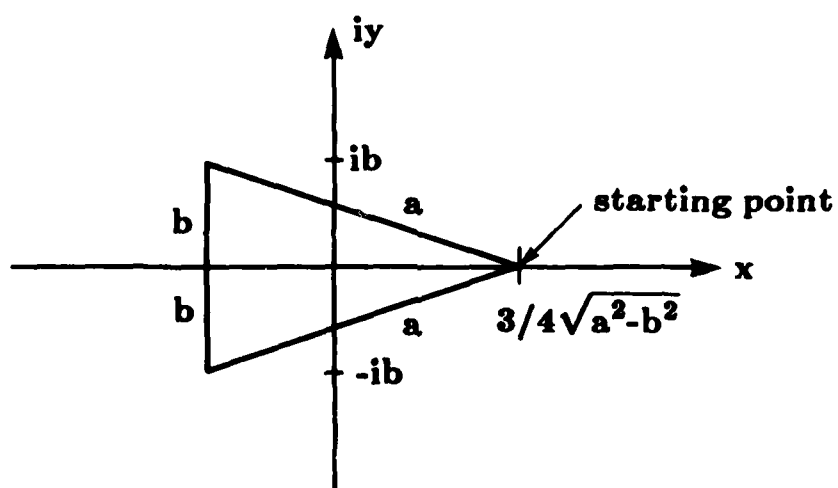


Figure 3.15 Isosceles triangle.

$$c_n = \frac{2(a+b)}{\pi^2 n^2} \sin \frac{n\pi}{2} \left\{ \cos \frac{n\pi\rho}{2} + \sin \frac{n\pi\rho}{2} \right\}$$

$$c_n = 0, (n-1) \bmod 2 \neq 0$$

$$c_0 = 0,$$

where $\rho = \frac{a}{a+b}$. Notice that the rectangle exhibits the properties of 2-fold and bilateral symmetry. The coefficients can be normalized to

$$\tilde{c}_n = \frac{\sin \frac{n\pi}{2} \left\{ \cos \frac{n\pi\rho}{2} + \sin \frac{n\pi\rho}{2} \right\}}{n^2 \left\{ \cos \frac{\pi\rho}{2} + \sin \frac{\pi\rho}{2} \right\}}$$

The n^{th} component is a function of only n and the parameter ρ .

The $\tilde{c}_1$ and $\tilde{c}_{-1}$ coefficients can be used to estimate ρ , as

$$\hat{\rho} = \frac{2}{\pi} \arctan \left(\frac{|c_1| + |c_{-1}|}{|c_1| - |c_{-1}|} \right).$$

A shape can be tested whether or not it is rectangular by computing coefficients of a rectangle based on the the estimate $\hat{\rho}$. Then decide based on the distance between the normalized unknown components and those for the rectangular model estimate. If $1/2 < \rho$ the rectangle is in standard position. If $\rho > 1/2$ then it is not and the roles of a and b should be reversed. The report by Mitchell and others [MITC81] describes the use of the above generic shape recognition procedures to classify shape in imagery to assist in aerial photo reconnaissance. The above formula for the rectangle also includes the special case of $\rho = 1/2$, corresponding to a square, when

$$c_n = \frac{2(a+b)}{\pi^2 n^2} \sin \frac{n\pi}{2} \left\{ \cos \frac{n\pi \rho}{2} + \sin \frac{n\pi \rho}{2} \right\}$$

$$c_n = 0, (n-1) \bmod 4 \neq 0.$$

- 4) *Isosceles triangle*: Shown in Figure 3.15 is an isosceles triangle whose two equal sides are of length a , and base is of length $2b$. The triangle is in standard position, and $c_0 = 0$. The list of $\{\Delta\gamma_i\}$, $\{\Delta t_i\}$, γ_0 , and T are

$$\{\Delta\gamma_i\} = \{ -\sqrt{a^2-b^2} + ib, -i2b, \sqrt{a^2-b^2} + ib \}$$

$$\{\Delta t_i\} = \{ a, 2b, a \}, T = 2(a+b), \gamma_0 = \frac{3}{4}\sqrt{a^2-b^2} + i0.$$

This gives the following formula for the coefficients:

$$c_n = \frac{a+b}{a} \frac{1}{\pi^2 n^2} \left\{ (a+b) \sin\left[\frac{\pi na}{a+b}\right] + \sqrt{(a+b)(a-b)} (1 - \cos\left[\frac{\pi na}{a+b}\right]) \right\}, n \neq 0$$

$$c_0 = 0.$$

Notice that the triangle has bilateral symmetry. For the case where $2b = a$, the isosceles triangle becomes an equilateral triangle with

$$c_n = \frac{3a}{4\pi^2 n^2} \left\{ 3\sin\left[\frac{2\pi n}{3}\right] + 2\sqrt{3}\sin^2\left[\frac{\pi n}{3}\right] \right\}.$$

$$c_n = 0, (n-1) \bmod 3 \neq 0.$$

Observe that the triangle has 3-fold symmetry.

5) *Line shape*: Shown in Figure 3.16 is the simplest line shape having

$$\{\Delta\gamma_i\} = \{-2, 2\}$$

$$\{\Delta t_i\} = \{2, 2\} \quad \gamma_0 = 1 + i0, \quad T = 4.$$

The coefficients are

$$c_n = \frac{4}{\pi^2 n^2}, \quad (n-1) \bmod 2 = 0$$

$$c_n = 0, \quad \text{otherwise.}$$

The normalized coefficients are

$$\tilde{c}_n = \frac{1}{n^2}, \quad (n-1) \bmod 2 = 0$$

$$\tilde{c}_n = 0, \quad \text{otherwise.}$$

The shape is in standard position and the starting point is on the right endpoint ($\alpha = 0$.)

6) *Thin cross*: Figure 3.17 depicts a thin cross. The list of $\{\Delta\gamma_i\}$ and $\{\Delta t_i\}$ are

$$\{\Delta\gamma_i\} = \{-1, i, -i, -1, 1, -i, i, 1\}$$

$$\{\Delta t_i\} = \{1, 1, 1, 1, 1, 1\}, \quad T = 8, \quad \gamma_0 = 1 + i0.$$

Notice that this is not a line shape. It does not retrace itself in the opposite direction in one period. The coefficients for this shape are

$$c_n = \frac{8}{\pi^2 n^2} \{ [1 - (-1)^n] - [\cos \frac{\pi n}{4} + \sin \frac{\pi n}{4}] + 2 \sin \frac{\pi n}{2} \}$$

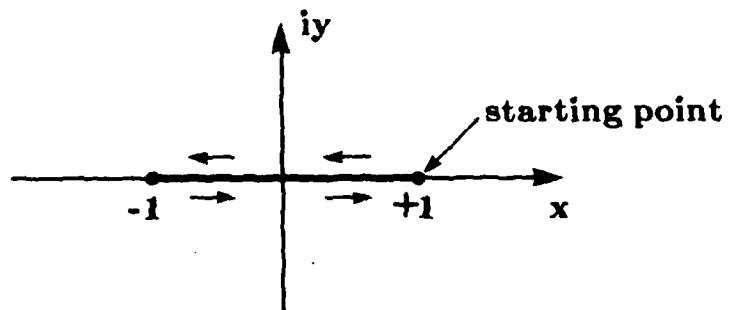


Figure 3.16 Simple line shape.

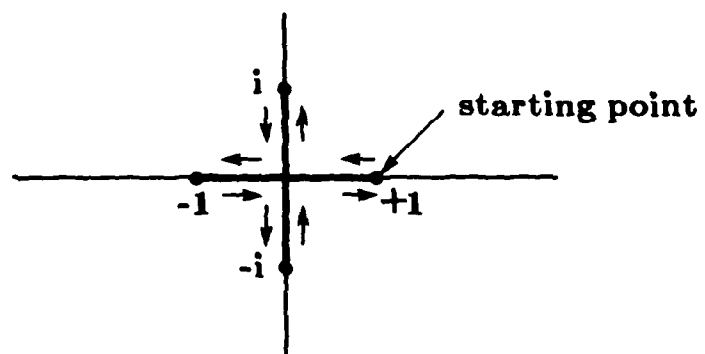


Figure 3.17 Thin cross.

$$+ [\cos \frac{3\pi n}{4} - \sin \frac{3\pi n}{4}], (n-1) \bmod 4 \neq 0$$

$c_n = 0$, otherwise.

The contour of a object can be reconstructed from its Fourier coefficients by taking the inverse Fourier transform. Figure 3.18 show an F104 contour reconstructed from finite sets of coefficients. Figure 3.19 shows the Fourier descriptor magnitudes for the contour. The original F104 contour is shown in Figure 4.2.

3.6 Experimental Results

The experiments in section 2.3 were carried out for the Fourier descriptors of the boundary. The feature vector was formed by calculating and normalizing the coefficients. After normalization, the $\tilde{c}_0$ and $\tilde{c}_1$ components did not carry any shape information and were dropped from the feature vector. The feature vector was then formed by listing the real and imaginary parts of the remaining components as

$$\underline{f} = (\text{Re}\{c_{-1}\}, \text{Im}\{c_{-1}\}, \text{Re}\{c_2\}, \text{Im}\{c_2\},$$

$$\text{Re}\{c_{-2}\}, \text{Im}\{c_{-2}\}, \dots, \text{Re}\{c_{N/2}\}, \text{Im}\{c_{N/2}\})^T.$$

This is the full feature vector. The 32 coefficients, c_{-16} to c_{15} , were used for most of the experiments. The feature vector was then reduced by retaining the 12 first elements of the feature vector after applying the principle components transformation. Now the results of each basic experiment will be presented and discussed.

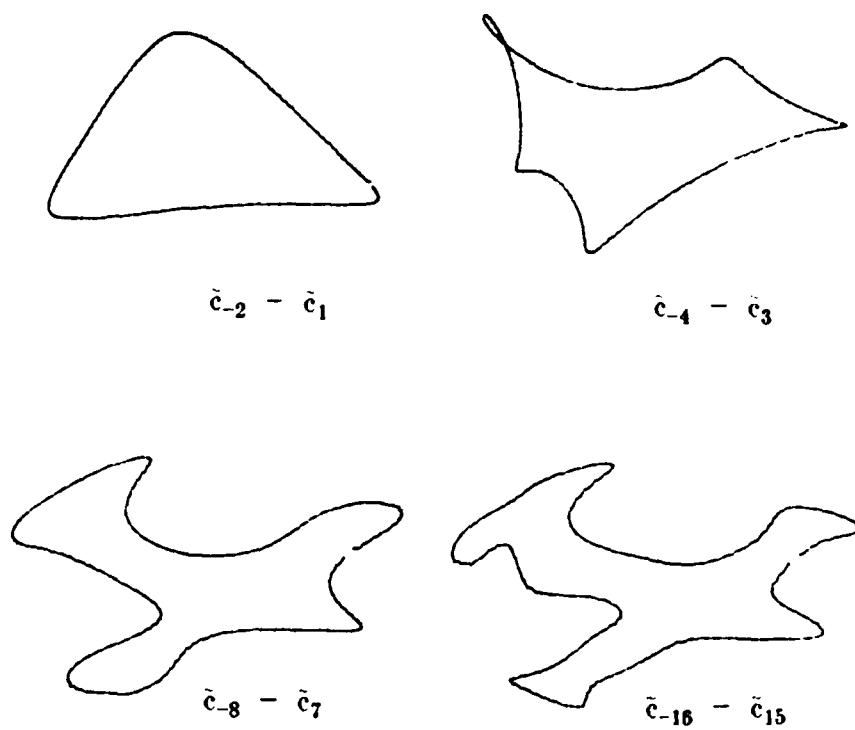


Figure 3.18 F104 contours reconstructed from the Fourier descriptors of the boundary.

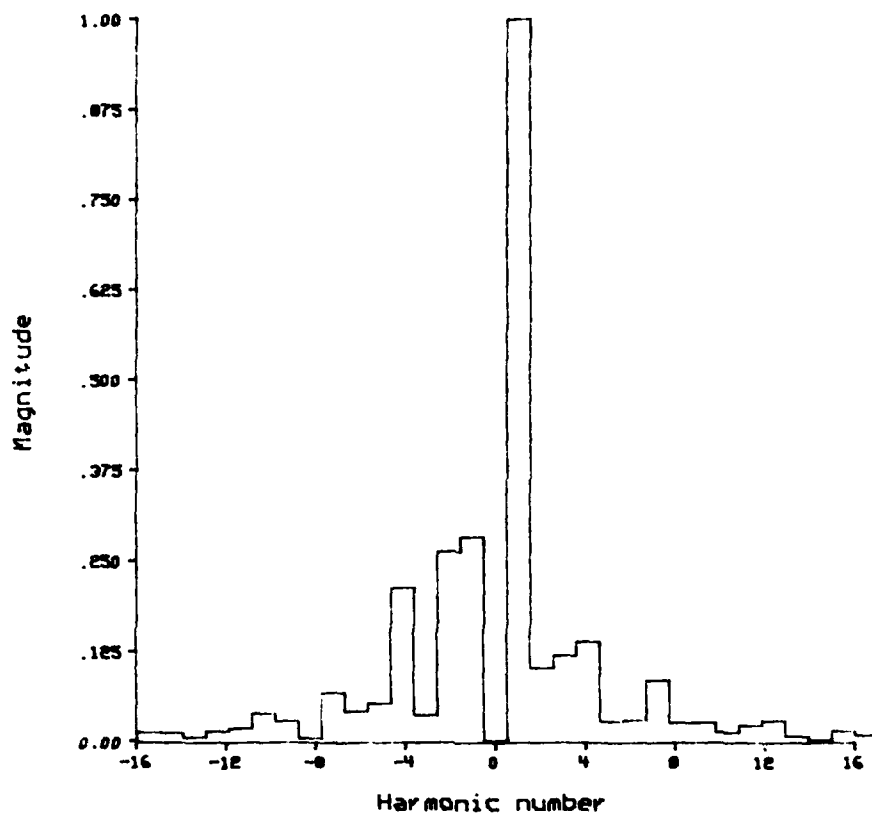


Figure 3.19 Fourier descriptors of the boundary (magnitude) for F104 contour.

The imaging resolution experiment was carried out. The classification results are presented in Table 3.1. The median x and y angle error between the correctly classified aircraft and the closest view in the library are shown in Table 3.2. The classification accuracy increases as the resolution of the unknown increases for a fixed library resolution until the resolution of the unknown and the library are the same. Then it begins to decrease. It is interesting that the accuracy decreases with increasing library resolution for a fixed unknown resolution, if the unknown is of less resolution. So, in order to obtain the best possible accuracy, the unknown and library resolution should be as close as possible. If the library must be at a fixed resolution, then the 128x128 library has the highest average classification accuracy for these shapes. This is contrary to the idea that the highest library resolution must be the best. The median angle errors follow the same trend. The y angle errors are in general smaller than those for the x angle errors. This is probably due to the actual aircraft shapes being used.

The results for the feature vector experiment are provided in Table 3.3. The upper portion of the table lists the results using the full feature vectors. The lower portion contains results for when the principle components transformation was used and the 12 features corresponding to the largest eigenvalues are retained. As the number of features is increased, the classification accuracy increases. From the sudden increase in classification accuracy, and the sharp decrease in median angle error between using 4 and 12 features indicates that most of the necessary information for these shapes are contained in the first 8 Fourier coefficients.

Table 3.4 contains the results of the imaging noise experiment. The feature vector consisted of 32 Fourier coefficients reduced to 12 real numbers.

Table 3.1 FDS image resolution experimental results: classification accuracy (%).

Unknown Resolution	Prototype Resolution					
	16	32	64	128	256	512
16	54.33	59.33	49.33	45.33	43.0	41.33
32	64.0	86.33	77.0	72.0	70.33	70.0
64	60.33	88.67	91.0	91.33	87.33	87.0
128	55.67	81.33	91.0	93.67	92.0	92.0
256	50.0	79.33	90.33	92.67	92.33	93.0
512	50.33	78.0	89.0	93.0	92.67	92.67

Table 3.2 FDS image resolution experimental results: median angle error $\Delta\phi_x, \Delta\phi_y$ (°).

Unknown Resolution	Prototype Resolution					
	16	32	64	128	256	512
16	11,7	12,8	10,8	10,9	12,9	10,8
32	10,7	7,5	8,5	7,6	7,6	7,5
64	10,7	7,5	7,4	6,4	7,4	7,4
128	11,7	7,5	7,4	4,4	4,4	4,4
256	11,7	7,5	7,4	4,4	4,4	4,4
512	14,8	7,5	7,5	6,4	6,4	4,4

Table 3.3 FDS feature vector experimental results.

Number of Fourier Coeffs.	Number of reduced Features	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
4	4	42.0	20,12
8	12	80.67	7,4
16	28	93.67	4,4
32	60	94.0	4,4
64	124	94.0	4,4
16	12	93.0	4,4
32	12	92.0	4,4
64	12	92.0	4,4

Table 3.4 FDS imaging noise experimental results.

Library Resolution	Unknown Resolution	SNR (dB)	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
256	64	3	26.67	25,20
		6	51.67	9,8
		10	74.0	7,6
		20	86.0	7,4
		∞	87.33	7,4
256	128	3	40.67	14,11
		6	58.67	9,7
		10	87.33	7,5
		20	92.0	4,4
		∞	92.0	4,4

The performance improves consistently with increasing signal-to-noise ratio.

Table 3.5 presents the results of performing the library sampling experiment. This data indicates that even more views might be necessary to obtain very high classification accuracy.

In Table 3.6 are the results for classifying partial shapes. The method performs better than might be expected for a global method up until more than 10% of the contour is chopped. Then the performance drops off quickly.

3.7 Conclusions

The Fourier descriptors of the boundary perform very well under varying circumstances. If more sophisticated classification procedures with interpolation were used then the results would even be better. The Fourier series is well understood. This greatly facilitates their usefulness for shape description and recognition.

Table 3.5 FDS library sampling experimental results.

Library Views	Library Resolution	Unknown Resolution	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
143	256	64	87.33	7,4
49			73.67	9,6
9			52.33	25,13
143	256	128	92.0	4,4
49			79.67	9,6
9			54.67	25,13
143	256	256	92.33	4,4
49			79.67	9,6
9			55.0	25,13

Table 3.6 FDS partial shape experimental results.

Percent Contour Chopped	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
0	92.0	4,4
10	83.67	7,5
20	58.0	9,7
30	30.33	34,15
40	22.0	57,27
50	12.67	62,33

CHAPTER 4

WALSH POINTS OF THE BOUNDARY

4.1 Introduction

The Walsh functions have often been used as the basis for functional approximation because the piecewise constant form leads to efficient computation [GIAR83]. Also, since the basis functions are non-zero over the entire interval of definition, this is a complete global shape analysis method. Instead of the transform coefficients, this method uses the basis functions to formulate the computation of *Walsh points* for the shape feature set. The Walsh points are a collection of points, $x_k + iy_k$, $k = 0, 1, 2, \dots, 2^m - 1$, that have a property derived from the fact that the Walsh functions were used in formulating their calculation. The projections $x(t)$ and $y(t)$ are approximated to within a prespecified degree of accuracy by a piecewise constant function, a truncated Walsh series.

The Walsh functions, $W_i(t)$, are products of the Rademacher functions, $r_i(t)$. The Walsh functions for a complete orthogonal basis set complete among square integrable functions on the interval $[0, 1]$. The Walsh functions are defined as

$$W_0(t) = 1.$$

$$W_n(t) = r_{n_1+1}(t) r_{n_2+1}(t) \cdots r_{n_p+1}(t),$$

where $n \geq 1$ is expressed in binary as

$$n = 2^{n_1} + 2^{n_2} + \dots + 2^{n_p}$$

and the integers n_i are ordered such that

$$n_1 < n_2 < \dots < n_p.$$

The Rademacher functions are

$$r_0(t) = 1$$

$$r_1(t) = \begin{cases} +1, & t \in [0, 1/2) \\ -1, & t \in [1/2, 1) \end{cases}$$

$$r_1(t+1) = r_1(t)$$

$$r_{k+1}(t) = r_1(2^k t), \quad k = 0, 1, 2, \dots$$

In the following only the x projection of the boundary calculation will be discussed. The same results follow for the y projection.

The Walsh series for $x(t)$ is

$$x(t) = \sum_{n=0}^{\infty} \alpha_n W_n(t/T)$$

where

$$\alpha_i = \int_0^T x(t) W_i(t/T) dt = T \int_0^1 x(Tt) W_i(t) dt$$

T = total arc length along $\gamma(t) = x(t) + iy(t)$.

Each α_i is made up of the area under the portions of $x(t)$ added and subtracted together. In approximating $x(t)$, only a finite number of terms in the Walsh series are retained - say the first 2^m . If the series is truncated to a finite

number of terms, there is an approximation error. It would be expected that as $N = 2^m$ increases, the closer $x_{2^m}(t)$ approximates $x(t)$. This error will be discussed later.

Let

$$x_{2^m}(t) = \sum_{k=0}^{2^m-1} \alpha_k W_k(t/T).$$

Since the function $x_{2^m}(t)$ (or $y_{2^m}(t)$) describes a piecewise constant function, there are 2^m discrete points, $x_k + iy_k$, that are the sum of the heights of the Walsh functions over each interval, $[\frac{k}{2^m} T, \frac{k+1}{2^m} T)$, $k = 0, 1, \dots, 2^m-1$, which approximate the x (or y) projection of $\gamma(t) = x(t) + iy(t)$.

It was mentioned earlier that the Walsh functions allow for an efficient computation of the function $x_{2^m}(t)$ (or $y_{2^m}(t)$.) The i^{th} Walsh function can be written as

$$W_i(t) = \sum_{k=0}^{2^m-1} \gamma_{ik} \chi_{[\frac{k}{2^m}, \frac{k+1}{2^m})}(t),$$

where

$$\gamma_{ik} = W_i(t) = \pm 1, \quad t \in [\frac{k}{2^m}, \frac{k+1}{2^m}).$$

The matrix of γ_{ik} , $H = [H_{ik}] = [\gamma_{ik}]$, is the Hadamard matrix with 2^m rows and columns.

The i^{th} coefficient of the Walsh series is

$$\alpha_i = T \int_0^1 x(Tt) W_i(t) dt = T \sum_{k=0}^{2^m-1} \gamma_{ik} \int_0^1 x(Tt) \chi_{[\frac{k}{2^m}, \frac{k+1}{2^m})} dt$$

$$\alpha_i = \sum_{k=0}^{2^m-1} \gamma_{ik} \int_{\frac{kT}{2^m}}^{\frac{(k+1)T}{2^m}} x(t) dt$$

$$\alpha_i = \sum_{k=0}^{2^m-1} \gamma_{ik} A_{xk}, \quad (1)$$

where

$$A_{xk} = T \int_{\frac{kT}{2^m}}^{\frac{(k+1)T}{2^m}} x(Tt) dt = \int_{\frac{kT}{2^m}}^{\frac{(k+1)T}{2^m}} x(t) dt.$$

A_{xk} is the area under $x(t)$ over the basic interval $[\frac{kT}{2^m}, \frac{(k+1)T}{2^m}]$. By letting

$$\bar{A}_x = \begin{pmatrix} A_{0x} \\ A_{1x} \\ \vdots \\ A_{2^m-1x} \end{pmatrix}$$

be the *Area* vector and

$$\bar{\alpha}_x = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{2^m-1} \end{pmatrix}$$

be the *sequency* vector, we can write

$$\bar{\alpha}_x = H \bar{A}_x.$$

The Truncated Walsh series is

$$x_{2^m}(t) = \sum_{i=0}^{2^m-1} \alpha_i W_i(t/T)$$

$$\begin{aligned}
&= \sum_{i=0}^{2^m-1} \alpha_i \left(\sum_{k=0}^{2^m-1} \gamma_{ik} \chi_{\left[\frac{k}{2^m}, \frac{k+1}{2^m}\right)}(t/T) \right) \\
&= \sum_{k=0}^{2^m-1} \left(\sum_{i=0}^{2^m-1} \alpha_i \gamma_{ik} \right) \chi_{\left[\frac{k}{2^m}, \frac{k+1}{2^m}\right)}(t/T) \\
&= \sum_{k=0}^{2^m-1} c_k \chi_{\left[\frac{k}{2^m}, \frac{k+1}{2^m}\right)}(t/T),
\end{aligned}$$

where

$$c_k = \sum_{i=0}^{2^m-1} \gamma_{ik} \alpha_i. \quad (2)$$

The constants c_k can be found by first forming the vector $\bar{x}_{2^m}$ as

$$\bar{x}_{2^m} = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{2^m-1} \end{pmatrix}$$

and noticing that as in equation (1) that equation (2) can be written as

$$\bar{x}_{2^m} = H^T \bar{a}_x$$

where H^T is the transpose of the Hadamard matrix, H .

Since H is symmetric $H = H^T$ and

$$\bar{x}_{2^m} = H^2 \bar{A}_x.$$

But because the Walsh functions are orthogonal, it follows that $H^2 = 2^m$.

Hence,

$$\bar{x}_{2^m} = 2^m \bar{A}_x.$$

So, the truncated Walsh series gives a piecewise constant function whose

heights are proportional to the area under the projection on each basic interval, i.e.

$$x_{2^m}(t) = 2^m \sum_{k=0}^{2^m-1} A_{kx} \chi_{\left[\frac{kT}{2^m}, \frac{(k+1)T}{2^m}\right)}(t).$$

4.2 Truncation Error and ϵ -Net

In the previous section the expansion for the truncated Walsh series was derived. The error in the expansion between $x(t)$ and $x_{2^m}(t)$ using the L_∞ norm is

$$\|x(t) - x_{2^m}(t)\|_\infty = \sup_{0 \leq t < 1} |x(t) - x_{2^m}(t)| \leq \frac{M}{2^{m+1}}$$

where M is the maximum slope in the x projection. This follows from the more general result [GIAR78], that if $x(t)$ satisfies a Holder condition of order α , $0 < \alpha \leq 1$, with constant M , that is

$$|x(t+h) - x(t)| \leq M |h|^\alpha, \quad M \geq 0$$

for t and h real, then

$$\|x(t) - x_{2^m}(t)\|_\infty \leq \frac{M 2^{\alpha-1}}{(2^\alpha - 1)(2^\alpha)^{m+1}}.$$

In particular if $\alpha = 1$ (which is the case for polygonal curves), we have

$$\|x(t) - x_{2^m}(t)\|_\infty \leq \frac{M}{2^{m+1}}.$$

An ϵ -net ($\epsilon > 0$) for an arbitrary bounded, closed curve $\gamma \ni [0, 1] \rightarrow \mathbb{R} \times \mathbb{R}$ is a collection of a finite number of points $W \ni [0, 1] \rightarrow \mathbb{R} \times \mathbb{R}$ where

$$W = \bigcup_{i=1}^L \{(x_i, y_i) | 1\} = \bigcup_{i=1}^L \{w_i\}.$$

The union of the closed balls, $\max_{i=1,L} [|x-x_i|, |y-y_i|] \leq \epsilon$, covers the set of points (x,y) of γ . That is, every point of γ must be within a distance ϵ from some point in W .

For a discrete image of an object silhouette, $x(t)$ (or $y(t)$) is the projection of a Freeman encoded boundary. Hence $x(t)$ is a piecewise linear function that satisfies Lipchitz condition with M equal to the maximum absolute slope of the linear elements. The important fact is that the Walsh points in the vector $\bar{x}_{2^m}$ form an ϵ -net for the curve γ . Figure 4.1 is a sketch of an ϵ -neighborhood. If we take any point p on the curve γ and draw a closed ball (a square with sides of length 2ϵ centered on p), then since W is an ϵ -net for γ there must exist a point $w_i \in W$. However, if W' is also an ϵ -net for the curve γ , then there must also be a point $w' \in W'$ in the same ϵ -neighborhood. So, the farthest Euclidean distance that w can be from w' is $2\sqrt{2}\epsilon$. This is key to a recognition procedure, because the Walsh points for a shape will be located within this degree of accuracy.

The curve γ can be approximated to a prescribed error ϵ or it can be approximated by a fixed number of points resulting in error bounded by ϵ . Given a prescribed error ϵ , the number of Walsh points will be 2^m where

$$m = \max(n, k)$$

where

$$n = \lceil \log_2 \left(\frac{M}{\epsilon} \right) \rceil - 1 \quad \text{and} \quad k = \lceil \log_2 \left(\frac{L}{\epsilon} \right) \rceil - 1$$

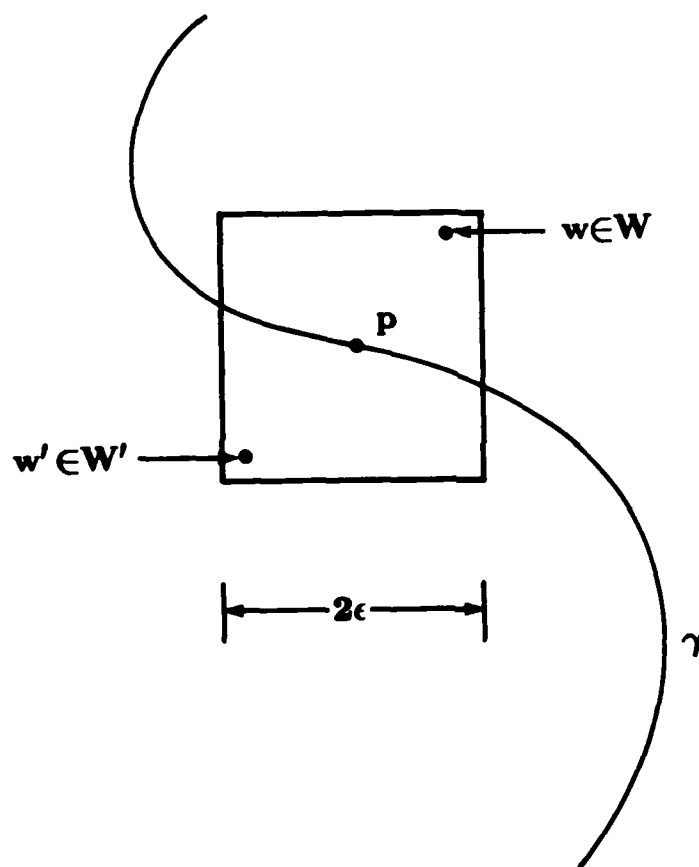


Figure 4.1 Sketch of an ϵ -neighborhood.

$$|x(t) - x_{2^n}(t)| \leq \frac{M}{2^{n+1}} \quad \text{and} \quad |y(t) - y_{2^k}(t)| \leq \frac{L}{2^{k+1}}$$

M and L are the Lipchitz constants for x and y respectively. Alternatively, the number of Walsh points can be fixed at $N = 2^m$ and the accuracy is determined by

$$\epsilon = \max\{\epsilon_x, \epsilon_y\}, \text{ where}$$

$$\epsilon_x = \frac{M}{2^{m+1}} \quad \text{and} \quad \epsilon_y = \frac{L}{2^{m+1}}.$$

In the recognition experiments, a fixed number of Walsh points were used.

4.3 Normalization

In order to compare the Walsh points feature vector of an unknown to a library prototype, normalization with respect to translation, scale, rotation, and shift in starting point must be accomplished.

The traced contour from the image grid is stored using Freeman chain code links. When the shape features are to be extracted, the chain code is converted to a complex vector $\vec{\gamma} = \vec{x} + i\vec{y}$. The increments in the arc length for the piecewise linear segments connecting the grid points are calculated. The c_0 , c_1 , and c_{-1} Fourier coefficients of the boundary are calculated to provide the global information necessary for translation, scale, and rotation. The translation is $-c_0$. The scale factor $s = |c_1|$. The rotation angle is

$$\psi = -\frac{\angle c_1 + \angle c_{-1}}{2}. \text{ The complex vector } \vec{\gamma} = \begin{pmatrix} \gamma_0 \\ \vdots \\ \gamma_k \end{pmatrix} \text{ is normalized to } \vec{\gamma}', \text{ where}$$

the i^{th} component is

$$\gamma'_i = \frac{1}{s} (\gamma_i - c_0) e^{i\psi}, \text{ and } T' = T/s.$$

This normalization moves the center of the shape to the origin. It scales and rotates the shape so that the major axis corresponding to the fundamental ellipse is along the x-axis and has a length of 2. The starting point is moved by finding the k^{th} point, γ_k , which is closest to one of the two points where the fundamental ellipse crosses the x-axis. An additional 180° rotation is applied if necessary to place γ'_k in the right half plane. So, the normalized Walsh points form the vector

$$\vec{\gamma}' = \begin{pmatrix} \gamma'_k \\ \vdots \\ \gamma'_{K-1} \\ \gamma'_0 \\ \vdots \\ \gamma'_{k-1} \end{pmatrix} = \underset{k}{\text{shift}} \vec{\gamma}.$$

or $\vec{\gamma}' = \underset{k}{\text{shift}} \vec{\gamma} \cdot e^{i\frac{\pi}{2}}$. Then the Walsh points are calculated by computing the Area vector and multiplying by 2^m . Figure 4.2 depicts an F104 contour. Figures 4.3 to 4.6 show the Walsh points with their corresponding ϵ -neighborhoods after normalization. Figures 4.7 to 4.11 show their corresponding x and y projections.

4.4 Experimental Results

The feature vector used in the experiments was formed by simply listing the Fourier points real and imaginary parts in order, i.e.

$$\underline{f} = (x_0, y_0, x_1, y_1, \dots, x_{K-1}, y_{K-1})^T$$

For most of the experiments the full feature vector consisted of 32 Walsh

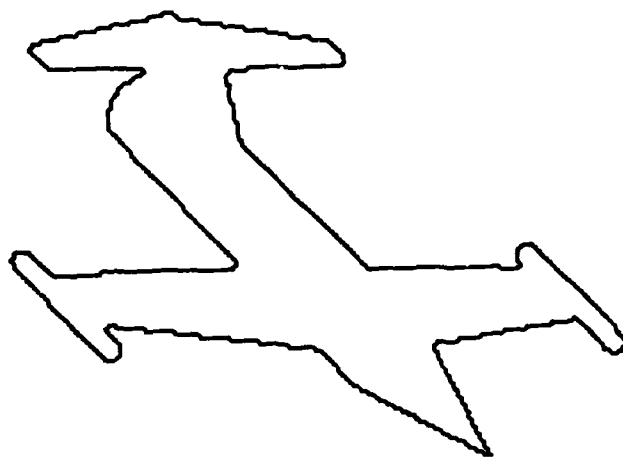


Figure 4.2 Original F104 contour.

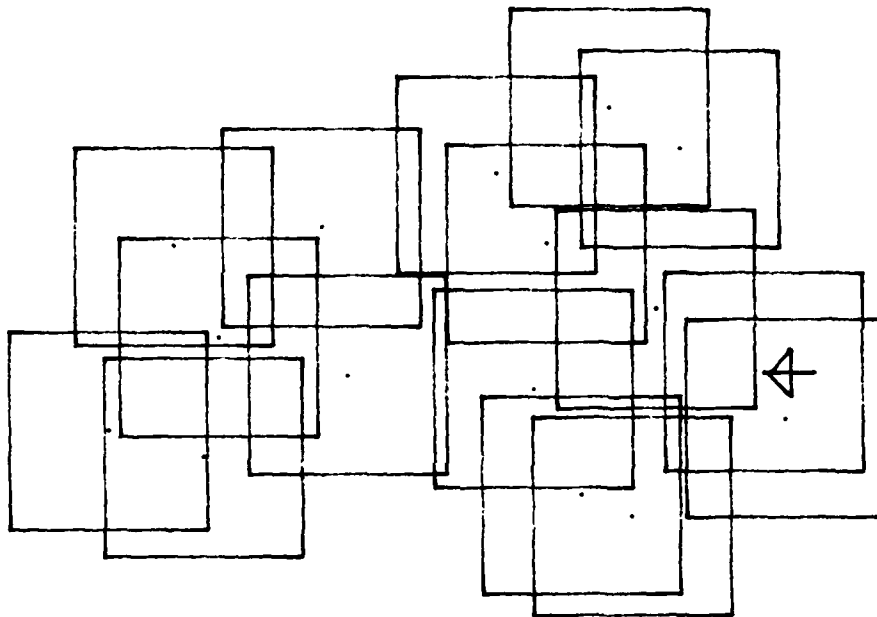


Figure 4.3 Walsh points with corresponding ϵ -neighborhood after normalization, $2^m = 16$ points.

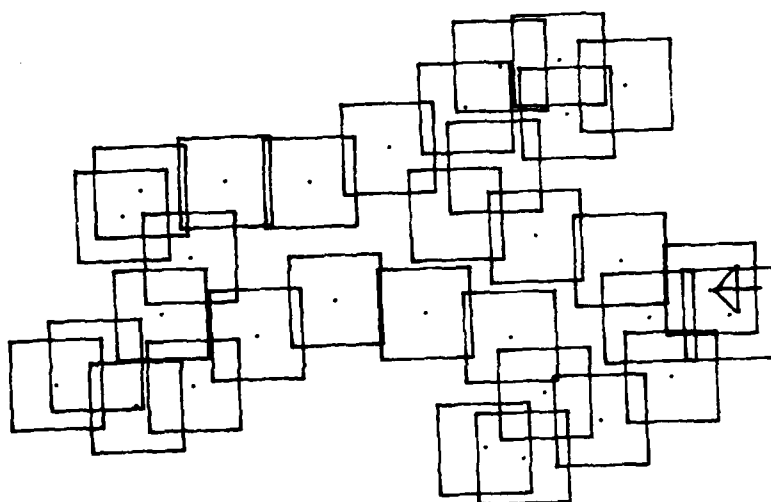


Figure 4.4 Walsh points with corresponding ϵ -neighborhood after normalization, $2^m = 32$ points.

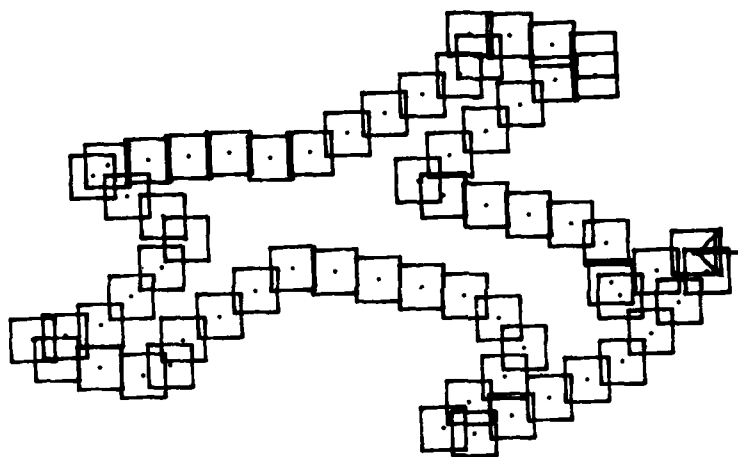


Figure 4.5 Walsh points with corresponding ϵ -neighborhood after normalization, $2^m = 64$ points.

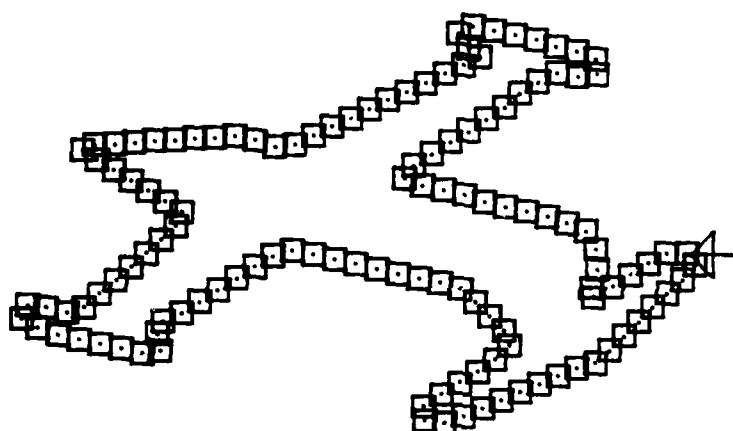


Figure 4.6 Walsh points with corresponding ϵ -neighborhood after normalization, $2^m = 128$ points.

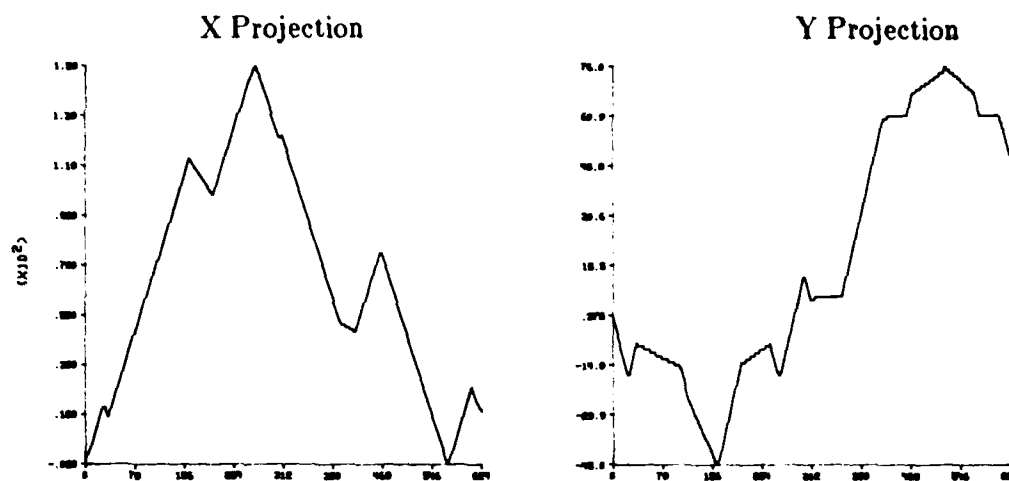


Figure 4.7 Original F104 contour projections (624 points).

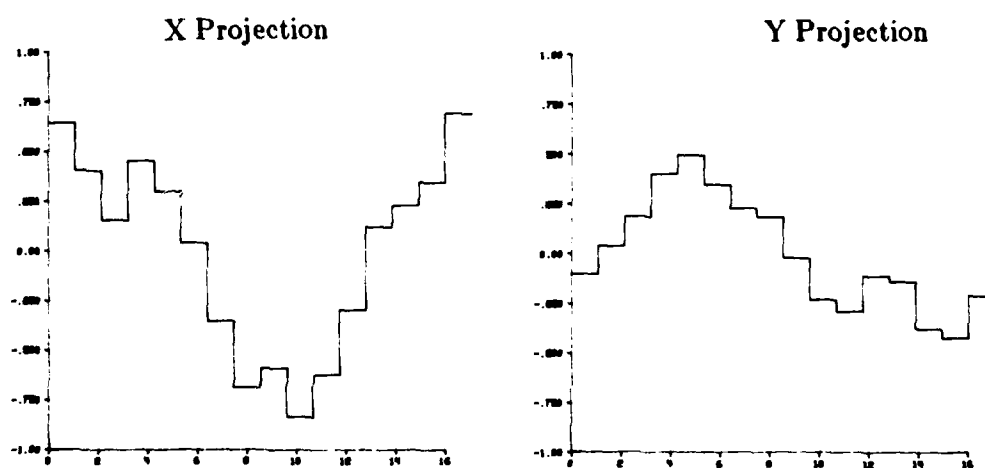


Figure 4.8 Normalized Walsh points projections ($2^m = 16$).

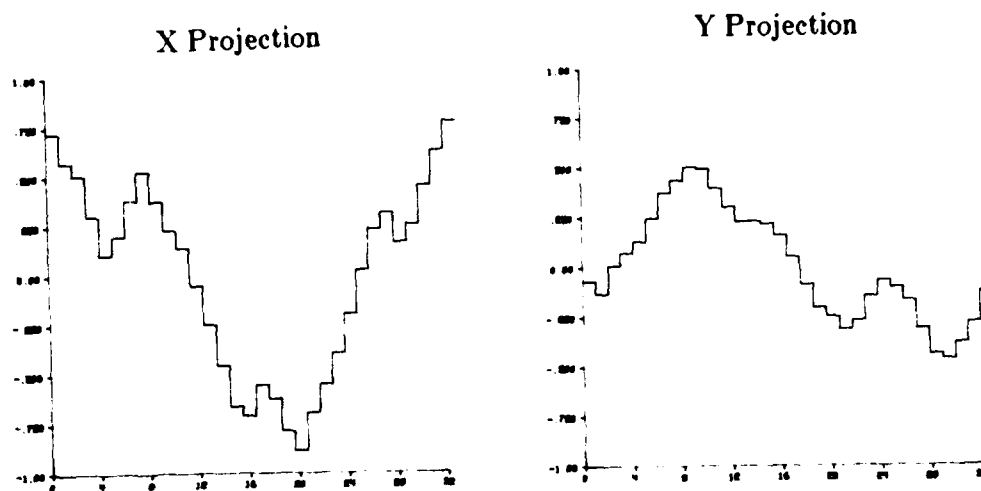


Figure 4.9 Normalized Walsh points projections ($2^m = 32$).

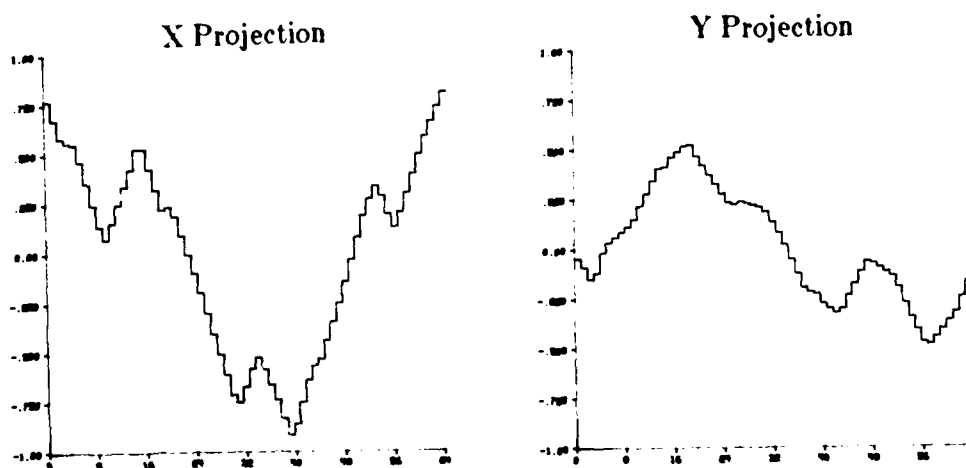


Figure 4.10 Normalized Walsh points projections ($2^m = 64$).

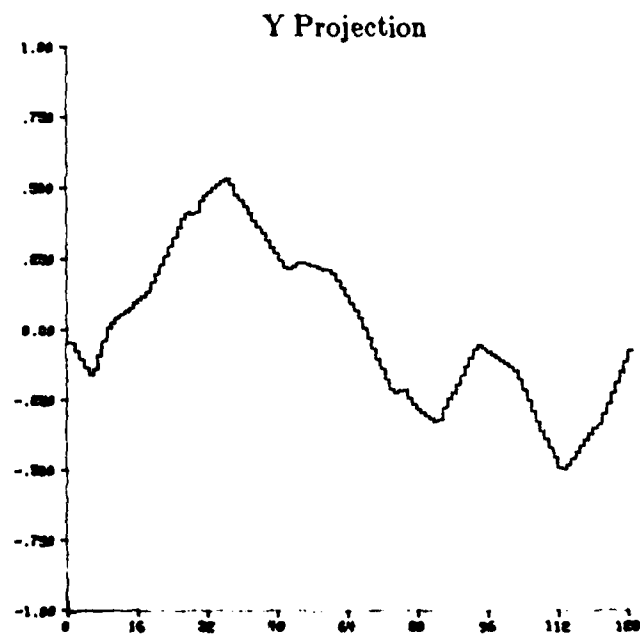
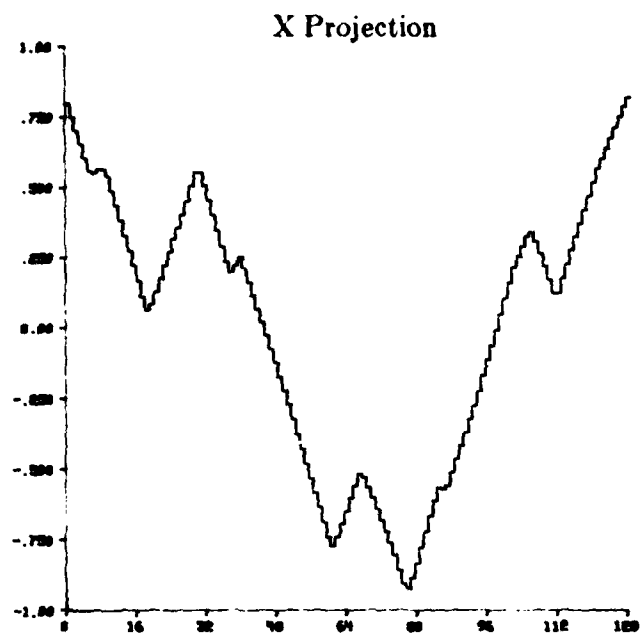


Figure 4.11 Normalized Walsh points projections ($2^m = 128$).

points. In most cases the feature is transformed using principle components and then truncated to 12 real numbers.

Table 4.1 and Table 4.2 list the results for the image resolution experiment. The feature consisted of 32 Walsh points reduced to the 12 real numbers. The best results are obtained when the prototype and unknown image resolutions match.

Table 4.3 presents the results of the feature vector experiment. The results shown in the top portion of the table are for the full feature vector. The lower portion corresponds to when the feature was reduced to 12 real numbers. The performance improves with increasing number of Walsh points.

Table 4.4 lists the results of the imaging noise experiment. The feature set is the 12 reduced real numbers. The performance improves with increasing signal-to-noise ratio.

Table 4.5 lists the results of the library sampling experiment. The performance improves with increasing numbers of library views.

The partial shape experiment results are contained in Table 4.6. The reduced vector was used here also. The method can obtain a correct classification $3/4$ of the time even with 10% of the contour chopped.

4.5 Conclusions

The Walsh points perform well. The speed at which the Walsh points can be computed make them a good competitor. The major problem is with the normalization. Normalization is cumbersome for the Walsh points. In fact, normalization is actually carried out with use of the Fourier descriptors of the boundary. The use of interpolation techniques would surely improve the

Table 4.1 WAL image resolution experimental results: classification accuracy (%).

Unknown Resolution	Prototype Resolution					
	16	32	64	128	256	512
16	45.0	44.67	40.0	38.33	37.33	35.33
32	52.33	73.33	71.33	62.67	58.67	57.0
64	49.0	72.33	83.0	80.0	75.67	72.0
128	46.67	67.67	82.0	86.33	83.33	82.67
256	44.33	66.0	81.33	85.67	85.33	83.67
512	44.0	64.33	79.0	84.67	83.33	85.0

Table 4.2 WAL image resolution experimental results: median angle error $\Delta\phi_x, \Delta\phi_y$ (°).

Unknown Resolution	Prototype Resolution					
	16	32	64	128	256	512
16	13,7	17,8	20,8	17,8	25,11	25,12
32	10,7	8,5	10,5	9,6	9,7	9,7
64	12,7	7,6	7,5	7,5	7,4	7,5
128	11,7	7,6	7,5	4,4	6,4	6,4
256	9,7	8,5	7,5	5,5	6,4	6,4
512	9,8	9,6	7,5	6,4	7,4	6,4

Table 4.3 WAL feature vector experimental results.

Number of Walsh Points	Number of reduced Features	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
16	32	88.0	6,4
32	64	90.67	6,4
64	128	91.67	6,4
32	12	83.0	6,4
64	12	83.33	6,4
128	12	84.0	6,4

Table 4.4 WAL imaging noise experimental results.

Library Resolution	Unknown Resolution	SNR (dB)	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
256	64	3	28.0	25,17
		6	43.33	14,9
		10	61.67	9,6
		20	71.33	7,5
		∞	75.67	7,4
256	128	3	39.0	21,12
		6	52.0	12,8
		10	72.0	7,5
		20	83.33	6,4
		∞	83.33	6,4

Table 4.5 WAL library sampling experimental results.

Library Views	Library Resolution	Unknown Resolution	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
143	256	64	75.67	7,4
49			64.33	12,6
9			41.0	32,13
143	256	128	83.33	6,4
49			70.67	9,6
9			40.0	26,12
143	256	256	85.33	6,4
49			70.33	9,5
9			41.67	32,13

Table 4.6 WAL partial shape experimental results.

Percent Contour Chopped	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
0	83.33	6,4
10	76.0	7,5
20	56.33	12,7
30	30.0	27,12
40	21.33	63,27
50	17.0	91,33

results. The Walsh points however do not provide the easy access to generic shape property measurements, such as for symmetry.

CHAPTER 5

CUMULATIVE ANGULAR DEVIANT FOURIER DESCRIPTOR

5.1 Introduction

This shape analysis method also uses the Fourier expansion in order to exploit its time shift properties. Instead of approximating the boundary function, the approach here is to use a function derived from the angle function [ZAHN72]. Physiological and psychological investigations of the human visual system suggest that curvature is very important in classifying shape [ATTN66].

The boundary is traced clockwise as a function of arc length. At each point the instantaneous direction of the velocity vector with respect to an orthogonal coordinate system defines the angle function, $\theta(\ell)$, as a function of arc length. Since this function repeats after one complete circuit, it is periodic and can be expanded into a complex Fourier series. But first it would be advantageous to normalize this function. The initial starting angle $\theta(0) = \delta_0$ is first subtracted to define

$$\phi(\ell) = \theta(\ell) - \delta_0, \quad \ell \in [0, L],$$

where L is the total arc length. The time scale is normalized to make a function that has a period of 2π . Also, if the contour is simple and closed, then $\phi(0) = 0$ and $\phi(L) = -2\pi$. This linearly decreasing angle is subtracted to obtain the cumulative angular deviant function $\phi^{\square}(t)$,

$$\phi^{\blacksquare}(t) = \phi\left(\frac{Lt}{2\pi}\right) + t, \quad t \in [0, 2\pi].$$

The complex Fourier series then is

$$\phi^{\blacksquare}(t) = \sum_{n=-\infty}^{+\infty} d_n e^{int}$$

where

$$d_n = \frac{1}{2\pi} \int_0^{2\pi} \phi^{\blacksquare}(t) e^{-int} dt.$$

But since $\phi^{\blacksquare}$ is purely real, then $d_n = d_{-n}^*$. So, it is only necessary to know d_n for $n \geq 0$. The series can therefore also be written as

$$\phi^{\blacksquare}(t) = \mu_0 + \sum_{n=1}^{\infty} A_n \cos(nt + \alpha_n),$$

where

$$\mu_0 = d_0$$

$$A_n = 2|d_n| \quad \text{and} \quad \alpha_n = \angle d_n.$$

5.2 Properties

The following is a list of properties of the Fourier series of the cumulative angular deviant [ZAHN72] as they pertain to shape geometry.

- 1) $\phi^{\blacksquare}$ is invariant under translation, rotation, and dilation (scale) of the contour γ .

- 2) For all simple closed curves

$$\phi^{\blacksquare}(2\pi) = \phi^{\blacksquare}(0) .$$

- 3) Since it has been required that $\phi^{\blacksquare}(0) = 0$, then

$$d_0 = -2 \sum_{n=1}^{\infty} \operatorname{Re}\{d_n\}$$

or

$$\mu_0 = - \sum_{n=1}^{\infty} A_n \cos(\alpha_n) .$$

- 4) *Reconstruction:* Given a contour described by $\theta(\ell)$ and starting point $\gamma(0)$, then

$$\gamma(\ell) = \gamma(0) + \int_0^{\ell} e^{i\theta(\lambda)} d\lambda .$$

- 5) Any function $\theta'(\ell) \equiv \theta(\ell) \bmod 2\pi$ describes the same contour.
- 6) A curve is completely represented by $\gamma(0)$, δ_0 , L , and $\phi^{\blacksquare}(t)$.
- 7) *Closure:* The function $\gamma(\ell)$ is a closed curve, if one of the following is true:
- A_1 is a zero of the first order Bessel function $J_1(x)$ and $A_n = 0$ for $n \geq 0$.
 - $A_n = 0$ for all $n \bmod k \neq 0$, where $k \geq 2$.
- 8) *Starting point:* A clockwise shift in the starting point of Δl implies that

$$d_{n'} = d_n e^{-i \frac{2\pi n \Delta l}{L}}$$

or

$$A_n' = A_n$$

and

$$\alpha_n' = \alpha_n - n \Delta\alpha, \quad \Delta\alpha = \frac{2\pi\Delta\ell}{L}.$$

9) *N-fold symmetry*: $\gamma(\ell)$ has N-fold symmetry, if and only if

$$d_n = 0, \quad \text{for } n \bmod N \neq 0.$$

10) *Axial symmetry*: If the curve γ has axial symmetry, then

$$d_n = 0, \quad \text{for } n \text{ odd.}$$

(This is different than that given in [ZAHN72]).

11) *Reflection*: For $\tilde{\gamma}$ the reflection of γ ,

$$\tilde{d}_n = -d_n^*, \quad n \neq 0$$

$$d_0 = -d_0 - \Delta\phi_0,$$

where $\Delta\phi_0$ is the change in angle at the starting point if γ is polygonal there (i.e. has a jump in angle of size $\Delta\phi_0$.)

5.3 Calculation

If the curve is a polygon, as in the case for the chain coded boundaries, the Fourier descriptors of the cumulative angular function can be calculated directly. Assuming the polygon has K vertices, let $\Delta\ell_i$ be the length of the side from the $(i-1)^{\text{th}}$ to the i^{th} vertex. The change in angular direction at the i^{th} vertex is $\Delta\phi_i$. The total arc length is then $L = \sum_{i=1}^K \Delta\ell_i$. Therefore,

$$\phi(\ell) = \sum_{i=1}^K \Delta\phi_i, \text{ for } \sum_{i=1}^k \Delta\ell_i \leq \ell < \sum_{i=1}^{k+1} \Delta\ell_i.$$

and

$$\phi(\ell) = 0, \text{ for } 0 \leq \ell < \ell_1.$$

With these definitions (see Figure 5.1), the Fourier coefficients are

$$d_0 = -\pi - \frac{1}{L} \sum_{k=1}^K \ell_k \Delta\phi_k$$

$$d_n = \frac{i}{n\pi} \sum_{k=1}^K \Delta\phi_k e^{i \frac{2\pi n \ell_k}{L}}, \quad n \neq 0,$$

where

$$\ell_k = \sum_{i=1}^K \Delta\ell_i.$$

The parameters $\Delta\ell_i$ and $\Delta\phi_i$ can be obtained directly from the chain codes, a_i , where $\Delta\ell_i = \sqrt{2}$ or 1 and

$$\Delta\phi_i = f(a_i - a_{i-1})$$

where

$$\begin{aligned} \{ (i, f(i)) : & (-7, \frac{\pi}{4}), (-6, \frac{\pi}{2}), (-5, \frac{3\pi}{2}), (-4, \pi), (-3, -\frac{3\pi}{2}), (-2, -\frac{\pi}{2}), \\ & (-1, -\frac{\pi}{4}), (0, 0), (1, \frac{\pi}{4}), (2, \frac{\pi}{2}), (3, \frac{3\pi}{2}), (4, -\pi), (5, -\frac{3\pi}{2}), (6, -\frac{\pi}{2}), (7, -\frac{\pi}{4}) \}. \end{aligned}$$

For a contour having a long chain code, this can be a lengthy calculation of order $N \cdot K$. This burden can be reduced by resampling $\Delta\ell_i$ and $\Delta\phi_i$ to $K' < K$. This will introduce distortion which is non-linear with respect to the

contour.

Another alternative is to resample the (x_i, y_i) sequence to a power of two. Compute the cumulative angular deviant function as

$$\phi_i = \tan^{-1} \left[\frac{y_i - y_{i-1}}{x_i - x_{i-1}} \right] - \tan^{-1} \left[\frac{y_i - y_0}{x_i - x_0} \right] + \frac{2\pi}{K'} i$$

Then use the FFT, retaining the positive frequencies only. The resampling and discrete approximation used to compute the angle function introduce distortion. Also, the FFT is periodic, so there is aliasing if the resampling does not effectively lowpass the contour sequence. Another source of distortion is that now the samples are not necessarily uniformly distributed along the contour.

In the experiments discussed later, the direct method of calculation was employed.

5.4 Normalization

The cumulative angular deviant is already normalized for translation, rotation, and dilation. This leaves only a starting point shift to normalize. The original authors [ZAHN72] used algebraic combinations of the phases of the coefficients to obtain starting point shift invariance. It is difficult however to determine if this transformation has retained all the necessary shape information.

The method used in this report normalizes the Fourier descriptors of the cumulative angular deviant, d_n , by employing the time shift property in a manner similar to that used in the Fourier descriptors of the boundary function normalization. The positive components having the first and second largest

moduli are found, $|d_{k1}|$ and $|d_{k2}|$, respectively. The starting point is shifted so that the angle of the $k1^{\text{th}}$ component has zero phase, i.e.

$$\tilde{d}_n' = d_n e^{-in \frac{\angle d_{k1}}{k1}}.$$

For the unknowns, if the $t \times |d_{k1}| \leq |d_{k2}|$, where $t = 0.95$, then a second feature vector is computed using the phase of d_{k2} , i.e.

$$\tilde{d}_n = d_n e^{-in \frac{\angle d_{k2}}{k2}}.$$

This was done to prevent noise from causing the wrong component to be chosen for the normalization. In the recognition experiments, both feature vectors are compared to all the library feature vectors. The unknown is labeled according to the smallest distance for both of the feature vectors. Since property 3) states that the d_0 coefficient is redundant, it is dropped from the feature vector.

5.5 Generic Shapes

Following are a few generic shapes with their Fourier descriptors of their cumulative angular deviant.

- 1) *Rectangle*: For a rectangle traced clockwise with the starting point on the upper right hand corner,

$$\{\Delta\phi_i\} = \left\{ \frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2} \right\}$$

$$\{\Delta\ell_i\} = \{b, a, b, a\}, \quad L = 2(a + b).$$

In that case the coefficients are

$$d_n = \frac{1}{n} \left[1 + (-1)^n \right] \cos\left(\frac{\pi n b}{2(a+b)}\right) e^{i \frac{\pi n b}{2(a+b)}} e^{i\pi/2}, \quad n \neq 0$$

or

$$d_{2n} = \frac{1}{n} \cos\left(\frac{\pi n b}{a+b}\right) e^{i \frac{\pi n b}{a+b}}, \quad d_{2n+1} = 0, \quad n = 1, 2, \dots$$

The normalized coefficients $\tilde{d}_n$ are

$$\tilde{d}_n = \frac{1}{n} \cos\left(\frac{\pi n b}{a+b}\right) e^{i \frac{\pi}{2}(n+1)}, \quad \tilde{d}_{2n+1} = 0, \quad n = 1, 2, \dots$$

In the special case of a square

$$d_{4n} = \frac{i}{2n}, \quad d_{4n+j} = 0, \quad j = 1, 2, 3, \quad n = 1, 2, \dots$$

2) *Isosceles triangle*: For an isosceles triangle traced clockwise

$$\{\Delta\phi_i\} = \{\cos^{-1}\left(\frac{-b}{a}\right), 2\cos^{-1}\left(\frac{b}{a}\right), \cos^{-1}\left(\frac{-b}{a}\right)\}$$

$$\{\Delta\theta_i\} = \{2b, a, a\}, \quad L = 2(a+b)$$

The Fourier coefficients are

$$d_n = \frac{2i}{n\pi} e^{i \frac{\pi n b}{a+b}} \left\{ \left(\pi - \cos^{-1}\left(\frac{b}{a}\right) \right) \cdot \cos\left(\frac{\pi n b}{a+b}\right) + \cos^{-1}\left(\frac{b}{a}\right) \cdot (-1)^n \right\}.$$

In the special case where the triangle is equilateral

$$d_{3n} = \frac{2i}{3n}, \quad d_{3n+j} = 0, \quad j = 1, 2, \quad n = 1, 2, \dots$$

3) *Circle*: A circle traced clockwise has an angle function $\theta(\ell) = -\frac{2\pi}{L}$. Hence

$$\phi^{\square} \equiv 0. \rightarrow d_n = 0, \forall n.$$

The Fourier descriptor magnitudes of the cumulative angular deviant for an F104 contour are shown in Figure 5.2. The contours reconstructed from increasing numbers of coefficients are presented in Figure 5.3.

5.6 Experimental Results

The feature vector for the Fourier descriptors of the cumulative angular deviant were formed by dropping the DC (d_0) component and listing the real and imaginary parts of the positive frequency components, i.e.

$$\underline{f} = (\text{Re}\{d_1\}, \text{Im}\{d_1\}, \dots, \text{Re}\{d_{n-1}\}, \text{Im}\{d_{n-1}\})^T.$$

In most cases the full feature vector was formed from the first 32 Fourier coefficients. Then the principle components transformation was used to reduce the feature to 12 real numbers.

The image resolution experiment was performed using the feature vector reduced using the principle components transformation to 12 real numbers. The results are provided in Table 5.1 and Table 5.2. The best results are obtained when the prototype image resolution matches that of the unknown.

In Table 5.3 are the results obtained for the feature vector experiment. It is not clear why the classification accuracy should decrease with an increasing number of features.

Table 5.4 provides the results for the imaging noise experiment. In general, the performance increases with increasing signal-to-noise ratio. There is a large increase in classification accuracy between the 10 and 20 dB signal-

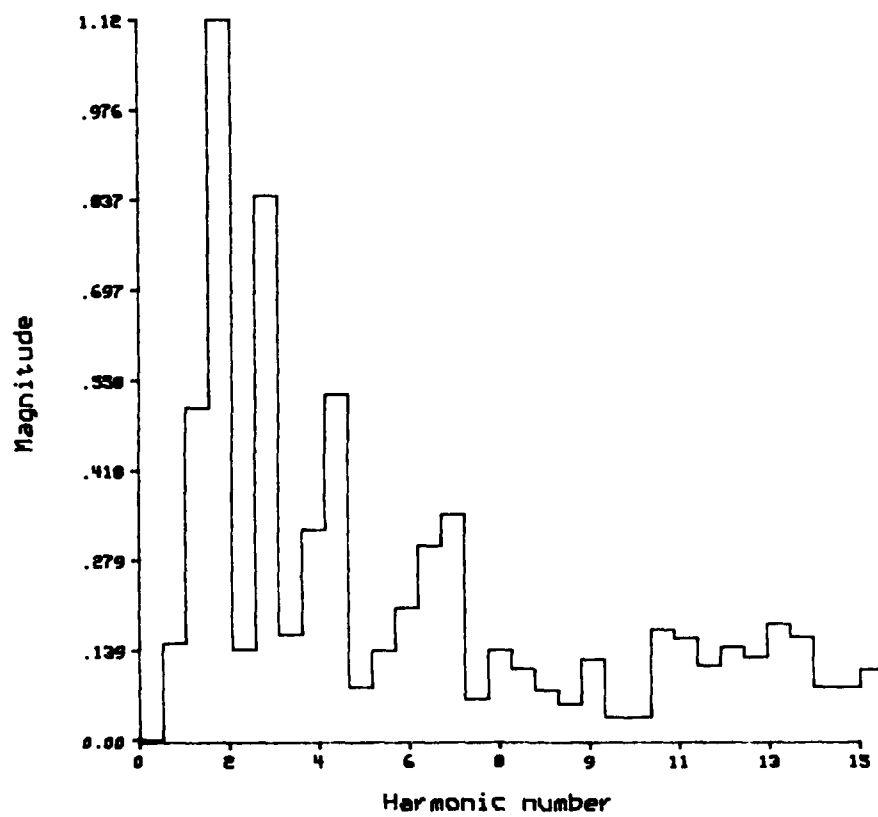


Figure 5.2 Cumulative angular deviant Fourier descriptors (magnitude) for F104 contour.

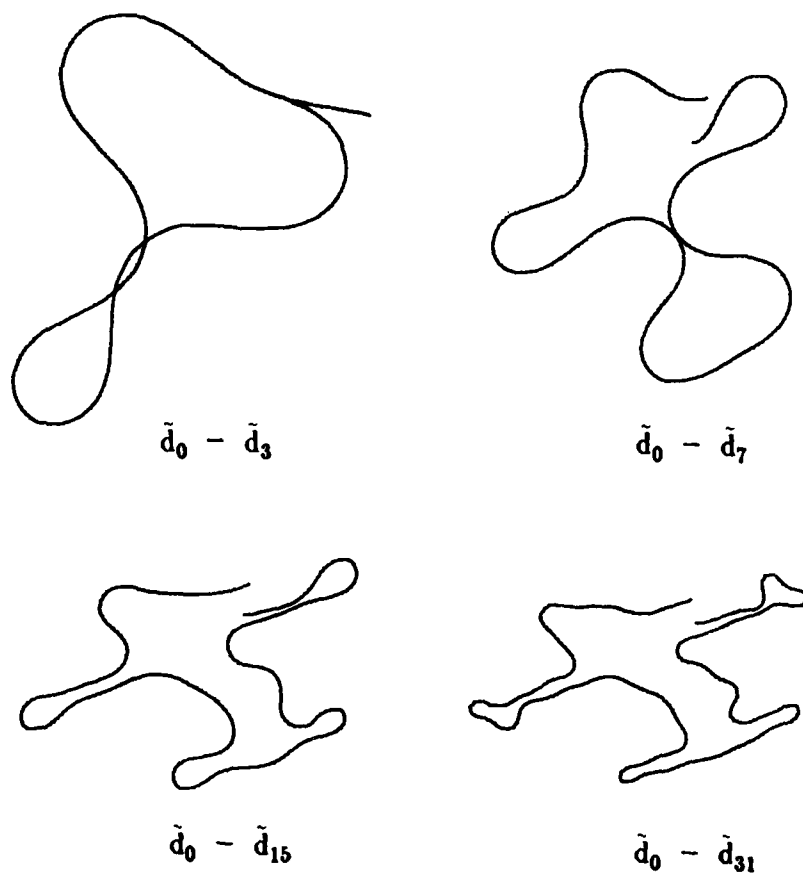


Figure 5.3 F104 contours reconstructed from the cumulative angular deviant Fourier descriptors.

Table 5.1 CAD image resolution experimental results: classification accuracy (%).

Unknown Resolution	Prototype Resolution					
	16	32	64	128	256	512
16	34.67	34.0	26.33	25.0	25.33	31.0
32	31.0	39.33	40.33	39.67	38.0	37.0
64	39.0	43.0	47.67	44.67	45.33	44.33
128	36.67	47.67	52.33	52.0	52.67	51.67
256	34.67	43.67	50.67	52.67	53.33	54.67
512	33.0	41.0	49.67	52.0	52.0	54.33

Table 5.2 CAD image resolution experimental results: median angle error $\Delta\phi_x, \Delta\phi_y$ (°).

Unknown Resolution	Prototype Resolution					
	16	32	64	128	256	512
16	22,15	30,17	54,20	75,21	78,21	67,21
32	32,15	18,9	22,12	43,13	41,12	36,12
64	41,15	25,12	23,9	21,9	32,9	32,12
128	36,15	24,10	17,9	16,9	16,9	18,10
256	34,17	34,10	17,9	17,9	17,9	20,9
512	38,20	32,10	18,10	16,9	20,9	17,9

Table 5.3 CAD feature vector experimental results.

Number of Fourier Coeffs.	Number of reduced Features	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
4	6	48.67	22,13
8	14	58.67	20,10
16	30	56.67	17,9
32	62	53.67	17,8
64	126	49.67	21,9
16	12	53.0	16,9
32	12	52.67	16,9
64	12	53.33	16,9

Table 5.4 CAD imaging noise experimental results.

Library Resolution	Unknown Resolution	SNR (dB)	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
256	64	3	15.67	92,54
		6	19.0	54,26
		10	23.33	54,26
		20	46.67	32,14
		∞	45.33	32,9
256	128	3	17.0	91,52
		6	20.33	85,52
		10	22.33	63,35
		20	50.33	30,9
		∞	52.67	16,9

to-noise ratios. There seems to be a threshold effect occurring.

Table 5.5 provides the results of the library sampling experiment. Classification accuracy does not increase dramatically with an increasing number of views as would be expected.

The partial shape experiment was also performed. The results are listed in Table 5.6. The performance quickly degrades as the contour is chopped.

5.7 Conclusions

The performance of the cumulative angular deviant Fourier descriptors is only moderate to even poor. Also the need to take the derivative and calculate the inverse tangent even before a Fourier transform is taken increases the computation complexity. It is also more difficult to obtain insight into the geometric properties of a shape because of the nonlinearity of the transformation. It is difficult to understand how the high classification accuracy results reported in the literature were obtained with this method in view of the facts as presented here.

Table 5.5 CAD library sampling experimental results.

Library Views	Library Resolution	Unknown Resolution	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
143	256	64	45.33	32,9
49			42.0	27,9
9			31.67	50,18
143	256	128	52.67	16,9
49			50.33	21,9
9			34.33	56,18
143	256	256	53.33	17,9
49			49.67	22,9
9			34.67	47,17

Table 5.6 CAD partial shape experimental results.

Percent Contour Chopped	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
0	52.67	16,9
10	34.67	28,16
20	26.0	30,11
30	26.0	32,23
40	25.33	68,27
50	19.0	90,40

CHAPTER 6

MOMENTS OF THE SILHOUETTE

6.1 Introduction

An internal spatial shape analysis method will now be introduced [HU62,TEAG80,DUDA77]. Let the image be a two dimensional function, $f(x, y)$. If $f(x, y)$ is piecewise continuous and has bounded support, then the sequence, $\{M_{pq}\}$, of moments describes $f(x, y)$, where

$$M_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy .$$

These moments are often called the conventional or raw moments. The moment M_{pq} is said to be of order $n = p + q$. This characterization is the projection of $f(x, y)$ on the basis set, the monomials, $x^p y^q$. The monomials are complete basis set for the functions of the class described above. The set however is not orthogonal.

The shapes dealt with here are silhouette. So, let the set of points contained in the object be O . Then

$$f(x, y) = \begin{cases} 1, & (x, y) \in O \\ 0, & (x, y) \notin O \end{cases} .$$

Since the monomials are non-zero almost everywhere, this is a global shape analysis method.

6.2 Properties

Listed below are a number of properties of the moment set for the function $f(x, y)$.

- 1) *Translation*: Let M_{pq} be the moment of order $n = p + q$ for $f(x, y)$. Translating $f(x, y)$ by the shift (a, b) , then

$$f'(x, y) = f(x - a, y - b) \rightarrow \mu_{pq} = M'_{pq} = \sum_{r=0}^p \sum_{s=0}^q \binom{p}{r} \binom{q}{s} a^{p-r} b^{q-s} M_{rs}.$$

When (a, b) is chosen such that $\mu_{00} = 0$, the set $\{\mu_{pq}\}$ is called the central moments.

Proof:

$$\begin{aligned} M'_{pq} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x - a, y - b) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + a)^p (y + b)^q f(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{r=0}^p \binom{p}{r} a^{p-r} x^r \sum_{s=0}^q \binom{q}{s} b^{q-s} y^s f(x, y) dx dy. \end{aligned}$$

Ω

- 2) *Dilation (scale)*: If the extent of the object, O , is increased by λ , then

$$f'(x, y) = f(x/\lambda, y/\lambda) \rightarrow M'_{pq} = \lambda^{p+q+2} M_{pq}.$$

Proof:

$$M'_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x/\lambda, y/\lambda) dx dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\lambda x)^p (\lambda y)^q f(x, y) \lambda dx \lambda dy .$$

Ω

3) *Rotation:* If an object is rotated through an angle θ about the origin, then

$$f'(x, y) = f(x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta)$$

$$\rightarrow M'_{pq} = \sum_{r=0}^p \sum_{s=0}^q \binom{p}{r} \binom{q}{s} (-1)^{q-s} (\cos \theta)^{p-r+s} (\sin \theta)^{q+r-s} M_{p+q-r-s, r+s} .$$

Proof:

$$\begin{aligned} M'_{pq} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x \cos \theta + y \sin \theta)^p (y \cos \theta - x \sin \theta)^q f(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{r=0}^p \binom{p}{r} (x \cos \theta)^{p-r} (y \sin \theta)^r \\ &\quad \cdot \sum_{s=0}^q \binom{q}{s} (y \cos \theta)^s (x \sin \theta)^{q-s} (-1)^{q-s} f(x, y) dx dy . \end{aligned}$$

4) *Reflection:* If an object is reflected about the y-axis, then

$$f'(x, y) = f(-x, y) \rightarrow M'_{pq} = (-1)^p M_{pq} .$$

Similarly, for a reflection about the x-axis,

$$f'(x, y) = f(x, -y) \rightarrow M'_{pq} = (-1)^q M_{pq} .$$

Proof:

$$\begin{aligned}
 M'_{pq} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(-x, y) dx dy \\
 &= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (-x)^p y^q f(x, y) dx dy \\
 &= (-1)^p \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy
 \end{aligned}$$

Ω

5) *Symmetry:* If an object is symmetric about the x-axis, then

$$f(x, y) = f(-x, y) \rightarrow M_{pq} = 0, \text{ for } p \text{ odd.}$$

Similarly, if the object is symmetric about the y-axis

$$f(x, y) = f(x, -y) \rightarrow M_{pq} = 0, \text{ for } q \text{ odd.}$$

Proof:

$$\begin{aligned}
 M_{pq} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(-x, y) dx dy \\
 &= (-1)^p M_{pq}
 \end{aligned}$$

$$M_{pq} [1 - (-1)^p] = 0.$$

Ω

- 6) *Combined moments*: If there are two objects O and P such that $O \cup P = \emptyset$, then the moments of the union of the two sets is just the sum of the corresponding moments, i.e.

$$M_{pq}^{O \cup P} = M_{pq}^O + M_{pq}^P.$$

In general, if $h(x,y) = f(x,y) + g(x,y)$, then

$$M_{pq}^h = M_{pq}^f + M_{pq}^g.$$

This property can make moments useful in combining the disconnected regions of a shape when noise causes segmentation errors. Also, if there were sufficient storage, a two dimensional shape could have its components stored separately. Then this method could be used for partial shape recognition by combining the moments of the components until a good match is made.

6.3 Other Moments Sets

The theory of orthogonal polynomials provides a host of complete sets of basis functions. In general, however, they do not have all the simple relationships for all the transformations mentioned early. Others [TEAG80, REEV81a, HU62] have used these alternate characterizations to exploit some particular property. Some of these alternatives are now introduced.

- 1) *Legendre moments*: The Legendre polynomials are a complete orthogonal basis set on the interval $[-1, 1]$. The n^{th} order Legendre polynomial is

$$P_n(x) = \sum_{j=0}^n c_{nj} x^j = \sum_{k=0}^{n/2} (-1)^k \frac{(2n-2k)!}{2^n k!(n-k)!(n-2k)!} x^{n-2k}.$$

The Legendre moment of order $m + n$ is

$$\lambda_{mn} = \frac{(2m+1)(2n+1)}{4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P_m(x) P_n(y) f(x,y) dx dy .$$

The function $f(x,y)$ can be approximated by a continuous function which is a truncated series

$$f(x,y) \simeq \sum_{m=0}^N \lambda_{N-m,m} P_{N-m}(x) P_m(y) .$$

- 2) *Zernike moments*: The Zernike polynomials, $V_{n\ell}(x,y)$, are complex valued functions orthogonal on the unit circle $x^2 + y^2 = 1$. They are defined as

$$V_{n\ell}(x,y) = R_{n\ell}(r) e^{i\ell\theta} ,$$

where n is the order, $n \geq \ell \geq 0$ and $n-\ell$ is even. The real valued polynomial $R_{n\ell}(r)$ is defined as

$$R_{n\ell}(r) = \sum_{s=0}^{\frac{n-\ell}{2}} (-1)^s \frac{(n-s)!}{s! (\frac{n+\ell}{2}-s)! (\frac{n-\ell}{2}-s)!} r^{n-2s} .$$

The Zernike moments, $A_{n\ell}$, of order n is defined as

$$A_{n\ell} = \frac{n+1}{\pi} \int_{-\pi}^{\pi} \int_0^{\infty} [V_{n\ell}(r,\theta)]^* g(r,\theta) r dr d\theta ,$$

where

$$g(r,\theta) = f(r \cos \theta, r \sin \theta) .$$

A rotation of the object by an angle α , implies

$$A_{n\ell}' = A_{n\ell} e^{i\ell\alpha} .$$

A translation of the object however is difficult.

The Zernike moments are related to the central moments. The radial polynomials have the form

$$R_{n\ell}(r) = \sum_{k=1}^n B_{n\ell k} r^k.$$

Then the Zernike moments are

$$A_{n\ell} = \frac{n+1}{\pi} \sum_{k=\ell}^n \sum_{j=0}^q \sum_{m=0}^{\ell} (-i)^m \begin{pmatrix} q \\ j \end{pmatrix} \begin{pmatrix} \ell \\ m \end{pmatrix} B_{n\ell k} \\ \times \mu_{k-2j+m, 2j+\ell-m},$$

where $q = (k - \ell)/2$.

If $f'(x, y)$ is the reflection of $f(x, y)$ about a line through the origin at an angle of θ with respect to the y-axis then

$$(A_{n\ell})' = (A_{n\ell})^* e^{-i2\ell\theta}.$$

3) *Rotational moments*: The rotational moments [REEV81a] are complex valued and defined as

$$F_{n\ell} = \int_{-\pi}^{\pi} \int_0^{\infty} r^n e^{i\ell\theta} g(r, \theta) r dr d\theta,$$

where n is the order and $n \geq \ell \geq 0$ and $n - \ell$ is even,

$$g(r, \theta) = f(r \cos \theta, r \sin \theta), \text{ for } r = \sqrt{x^2 + y^2} \text{ and } \theta = \tan^{-1}(\frac{y}{x}).$$

The rotational moments can be obtained from the conventional moments by

$$F_{n\ell} = \sum_{j=0}^{\frac{n-\ell}{2}} \sum_{k=0}^{\ell} (-i)^k \begin{pmatrix} \frac{n-\ell}{2} \\ j \end{pmatrix} \begin{pmatrix} \ell \\ k \end{pmatrix} M_{n-\ell+k-2j, \ell-k+2j}.$$

For the special case of $\ell = n$, then

$$F_{nn} = \sum_{j=0}^n \binom{n}{j} M_{j, n-j}.$$

So, the n^{th} order rotational moment is a linear combination of the conventional moments of order less than and equal to n . If the image is rotated by an angle θ , then

$$F'_{n\ell} = F_{n\ell} e^{i\ell\theta}.$$

If the image is reflected, then

$$F'_{n\ell} = F_{n\ell}^* e^{-i\ell\theta}$$

If the function has N -fold symmetry, then

$$F_{n\ell} = F_{n\ell} e^{i\ell \frac{2\pi}{N} m}, \quad m=0, 1, \dots, N-1.$$

So,

$$F_{nn} = F_{nn} e^{in \frac{2\pi}{N} m}$$

$$F_{nn} [1 - e^{i2\pi n \frac{m}{N}}] = 0, \quad m = 0, 1, \dots, N-1$$

$$\rightarrow F_{nn} = 0, \text{ for } n \bmod N \neq 0.$$

This provides a way to detect symmetry. Because of noise, the higher order moments become less reliable. Therefore, a test on the magnitude of F_{nn} should have a threshold increasing with n .

6.4 Normalization

In order to perform shape recognition it is often necessary to normalize the shape for translation, scale, and rotation. Translation normalization is accomplished by computing the central moments, i.e.

$$\mu_{pq} = \sum_{r=0}^p \sum_{s=0}^q \binom{p}{r} \binom{q}{s} (-M_{10})^{p-r} (-M_{01})^{q-s} M_{rs}.$$

Scale normalization is performed such that $\mu_{00}' = 1$. Thus,

$$\mu_{pq}' = (\mu_{00}^{-1})^{p+q+2} \mu_{pq}.$$

Rotation normalization is much more difficult. Many investigators have resorted to the theory of algebraic invariants [HU62, DUDA77]. This theory provides measures calculated from polynomials of the central moments that are invariant under linear transformations such as a rotation.

A popular alternative is to determine the principle axis of the image. Then rotate the shape by transforming its moments, to place the principle axis along the x-axis [HU62, REEV81a, REEV81b].

Normalize for rotation by the angle α , where

$$\tan(2\alpha) = 2 \frac{\mu'_{11}}{\mu'_{20} - \mu'_{02}}.$$

This angle α to rotate the object however is ambiguous. An additional 180°

rotation will also place the principle axis along the x-axis. It therefore is necessary to add the constraints that $\tilde{\mu}_{20} > \tilde{\mu}_{02}$ and $\tilde{\mu}_{30} > 0$, where $\tilde{\mu}_{pq}$ are the rotated moments. Shapes with N-fold symmetry however have N principle axes.

For the experiments carried out in this report, the rotation normalization is carried out with the help of the rotational moments, F_{nn} . The moments are normalized for translation and scale. Then the rotational moments up to the 5th order are calculated. The phase of the nth order moment ($2 \leq n \leq 5$) having the largest magnitude is used for rotation normalization. The moments are rotated so the phase of the rotated nth order rotational moment is zero, i.e. the rotation angle is $\alpha = -\angle F_{nn}/n$. Then the rotational moments are calculated again. There are still n-1 rotations of $\frac{2\pi k}{n}$, $k = 1, \dots, n-1$, that would allow the nth order rotational moment to have zero phase. To choose among these possible rotations, the following criterion is maximized for the rotational moments of the rotated moments:

$$\sum_{n=2}^5 \text{Re}\{F_{nn}\}.$$

This essentially chooses the normalization that gives the highest degree of symmetry about the principle axis.

As with the other shape analysis methods, an additional feature vector is provided for the unknown shape if the magnitude of the next largest rotational moment is greater than 95% of the largest. This rotational moment is used for normalization in the same way as mentioned above. ..

6.5 Calculation

The images generated for the experiments are discrete, so

$f(x,y) = \sum_{i=1}^N \sum_{j=1}^N f(x_i, y_j) \delta(x - x_i, y - y_j)$. Therefore, the double integral becomes

$$M_{pq} = \sum_{i=1}^N \sum_{j=1}^N x_i^p y_j^q f(x_i, y_j) .$$

A discrete version of Green's theorem has been derived [TANG82] that makes it possible to compute the moments from the chain coded boundary instead of the discrete image.

$$M_{pq} = \sum_{i=1}^{K-1} F_x(x_i, y_i) D_Y(a_{i-1}, a_i) + x_i^p y_i^q C_Y(a_{i-1}, a_i) ,$$

where the (x_i, y_i) are the sequential boundary points (positive integers),

$F_x(x_i, y_i) = y_i^q \sum_{i=1}^{x_i} i^p$, and a_i is the direction chain code. Table 6.1 and Table 6.2

enumerate D_Y and C_Y . The computation of the moments can be speeded up

considerably by having a table of functions for $\sum_{i=0}^{x_i} i^p$. For example,

$$\sum_{i=0}^n i^3 = \frac{n^2(n+1)^2}{4} .$$

6.6 Reconstruction

It is often important to be able to reconstruct an approximation to the original intensity function $f(x, y)$ from the moments set. If $f(x, y)$ is of the class of functions already stated, then its Fourier transform exists

Table 6.1 Table of the function $D_Y(a_{i-1}, a_i)$.

a_i a_{i-1}	0	1	2	3	4	5	6	7
0	0	1	1	1	1	0	0	0
1	0	1	1	1	1	0	0	0
2	0	1	1	1	1	0	0	0
3	0	1	1	1	1	0	0	0
4	-1	0	0	0	0	-1	-1	-1
5	-1	0	0	0	0	-1	-1	-1
6	-1	0	0	0	0	-1	-1	-1
7	-1	0	0	0	0	-1	-1	-1

Table 6.2 Table of the function $C_Y(a_{i-1}, a_i)$.

a_i	0	1	2	3	4	5	6	7
a_{i-1}								
0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	1	0	0
2	0	0	0	0	0	1	1	0
3	0	0	0	0	0	1	1	1
4	1	0	0	0	0	1	1	1
5	1	1	0	0	0	1	1	1
6	1	1	1	0	0	1	1	1
7	1	1	1	1	0	1	1	1

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i2\pi(ux + vy)} f(x, y) dx dy .$$

Since, $f(x, y)$ is at worst piecewise continuous and of bounded support, then $F(u, v)$ can be expanded into a power series

$$F(u, v) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(-i2\pi)^{j+k}}{j! k!} M_{jk} u^j v^k ,$$

where M_{jk} is the $(j, k)^{th}$ moment of $f(x, y)$. If the inverse Fourier transform is attempted, then

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(-i2\pi)^{j+k}}{j! k!} M_{jk} u^j v^k e^{i2\pi(ux + vy)} dx dy .$$

Since the limits of integration and summation are infinite, they cannot be interchanged. So, some other alternative must be explored.

The theory of infinite dimensional normed linear spaces provides the necessary tools. A theorem states that the closest fit to the function on the subspace formed by a finite set of the basis functions is the orthogonal projection onto that subspace. Then the approximation $g(x, y)$ to $f(x, y)$ is just the sum of the reciprocal basis functions for that subspace weighted by the coefficients obtained in the projection, i.e.

$$f(x, y) \simeq g(x, y) = \sum_{i \in \mathbb{N}} \alpha_i \theta_i(x, y) ,$$

where

$$\alpha_i = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \phi_i(x, y) dx dy$$

and θ_i are such that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(x, y) \theta_k(x, y) dx dy = \delta_{ik}.$$

The reciprocal basis for the monomials $\{x^p y^q\}$ is very complicated. To avoid computing the reciprocal basis, the moment set for the monomials are converted to Legendre moments, λ_{mn} , where

$$\lambda_{mn} = \frac{(2m+1)(2n+1)}{4} \sum_{j=0}^m \sum_{k=0}^n C_{mj} C_{nk} M_{jk},$$

where C_{mj} are the Legendre coefficients. That is, the Legendre polynomials are

$$P_m(x) = \sum_{j=0}^m C_{mj} x^j, \quad x \in [-1, 1].$$

Since the Legendre polynomials are orthogonal, they are their own reciprocal basis (except for a multiplicative normalization constant.) So,

$$f(x, y) \simeq g(x, y) = \sum_{n=0}^N \sum_{m=0}^n \lambda_{n-m, m} P_{n-m}(x) P_m(y),$$

where $g(x, y)$ can also be written as

$$g(x, y) = \sum_{n=0}^N \sum_{m=0}^n g_{n-m, m} x^{n-m} y^m.$$

So, $g(x, y)$ is a polynomial whose moments up to order N match those of $f(x, y)$. Therefore, care must be taken to make sure that the moments are for an object which is contained in such a region. If not, the moments must be suitably scaled.

Another reconstruction method is based on the discrete nature of the $(N \times N)$ images and their transformed moments [ROSE76]. Since,

$$M_{pq} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_i^p y_j^q f(x_i, y_j) ,$$

a matrix $[M_{pq}]$ can be written as

$$[M_{pq}] = [P(p, i)] [f(i, j)] [Q(j, q)] ,$$

where

$$f(i, j) = \begin{cases} 1, & \text{if } f(x_i, y_j) = 1, \\ 0, & \text{if } f(x_i, y_j) = 0 \end{cases}$$

$$P(p, i) = x_i^p \quad \text{and} \quad Q(j, q) = y_j^q .$$

The transform can be inverted so that

$$[f(i, j)]_{N \times N} = [P^{-1}(i, p)]_{N \times n} [M_{pq}]_{n \times n} [Q^{-1}(q, j)]_{n \times N} ,$$

where

$$[P^{-1}(i, p)] = [P(p, i)]^{-1} \quad \text{and} \quad [Q^{-1}(q, j)] = [Q(j, q)]^{-1} .$$

The reconstruction is performed over the region $[-A, A] \times [-A, A]$. So, $x_i = \frac{2A}{N} i$ and $y_j = \frac{2A}{N} j$, $i = -N/2, \dots, N/2$. It is important to choose A so that the image described by the moments is contained in the region of definition. Figure 6.1 shows the original silhouette for the letter 'E'. Figure 6.2 shows the reconstruction of the letter 'E' from its moments up to 15th order using this discrete inverse method.

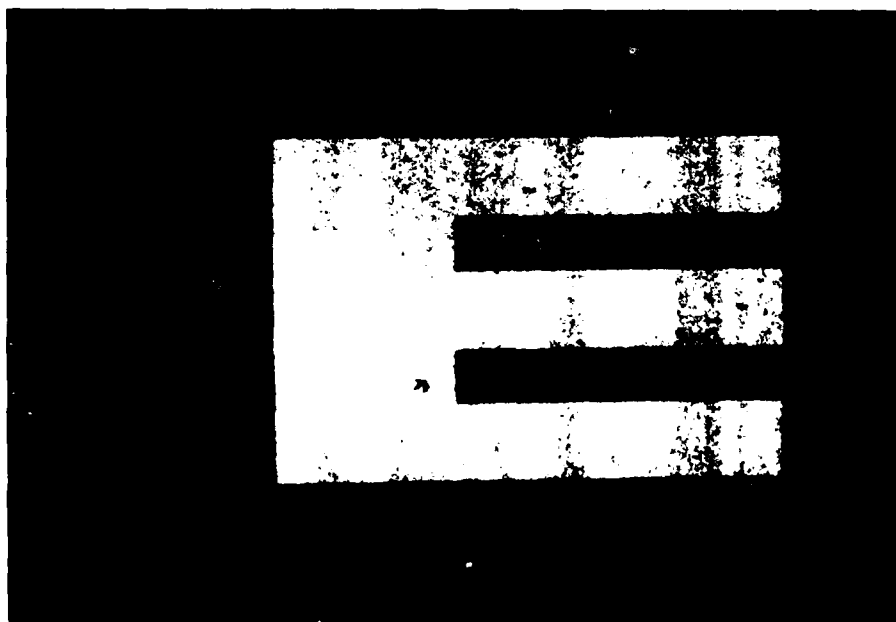


Figure 6.1 Letter 'E' silhouette.

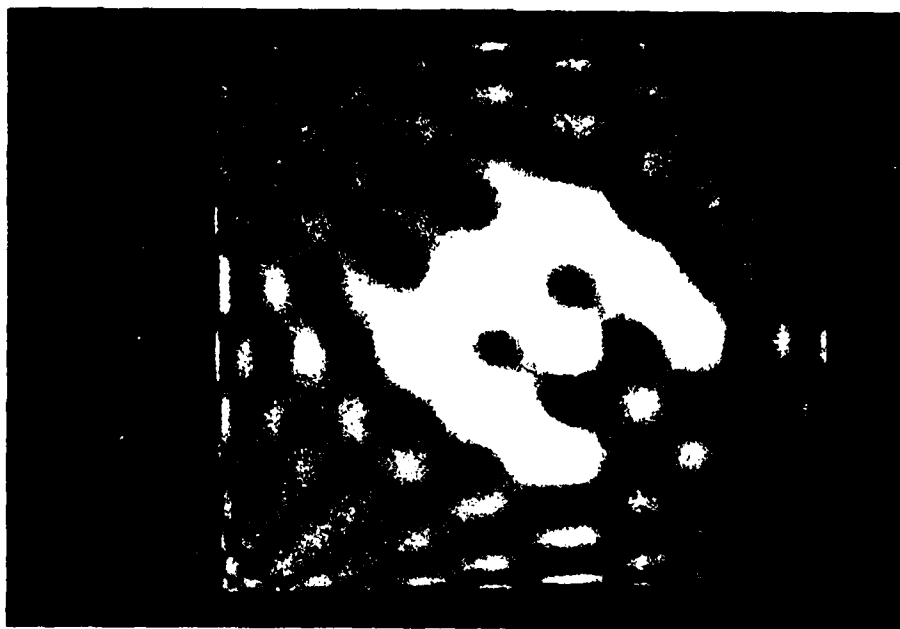


Figure 6.2 The letter 'E' reconstructed from the 15th order moments using the discrete inverse method.

6.7 Generic Shapes

The following are the exact formulae for the moments of an ellipse (circle) and a rectangle (square).

- 1) *Rectangle*: For a rectangle of length 'a' and height 'b' centered at the origin

$$M_{pq} = \frac{[1 + (-1)^p][1 + (-1)^q]}{(p+1)(q+1)2^{p+q+2}} a^{p+1}b^{q+1}$$

$$= \begin{cases} \frac{a^{p+1}b^{q+1}}{(p+1)(q+1)2^{p+q}}, & p \text{ and } q \text{ even} \\ 0, & p \text{ or } q \text{ odd} \end{cases}$$

The rectangle is symmetric about both the x and y axes, so notice that $M_{pq} = 0$ for p and q odd.

There is no need to normalize for translation since $M_{01} = M_{10} = 0$. After normalizing for scale

$$M'_{pq} = \begin{cases} \frac{\left(\frac{a}{b}\right)^{\frac{p-q}{2}}}{(p+1)(q+1)2^{p+q}}, & p \text{ and } q \text{ even} \\ 0, & p \text{ or } q \text{ odd} \end{cases}$$

For the special case of a square having side of length 'a',

$$M'_{pq} = \begin{cases} \frac{1}{(p+1)(q+1)2^{p+q}}, & p \text{ and } q \text{ even} \\ 0, & p \text{ or } q \text{ odd} \end{cases}$$

The rotational moments, F_{nn} , for a rectangle for $n=1$ to 5 are

$$F_{00} = M_{00}' = 1, \quad F_{11} = 0,$$

$$F_{22} = \frac{1}{12ab} (b^2 - a^2), \quad F_{33} = 0,$$

$$F_{44} = \frac{1}{80} \left(\frac{b^2}{a^2} + \frac{a^2}{b^2} \right), \quad F_{55} = 0.$$

If $a > b$ then $F_{22} < 0$. So, the normalization procedure will rotate the rectangle $\frac{\pi}{2}$, so the major axis is along the y axis.

2) *Ellipse*: For an ellipse with major axis of length $2a$ and minor axis of $2b$ ($a \geq b$) centered at the origin, the moments are

$$M_{pq} = \frac{[1 + (-1)^p][1 + (-1)^q]}{2(p+q+2)} a^{p+1} b^{q+1} B\left(\frac{q+1}{2}, \frac{p+1}{2}\right),$$

where $B(m, n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$ [GRAD80]. Again the ellipse is already normalized for translation. Now to normalize for scale, $\lambda = 1/\sqrt{M_{00}} = 1/\sqrt{\pi ab}$. So,

$$M_{pq}' = \frac{[1 + (-1)^p][1 + (-1)^q]}{2(p+q+2) \pi^{\frac{p+q+2}{2}}} \left(\frac{a}{b}\right)^{\frac{p-q}{2}} B\left(\frac{q+1}{2}, \frac{p+1}{2}\right).$$

In the special case where $a = b$, we have a circle whose moments are

$$M_{pq}' = \frac{[1 + (-1)^p][1 + (-1)^q]}{2(p+q+2) \pi^{\frac{p+q+2}{2}}} B\left(\frac{q+1}{2}, \frac{p+1}{2}\right).$$

6.8 Experimental Results

The feature vector for the experiments was formed by dropping the (0,0), (0,1), and (1,0) moments since they are identical for every normalized moment set. The remaining moments up to the n^{th} order were then listed from the lowest to highest as follows:

$$\underline{f} = (\tilde{\mu}_{20}, \tilde{\mu}_{11}, \tilde{\mu}_{02}, \tilde{\mu}_{30}, \tilde{\mu}_{21}, \dots, \tilde{\mu}_{0n})^T.$$

For most of the experiments this full feature vector of 9th order moments is used to 12 real numbers using the principle components transformation.

The image resolution experiment was performed. The results are provided in Table 6.3 and Table 6.4. The performance is low, but follows the same trend as the other methods discussed previously.

The feature vector experiments results are listed in Table 6.5. The unusual fact to note here is that the performance degrades with an increasing number of features. In an attempt to provide insight into this phenomenon, an additional experiment was carried out. It was observed that the high order moments were on the order of 10^4 times smaller than the low order moments. So, an identical experiment was performed except the moments were converted to Legendre moments before being reduced or classified. The Legendre moments appear to have values in a smaller dynamic range. The results of this experiment, however, were nearly identical to the original experiment. An another experiment was performed where the eigenvectors used in the principle components transformation were normalized by dividing by the eigenvalue corresponding to that eigenvector. The results for this experiment were even worse than with no eigenvalue normalization.

Table 6.3 MOM image resolution experimental results: classification accuracy (%).

Unknown Resolution	Prototype Resolution					
	16	32	64	128	256	512
16	28.67	21.67	19.67	16.33	16.0	16.33
32	25.0	37.0	25.0	24.0	25.0	24.0
64	18.33	31.33	36.67	33.67	29.33	28.0
128	19.0	29.0	38.33	40.67	37.67	34.0
256	18.33	25.67	31.33	40.0	39.67	38.33
512	19.0	22.0	31.67	38.0	40.67	40.33

Table 6.4 MOM image resolution experimental results: median angle error $\Delta\phi_x, \Delta\phi_y$ (°).

Unknown Resolution	Prototype Resolution					
	16	32	64	128	256	512
16	39,17	67,24	67,20	68,21	70,21	70,21
32	45,22	47,15	68,17	68,24	63,24	65,24
64	45,28	35,15	40,16	57,17	58,20	59,23
128	34,37	41,28	30,16	32,15	45,17	45,17
256	40,29	32,18	32,15	32,13	28,15	32,16
512	45,37	41,19	41,16	49,15	32,16	28,15

Table 6.5 MOM feature vector experimental results.

Order of Moments	Number of reduced Features	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
3	7	37.67	26,16
7	33	40.0	32,15
9	52	39.33	39,16
14	117	37.33	39,15
7	12	37.0	41,16
9	12	37.67	45,17
14	12	31.67	41,17

The imaging experiment was also carried out. The results are presented in Table 6.6. The performance improves with signal-to-noise ratio. There is a slight threshold effect between 3 and 6 dB for the 64x64 image resolution unknowns. However, the performance increases only slightly over the entire range of signal-to-noise ratios.

Table 6.7 lists the results obtained when the library sampling experiment was performed. Again, the performance improves with increasing library density, but the increase is minimal.

Finally, the partial shape experiment was carried out with its results listed in Table 6.8. The performance of the moments is low to begin with, but it degrades slowly with increasing amounts of the contour being chopped.

6.9 Conclusions

The performance of the moments of the silhouette is unexpectedly very poor. The fact that the moments are used to attempt an approximation to a discontinuous silhouette may explain the decrease in performance with even higher order moments (a Gibb's phenomenon.) But it seems more likely that crucial information is lost when the raw or conventional moments are computed using the monomials, $x^p y^q$, with p and q large. That is, maybe the problem is with the numerical precision of the calculations. It is possible that a nonlinear combination of the moments such as moments invariants would provide the transformation necessary to further extract the shape information. This seems unlikely, however. A great deal of insight might be obtained by improving the reconstruction methods. It is also hard to explain why the results for these particular shape experiments are so much lower than those quoted in the literature.

Table 6.6 MOM imaging noise experimental results.

Library Resolution	Unknown Resolution	SNR (dB)	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
256	64	3	20.0	68,24
		6	26.67	49,21
		10	29.33	50,21
		20	31.33	57,21
		∞	29.33	58,20
256	128	3	32.67	45,23
		6	33.67	41,17
		10	35.33	41,15
		20	37.67	45,17
		∞	37.67	45,17

Table 6.7 MOM library sampling experimental results.

Library Views	Library Resolution	Unknown Resolution	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
143	256	64	29.33	58,20
49			23.67	46,21
9			26.33	45,21
143	256	128	37.67	45,17
49			32.33	26,15
9			27.33	32,17
143	256	256	39.67	28,15
49			35.33	26,15
9			27.0	40,18

AD-A132 842

SHAPE RECOGNITION AND DESCRIPTION: A COMPARATIVE STUDY

3/3

(U) PURDUE UNIV LAFAYETTE IN SCHOOL OF ELECTRICAL

ENGINEERING T A GROGAN ET AL. JUL 83 TR-EE83-22

UNCLASSIFIED

ARO-18016.11-EL-A DAAG29-81-K-0088

F/G 12/1

NL

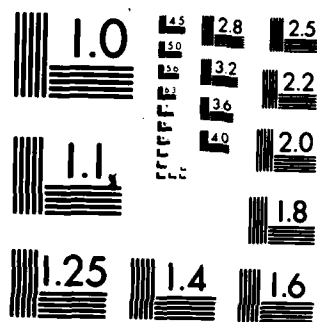
END

DATE

FILED

10 83

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Table 6.8 MOM partial shape experimental results.

Percent Contour Chopped	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y$ (°)
0	37.67	45,17
10	36.33	57,20
20	25.33	53,24
30	25.67	67,25
40	21.0	62,27
50	15.33	46,43

CHAPTER 7

MOMENTS OF THE BOUNDARY CURVE

7.1 Introduction

If a region in the plane is simply connected and its boundary is smooth, then this silhouette is completely specified by its boundary. Moments of the boundary curve have been used [DUDA77] before for shape recognition, but there has been a lack of a theoretical basis. Let all the "mass" of the object be concentrated on the boundary. Let $\gamma(x, y) = 0$ be the equation that describes the boundary curve. Then the moments of the boundary, $B_{M_{pq}}$, are

$$B_{M_{pq}} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q \delta(\gamma(x, y)) dx dy ,$$

where $\delta(\cdot)$ is the generalized delta-function [GELF64] . This is a global shape recognition method.

7.2 Normalization

The operations needed to normalize the moments of the boundary are exactly the same as for the moments of the silhouette with the exception of dilation (or scale.) To find the proper relationship, the moments of the circle boundary will be used.

The equation for a circle of radius c centered at the origin is $x^2 + y^2 = c^2$, or in polar coordinates $\rho = c$. So,

$$B_{M_{pq}} = \int_0^{2\pi} \int_0^{\infty} \rho^{p+q+1} \cos^p \theta \sin^q \theta \delta(\rho-c) d\rho d\theta$$

$$= c^{p+q+1} \int_0^{2\pi} \cos^p \theta \sin^q \theta d\theta$$

$$B_{M_{pq}} = \frac{[1 + (-1)^p][1 + (-1)^q]}{2} B\left(\frac{p+1}{2}, \frac{q+1}{2}\right) c^{p+q+1}.$$

Now, if the circle is scaled so it has a radius of λc , then

$$B_{M_{pq}} = \int_0^{2\pi} \int_0^{\infty} \rho^{p+q+1} \cos^p \theta \sin^q \theta \delta(\rho-\lambda c) d\rho d\theta$$

$$= \lambda^{p+q+1} c^{p+q+1} \int_0^{2\pi} \cos^p \theta \sin^q \theta d\theta.$$

This would seem to imply that

$$B_{M_{pq}}' = \lambda^{p+q+1} B_{M_{pq}}.$$

So, to normalize for scale changes, let $\lambda = 1/B_{M_{00}}$. When $\delta(\gamma(x,y))$ is written as $\delta(\rho-c)$ for a circle, we are assuming that the mass is of unit density and uniform mass distribution. On the other hand, if $\delta(\gamma(x,y))$ had been written as $\delta(\rho^2-c^2)$ for a circle, it would mean there is unit mass uniformly distributed.

7.3 Generic Shapes

Following are the moments of the boundary curve (MOMB) for the rectangle (square) and a circle.

- 1) *Rectangle*: For a rectangle centered at the origin with length 'a' and height 'b' with major axis along the x axis, the moments of the boundary curve are

$$B_{M_{pq}} = \frac{[1 + (-1)^p][1 + (-1)^q]}{2^{p+q+1}} \left\{ \frac{a}{p+1} + \frac{b}{q+1} \right\} a^p b^q.$$

For example, $B_{M_{00}} = 2(a+b)$, $B_{M_{10}} = B_{M_{01}} = 0$, and

$B_{M_{20}} = \frac{a^2}{6}(a+3b)$. After normalizing for scale, the moments are

$$B_{M_{pq}}' = \frac{[1 + (-1)^p][1 + (-1)^q]}{4^{p+q+1}} \left\{ \frac{a}{p+1} + \frac{b}{q+1} \right\} \frac{a^p b^q}{(a+b)^{p+q+1}}.$$

For the special case where $a = b$, a square,

$$B_{M_{pq}}' = \frac{[1 + (-1)^p][1 + (-1)^q]}{8^{p+q+1}} \cdot \frac{p+q+1}{(p+1)(q+1)}.$$

- 2) *Circle*: For a circle of radius c centered at the origin, the moments of the boundary curve are

$$B_{M_{pq}}' = c^{p+q+1} \frac{[1 + (-1)^p][1 + (-1)^q]}{2} B\left(\frac{p+1}{2}, \frac{q+1}{2}\right).$$

For example, $B_{M_{00}} = 2\pi c$, $B_{M_{10}} = B_{M_{01}} = 0$, and $B_{M_{20}} = \pi c^3$. After normalizing for scale, the moments become

$$B_{M_{pq}}' = \frac{1}{(2\pi)^{p+q+1}} \cdot \frac{[1 + (-1)^p][1 + (-1)^q]}{2} \cdot B\left(\frac{p+1}{2}, \frac{q+1}{2}\right).$$

For example, $B_{M_{00}}' = 1$, $B_{M_{10}}' = B_{M_{01}}' = 0$, and $B_{M_{20}}' = \frac{1}{8\pi^2}$.

7.4 Calculation

Since the images are discrete, the moments of the boundary curve are calculated from the (x_i, y_i) coordinates pairs obtained from the Freeman chain code. So,

$${}^B M_{pq} = \sum_{i=0}^{K-1} x_i^p y_i^q .$$

Then the moments are normalized for translation, scale, and rotation. The translation and rotation normalization is the same for the moments of the silhouette. The scale normalization is carried out as described in this chapter.

7.5 Experimental Results

The feature vector for the moments of the boundary is formed in exactly the same way as for the moments of the silhouette. The (0,0), (0,1), and (1,0) moments are dropped and the rest listed as follows:

$$\underline{f} = ({}^B \tilde{\mu}_{20}, {}^B \tilde{\mu}_{11}, {}^B \mu_{02}, \dots, {}^B \tilde{\mu}_{0n})^T .$$

For most of the experiments, the 9th order moments were used followed by the principle components transformation retaining 12 real numbers.

The results of the image resolution experiment are provided in Table 7.1 and Table 7.2. The performance is best when the prototype is of a resolution approximately the same as the unknown. The 256x256 library resolution has the highest average classification accuracy.

The results of the feature vector experiment are shown in Table 7.3. The performance is the same for all the different feature vectors investigated. This would imply that all the shape information (though it is not much) is contained in the 3rd order moments or less. It was observed that the low order moments

Table 7.1 MOMB image resolution experimental results: classification accuracy (%).

Unknown Resolution	Prototype Resolution					
	16	32	64	128	256	512
16	27.33	19.0	17.0	17.67	17.67	19.0
32	25.0	25.33	21.33	23.67	20.33	19.0
64	21.0	25.33	25.0	24.0	26.0	25.67
128	17.67	26.67	26.67	26.33	28.33	29.0
256	18.33	22.33	25.0	26.33	28.0	28.33
512	18.0	19.33	24.0	26.0	29.67	27.33

Table 7.2 MOMB image resolution experimental results: median angle error $\Delta\phi_x, \Delta\phi_y$ (°).

Unknown Resolution	Prototype Resolution					
	16	32	64	128	256	512
16	45,18	39,19	54,25	75,26	83,37	84,29
32	36,26	24,17	34,17	54,21	63,19	63,17
64	46,29	56,17	31,17	42,13	49,17	56,20
128	46,27	46,21	45,17	34,17	29,22	34,16
256	46,26	46,23	29,17	31,19	25,16	39,21
512	46,26	49,21	27,18	27,17	27,21	24,21

Table 7.3 MOMB feature vector experimental results.

Order of Moments	Number of reduced Features	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
3	7	28.33	29,22
5	18	28.33	29,22
7	33	28.33	29,22
9	52	28.33	29,22
10	63	28.33	29,22
5	12	28.33	29,22
7	12	28.33	29,22
9	12	28.33	29,22
10	12	28.33	29,22

were of many orders of magnitude larger than the high order moments. Since the square of the error distance measure was used in the classification, the differences among shapes expressed in the high order moments would be completely dominated by the low order moments. This problem also occurred for the moments of the silhouette, but to a somewhat lesser extent.

The imaging noise experiment was also performed. Its results are provided in Table 7.4. As expected, the performance improves gradually with increasing signal-to-noise ratio.

The library sampling experiment results are listed in Table 7.5. The performance improves moderately with increasing number of library views.

Table 7.6 provides the results for the partial shape experiment. The performance degrades very slowly as more of the contour is chopped.

7.6 Conclusions

The overall performance of this method is very poor. The dynamic range problem indicated for the moments of the silhouette seems to be accentuated for the moments of the boundary. Something would have to be done to address this problem before this method would be of much use for shape recognition.

Table 7.4 MOMB imaging noise experimental results.

Library Resolution	Unknown Resolution	SNR (dB)	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
256	64	3	17.33	39,31
		6	22.33	46,36
		10	26.0	31,27
		20	24.67	39,17
		∞	26.0	49,17
256	128	3	15.67	47,27
		6	15.67	49,32
		10	22.33	47,21
		20	28.33	29,22
		∞	28.33	29,22

Table 7.5 MOMB library sampling experimental results.

Library Views	Library Resolution	Unknown Resolution	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y$ (°)
143	256	64	26.0	49,17
49			21.67	27,17
9			22.0	46,16
143	256	128	28.33	29,22
49			25.0	25,17
9			22.33	46,17
143	256	256	28.0	25,16
49			24.67	21,13
9			22.33	46,18

Table 7.6 MOMB partial shape experimental results.

Percent Contour Chopped	Classification Accuracy (%)	Median Angle Error $\Delta\phi_x, \Delta\phi_y (^\circ)$
0	28.33	29,22
10	28.67	23,18
20	24.67	34,26
30	24.33	63,26
40	18.33	50,33
50	22.33	59,30

CHAPTER 8

COMPARISON OF GLOBAL SHAPE METHODS

8.1 Introduction

In this chapter the shape methods are compared in two ways. First, the analytic relationships between pairs of shape methods are discussed. Then, a comparison is made based upon the shape recognition experiments.

8.2 Analytic Comparisons

The Fourier descriptors of the boundary function and the Walsh points are very similar. The basis functions for both methods are characterized by the number of zero crossings in one basic interval. The Fourier component of a particular frequency has a form very similar to the corresponding sequence Walsh function. A basic difference, however, is that the Fourier functions are everywhere continuous, while the Walsh functions have a finite number of discontinuities. The discontinuous nature of the Walsh functions is what causes them to lack a simple time (starting point) shift property such as the Fourier functions exhibit. This difference is also what often causes the distance measure used in analyzing approximation errors for the Fourier and Walsh functions to be different. Often times the "sup max" norm is used for the Walsh method and an Euclidean norm for the Fourier method.

The boundary function and cumulative angular deviant are related through a nonlinear transformation. If $\gamma(t)$ is the boundary function, then its

derivative can be written as

$$\dot{\gamma}(t) = |\dot{\gamma}(t)| e^{iL\dot{\gamma}(t)}.$$

If the contour was traced uniformly (i.e. as a function of arc length), then

$$|\dot{\gamma}(t)| = K \equiv \text{constant}.$$

So,

$$\dot{\gamma}(t) = K e^{iL\dot{\gamma}(t)}.$$

The angle function, $\theta(l)$, of chapter 5 is then

$$L\dot{\gamma}(t) = \theta(-t).$$

$\theta(l)$ was defined for the object traced clockwise. Thus,

$$\dot{\gamma}(t) = K e^{i\{\phi^{\square}(\frac{2\pi t}{T}) - \frac{2\pi t}{T} + \delta_0\}}, \quad t \in [0, T].$$

Hence,

$$\sum_n i \frac{2\pi n}{T} c_n e^{i \frac{2\pi n}{T} t} = K e^{i\{\sum_n d_n e^{i \frac{2\pi n}{T} t} - \frac{2\pi t}{T} + \delta_0\}}.$$

Determining the coefficients of γ from the coefficients of $\phi^{\square}$ is similar to the problem of finding the spectrum of a phase modulated signal from the spectrum of the modulation signal.

Now, in discussing the relationship between the Fourier descriptors of the boundary and the moments of the silhouette, the class of objects will be restricted to those which are *star shaped* with respect to a (reference) point. An object is star shaped with respect to a point if a line segment joint that point and any other point in the object is completely contained in the object [LAY82]. The $(p, q)^{\text{th}}$ moment of a silhouette, O , is defined as

$$M_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy ,$$

where

$$f(x, y) = \begin{cases} 1, & (x, y) \in O \\ 0, & (x, y) \notin O \end{cases}$$

Writing x and y in polar coordinates

$$x = \rho \cos \theta = \rho \left(\frac{e^{i\theta} + e^{-i\theta}}{2} \right)$$

$$y = \rho \sin \theta = \rho \left(\frac{e^{i\theta} - e^{-i\theta}}{2i} \right)$$

Expanding x^p and y^q

$$x^p = \left[\rho \left(\frac{e^{i\theta} + e^{-i\theta}}{2} \right) \right]^p = \left(\frac{\rho}{2} \right)^p \sum_{m=0}^p \binom{p}{m} (e^{i\theta})^{p-m} (e^{-i\theta})^m$$

$$= \left(\frac{\rho}{2} \right)^p \sum_{m=0}^p \binom{p}{m} e^{i(p-2m)\theta}$$

$$y^q = \left[\rho \left(\frac{e^{i\theta} - e^{-i\theta}}{2i} \right) \right]^q = \left(\frac{\rho}{2} \right)^q \sum_{k=0}^q (i)^{2k-q} e^{i(q-2k)\theta} .$$

Substituting into the moment equation

$$M_{pq} = \int_0^{2\pi} \int_0^{\infty} \left[\left(\frac{\rho}{2} \right)^p \sum_{m=0}^p \binom{p}{m} e^{i(p-2m)\theta} \right] \cdot \left[\left(\frac{\rho}{2} \right)^q \sum_{k=0}^q (i)^{2k-q} e^{i(q-2k)\theta} \right] g(\rho, \theta) \rho d\rho d\theta .$$

$$= \sum_{m=0}^p \sum_{k=0}^q \frac{(i)^{2k-q}}{2^{p+q}} \binom{p}{m} \binom{q}{k} \int_0^{2\pi} \int_0^{\infty} \rho^{p+q+1} e^{i[p+q-2(m+k)]\theta} g(\rho, \theta) d\rho d\theta,$$

where $g(\rho, \theta) = f(x \cos \theta, y \sin \theta)$. Then for a fixed θ

$$g(\rho, \theta) = \begin{cases} 1, & \rho \leq r_b(\theta) \\ 0, & \rho > r_b(\theta) \end{cases} = u(\rho - r_b(\theta)).$$

$r_b(\theta)$ is the distance from the reference point to the boundary, and $u(t)$ is the Heaviside step function. Since, the object is star shaped, $r_b(\theta)$ is unique for a fixed θ . The double integral can be written as

$$\begin{aligned} & \int_0^{2\pi} \int_0^{\infty} \rho^{p+q+1} e^{i[p+q-2(m+k)]\theta} u(\rho - r_b(\theta)) d\rho d\theta \\ &= \int_0^{2\pi} \int_0^{r_b(\theta)} \rho^{p+q+1} e^{i(p+q+2)\theta} e^{-i2(m+k+1)\theta} d\rho d\theta \\ &= \frac{1}{p+q+2} \int_0^{2\pi} [r_b(\theta)]^{p+q+2} e^{i(p+q+2)\theta} e^{-i2(m+k+1)\theta} d\theta. \end{aligned}$$

If the object is traced so that $\gamma(\theta) = r_b(\theta) e^{i\theta}$, then the above integral becomes

$$\frac{2\pi}{(p+q+2)} \cdot \frac{1}{2\pi} \int_0^{2\pi} [\gamma(\theta)]^{p+q+2} e^{-i2(m+k+1)\theta} d\theta.$$

Notice that

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} \gamma(\theta) e^{-in\theta} d\theta.$$

If we let $\{*^n c\}_i$ to be the i^{th} element of the n^{th} order convolution of the $\{c_n\}$ sequence with itself, then

$$M_{pq} = \sum_{m=0}^p \sum_{k=0}^q \frac{(i)^{2k-q}}{2^{p+q}} \binom{q}{k} \binom{p}{m} \frac{2\pi}{(p+q+2)} [x^{p+q+2}]_{2(m+k+1)}.$$

Notice that if the boundary is traced so that $\gamma(\theta) = r_b(\theta) e^{i\theta}$, then γ is not uniformly traced.

A more general relationship between the moments of an object O and the Fourier descriptors of the boundary function γ can be obtained using Green's theorem. The $(p, q)^{\text{th}}$ moment of a simply connected object is

$$M_{pq} = \int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} x^p y^q f(x, y) dx dy.$$

Green's theorem states that

$$\iint_O \left(\frac{\partial M}{\partial x} - \frac{\partial N}{\partial y} \right) dx dy = \int_{\gamma} M dx + \int_{\gamma} N dy.$$

Letting

$$N = \frac{1}{2} \frac{x^{p+1} y^q}{p+1} \quad \text{and} \quad M = -\frac{1}{2} \frac{x^p y^{q+1}}{q+1},$$

then

$$\frac{\partial N}{\partial x} = \frac{1}{2} x^p y^q \quad \text{and} \quad \frac{\partial M}{\partial y} = \frac{-1}{2} x^p y^q.$$

So,

$$\begin{aligned} M_{pq} = \iint_O x^p y^q dx dy &= -\frac{1}{2(q+1)} \int_{\gamma} x^p y^{q+1} dx \\ &+ \frac{1}{2(p+1)} \int_{\gamma} x^{p+1} y^q dy \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \int_0^T \left[\frac{x^{p+1}y^q}{(p+1)} \cdot \frac{dy}{dt} - \frac{x^py^{q+1}}{(q+1)} \cdot \frac{dy}{dt} \right] dt \\
&= \frac{1}{2} \int_0^T x^py^q \left[\frac{x}{p+1} \cdot \frac{dy}{dt} - \frac{y}{q+1} \cdot \frac{dx}{dt} \right] dt.
\end{aligned}$$

Since,

$$x(t) = \frac{\gamma(t) - \gamma^*(t)}{2} \quad \text{and} \quad y(t) = \frac{\gamma(t) + \gamma^*(t)}{2i}$$

Then

$$\begin{aligned}
M_{pq} &= \frac{1}{4(p+1)(q+1)} \sum_{m=0}^p \sum_{k=0}^q \binom{p}{m} \binom{q}{k} \int_0^T [(p-q)(\dot{\gamma}\gamma - \gamma^*\dot{\gamma}^*) \\
&\quad + (p+q+2)(\dot{\gamma}\dot{\gamma}^* - \gamma^*\dot{\gamma})] \gamma^{p+q-m-k} (\gamma^*)^{m+k} dt
\end{aligned}$$

If $\gamma(t)$ can be written as a finite sum of Fourier components then, in principle, this expression can be evaluated. In chapter 3, the area formula in terms of the Fourier descriptor coefficients was derived using this expression, i.e. Area = M_{00} .

8.3 Empirical Comparisons

In this section the global shape methods will be compared on the basis of the evidence provided by results of the various experiments.

First, the image resolution experiment will be discussed. The diagonal of the tables of results for the image resolution experiment are plotted together in Figure 8.1, Figure 8.2, and Figure 8.3. This experiment sets the stage for the rest of the combined results in that it clearly ranks the methods performances.

Fourier descriptors of the boundary (FDS) and Walsh points (WAL) have the best results. Then the cumulative angular deviant Fourier descriptor method (CAD) is next, but at a significantly lower performance. The moments of the silhouette (MOM) and the moments of the boundary (MOMB) have the worst performance. The performances for all the methods quickly deteriorate below the 32x32 image resolution. Since, this occurs for all the methods, it is likely that this is due to the particular aircraft shapes being used. The median y angle error is usually somewhat lower than the median x angle error. This is also probably due to the aircraft shapes themselves.

The combined results for the feature vector experiment are plotted in Figure 8.4, Figure 8.5, and Figure 8.6. The results for this experiment are the most difficult to explain. The behavior of the Fourier descriptors and Walsh points are as would be expected. As the number of features increase, so does the classification accuracy. The performance for the cumulative angular deviant and moments of the boundary increase but then decrease. It is difficult to explain why this occurs. The moments of the boundary on the other hand, neither increase or decrease over the range investigated. This would indicate that all the information available is contained in the first few moment values.

The results for the library sampling experiment are plotted in Figure 8.7, Figure 8.8, and Figure 8.9. As the number of library views are increased the classification results improve. Again, the Walsh points and Fourier descriptors of the boundary are very close with the other methods following a similar trend but with a poorer performance level.

The imaging noise experiment results are plotted in Figure 8.10, Figure 8.11, and Figure 8.12. All the methods seem to exhibit a threshold effect. The Fourier descriptors and Walsh points effectiveness degrades markedly between

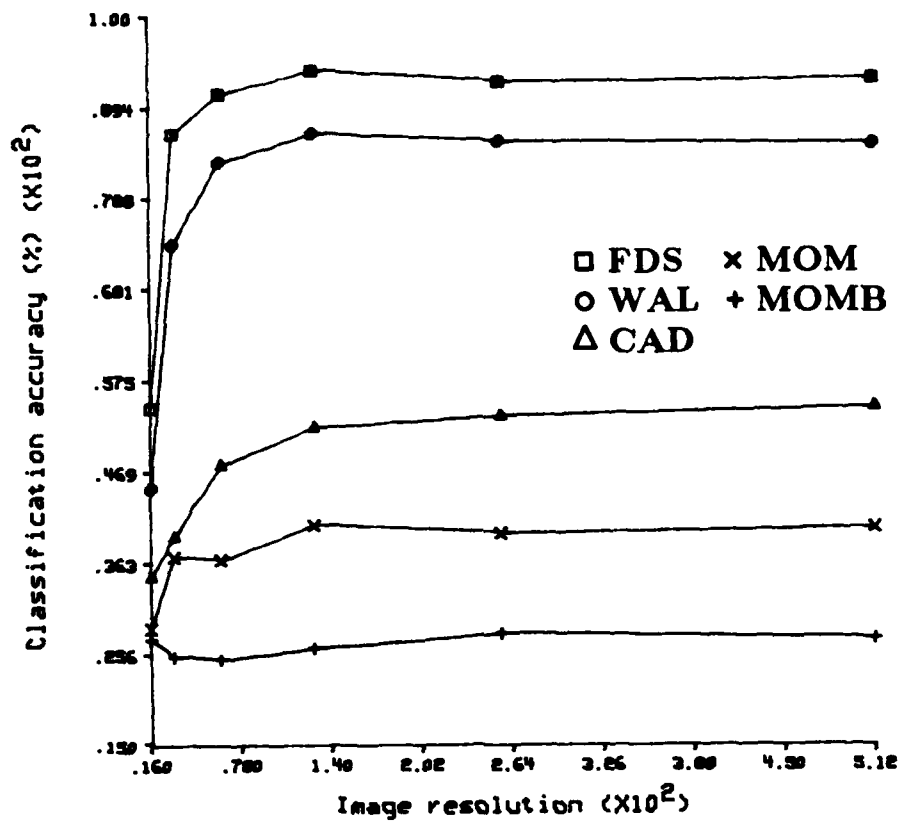


Figure 8.1 Combined results for the image resolution experiment: classification accuracy.

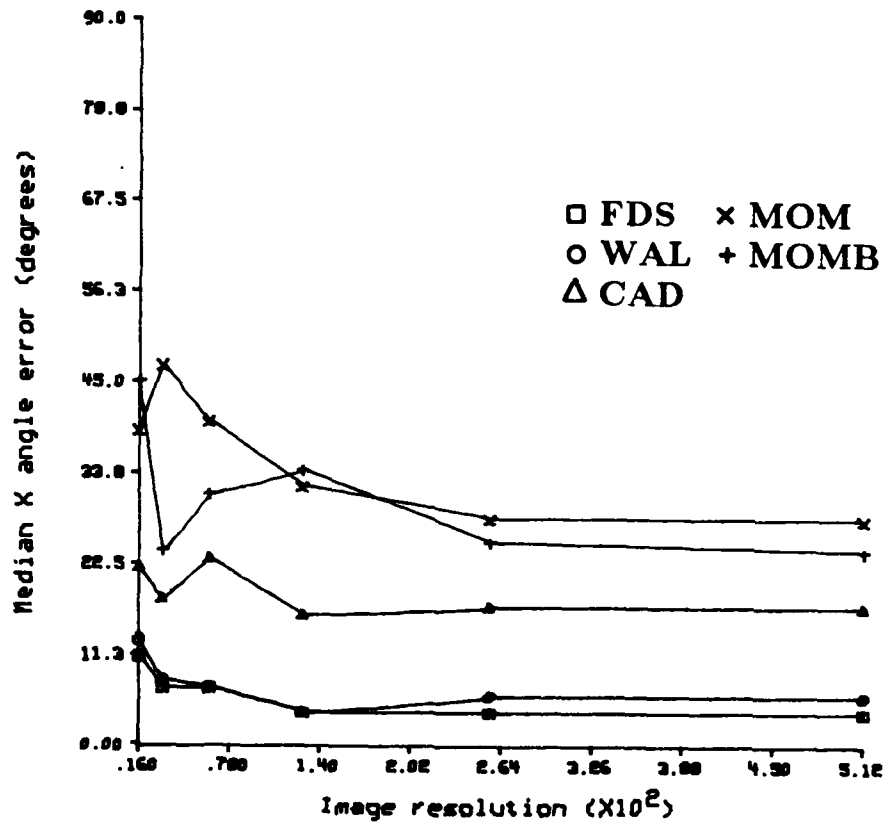


Figure 8.2 Combined results for the image resolution experiment: median x angle error.

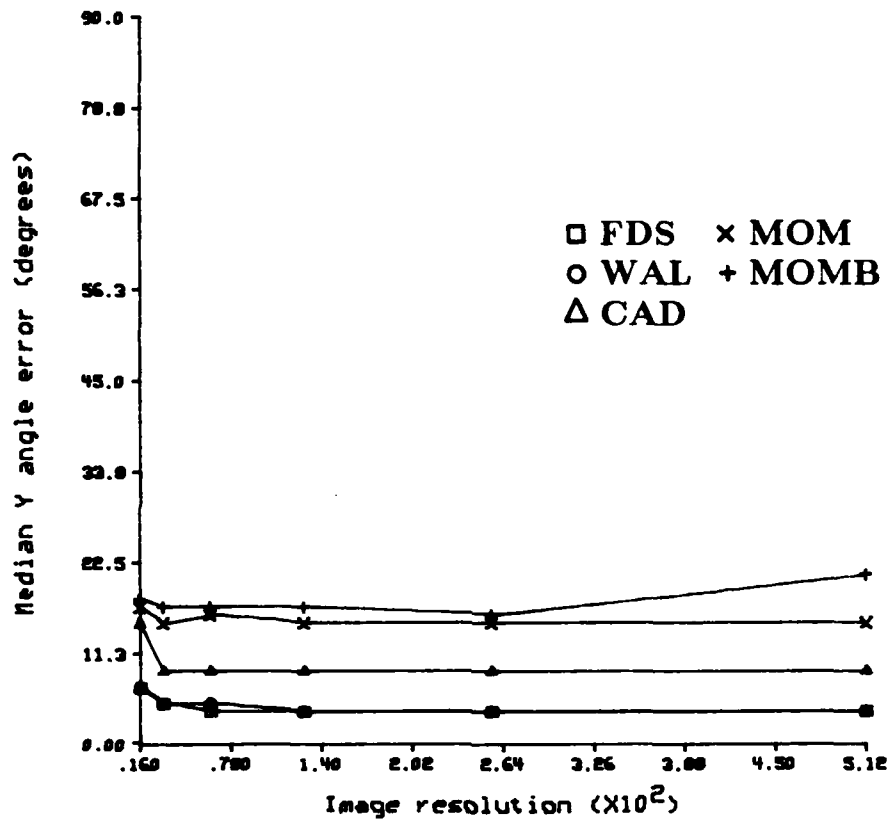


Figure 8.3 Combined results for the image resolution experiment: median y angle error.

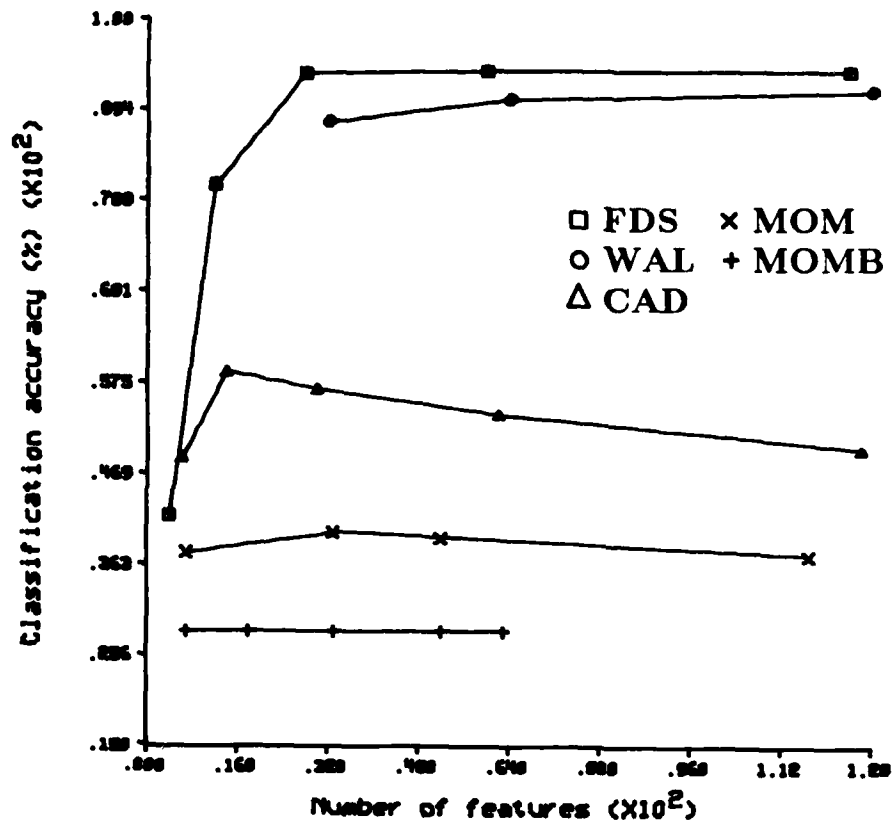


Figure 8.4 Combined results for the feature vector experiment: classification accuracy.

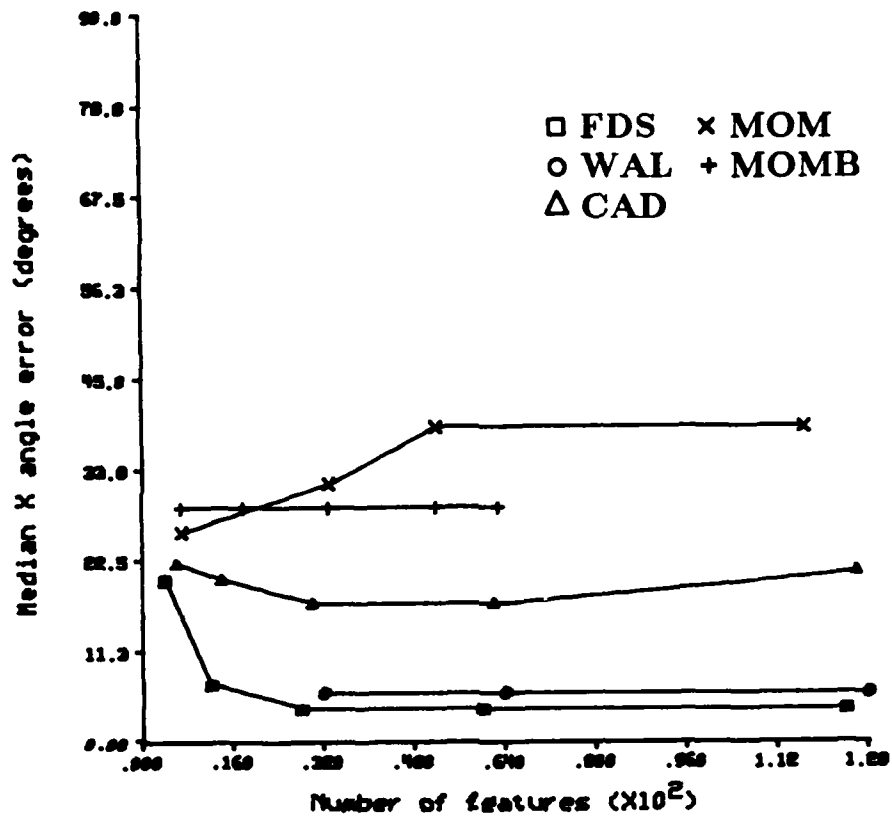


Figure 8.5 Combined results for the feature vector experiment: median x angle error.

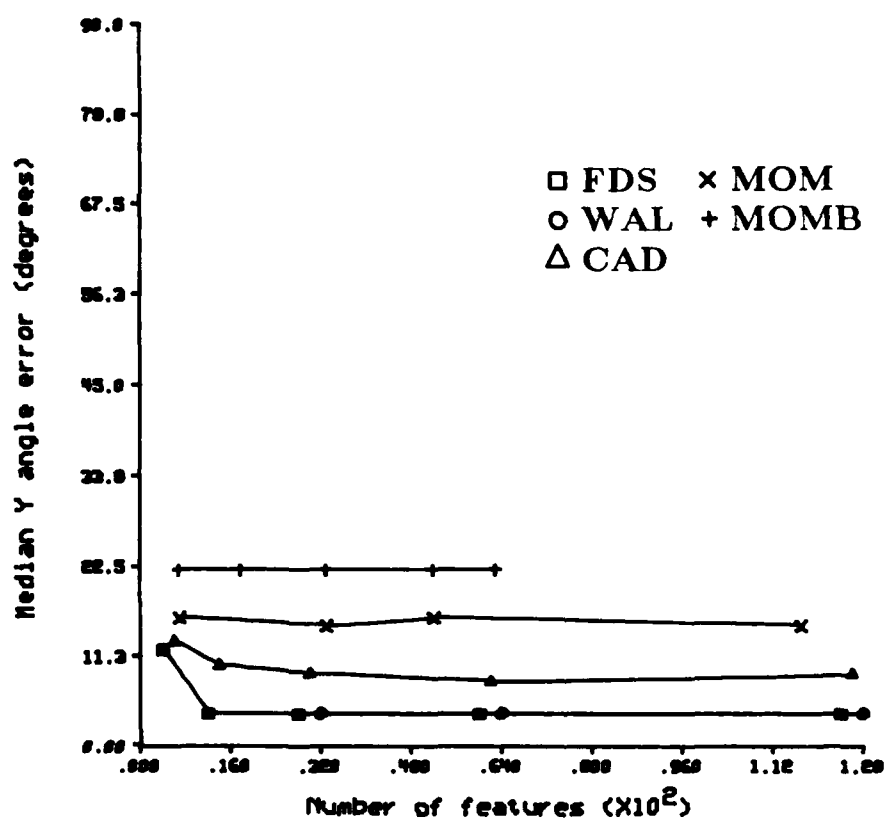


Figure 8.6 Combined results for the feature vector experiment: median y angle error.

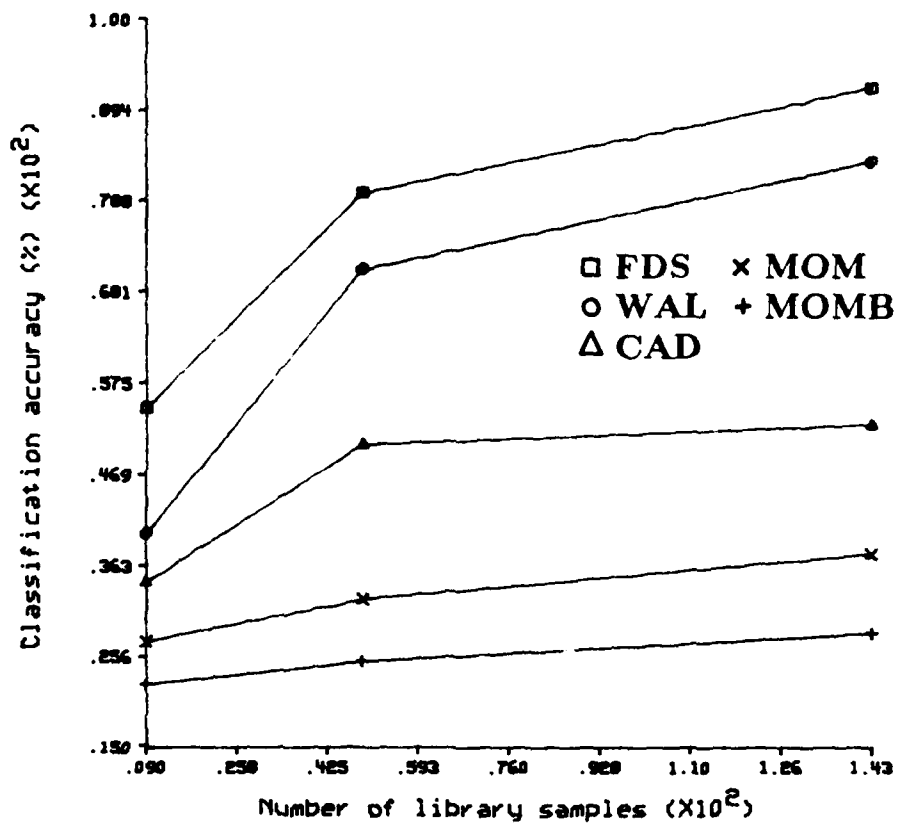


Figure 8.7 Combined results for the library projection experiment: classification accuracy.

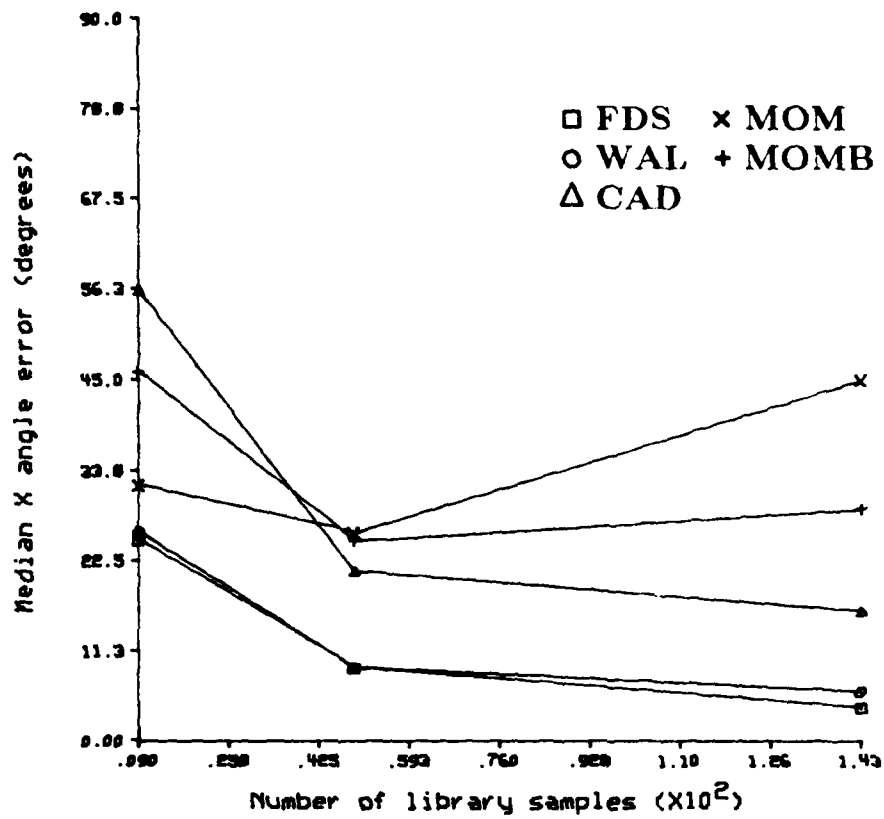


Figure 8.8 Combined results for the library projection experiment: median x angle error.

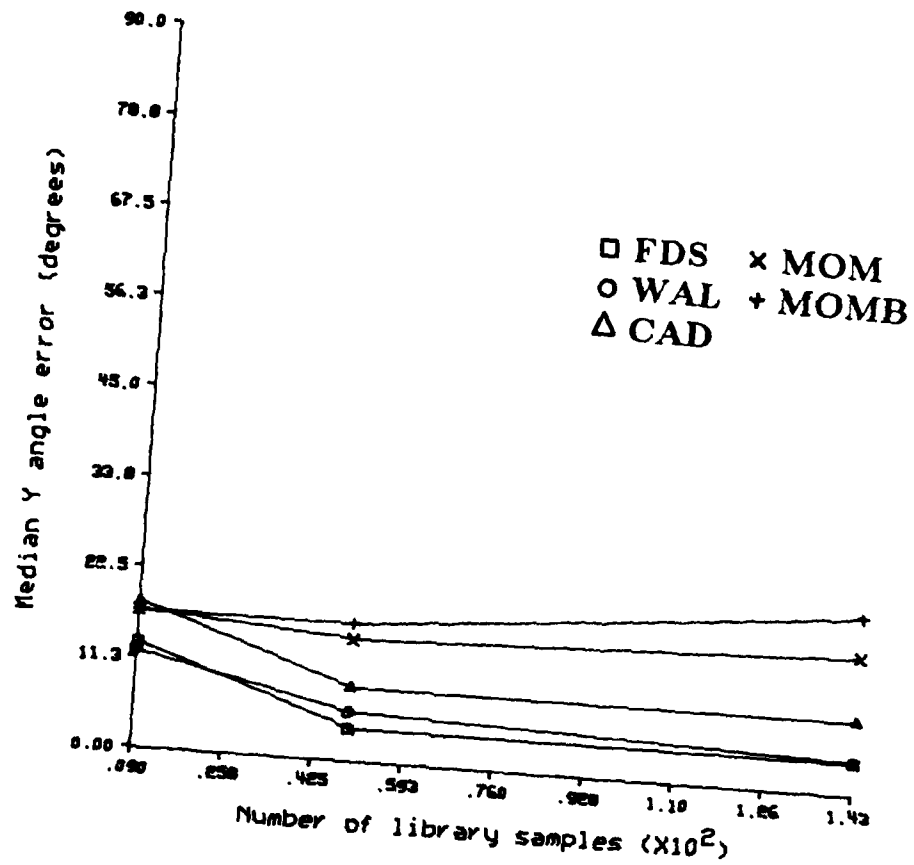


Figure 8.9 Combined results for the library projection experiment: median y angle error.

6 and 10 dB signal-to-noise ratio. The moments methods exhibit a very soft threshold. The cumulative angular deviant Fourier descriptors method exhibits a very prominent threshold between 10 and 20 dB signal-to-noise ratio.

The results for the partial shape experiment are plotted in Figure 8.13, Figure 8.14, and Figure 8.15. The Fourier descriptors and Walsh points have similar characteristics. The performance of both of these methods decreases rapidly once more than 10% of the contour is chopped. Since, these are all global shape methods, this is an expected result. The moments methods, however, degrade very slowly. Their performance is so poor to begin with, this ability is still not very useful. The performance of the cumulative angular deviant, on the other hand, degrades immediately from its already poor level as the contour is chopped.

8.4 Conclusions

The Fourier descriptors of the boundary and the Walsh points are very similar methods. Both perform excellently in comparison to the other shape methods. The cumulative angular deviant is a close relative of the Fourier descriptors. So its performance is closer to that of the Fourier descriptors of the boundary. But it is still significantly worse in classification accuracy with respect to the Fourier descriptors of the boundary.

The moments methods had a bad performance. From the results published in the literature, this was unexpected. It is difficult to explain the poor results, but as noted earlier, it seems that the large dynamic range in the moments might cause this poor performance. Another possibility might be that smooth polynomials are being asked to approximate a discontinuous binary function.

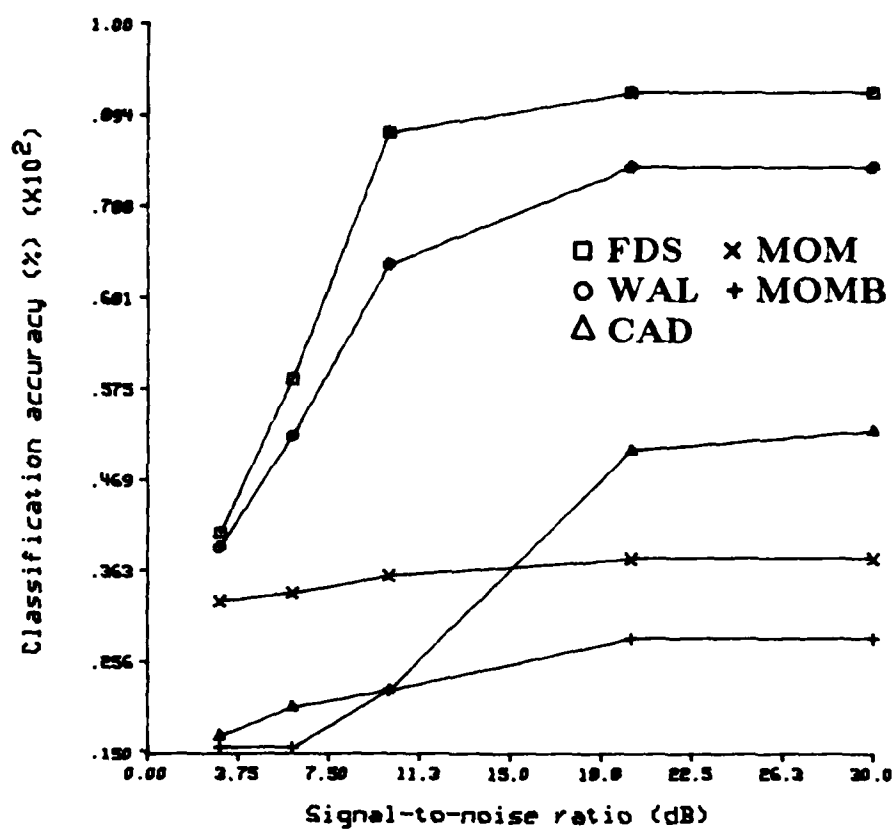


Figure 8.10 Combined results for the imaging noise experiment: classification accuracy.

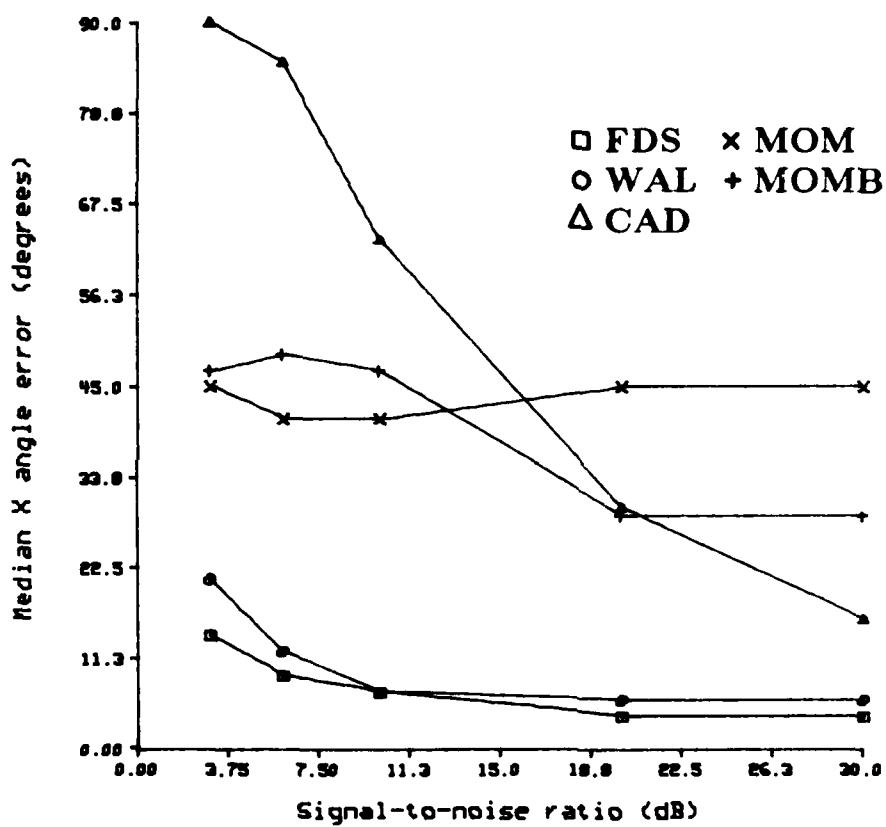


Figure 8.11 Combined results for the imaging noise experiment: median x angle error.

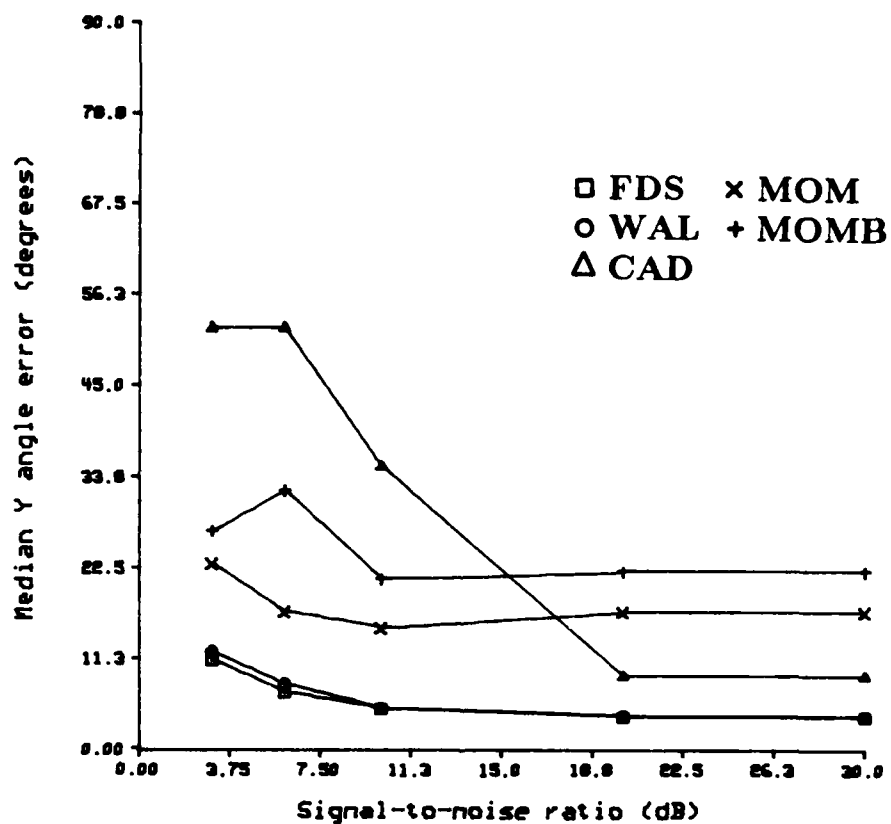


Figure 8.12 Combined results for the imaging noise experiment: median y angle error.

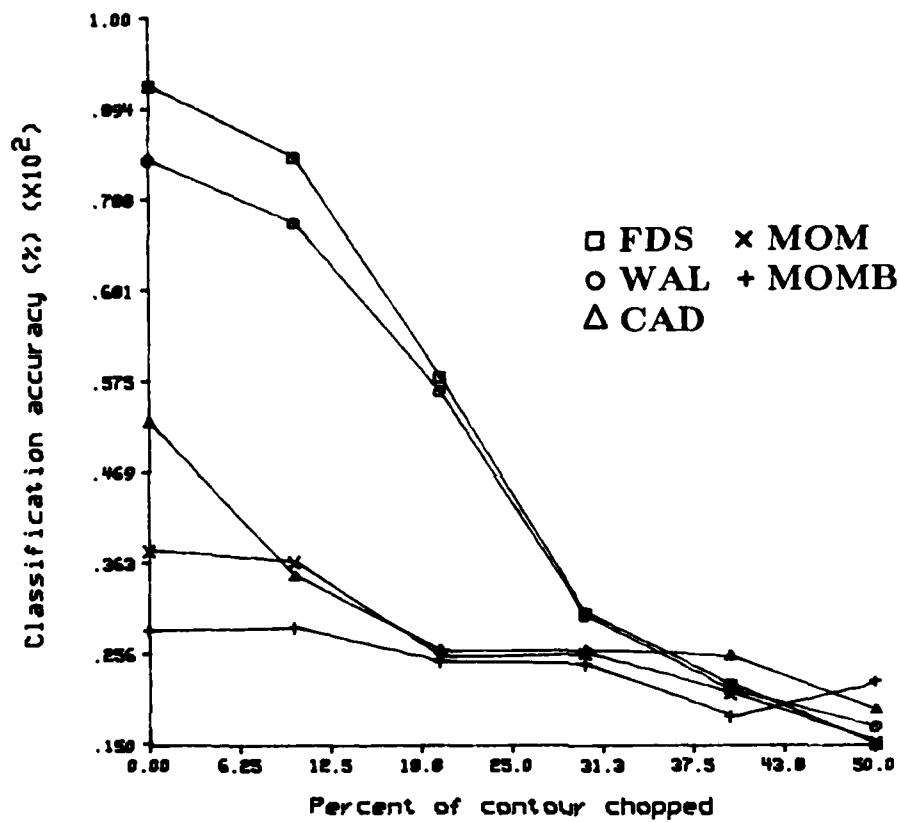


Figure 8.13 Combined results for the partial shape experiment: classification accuracy.

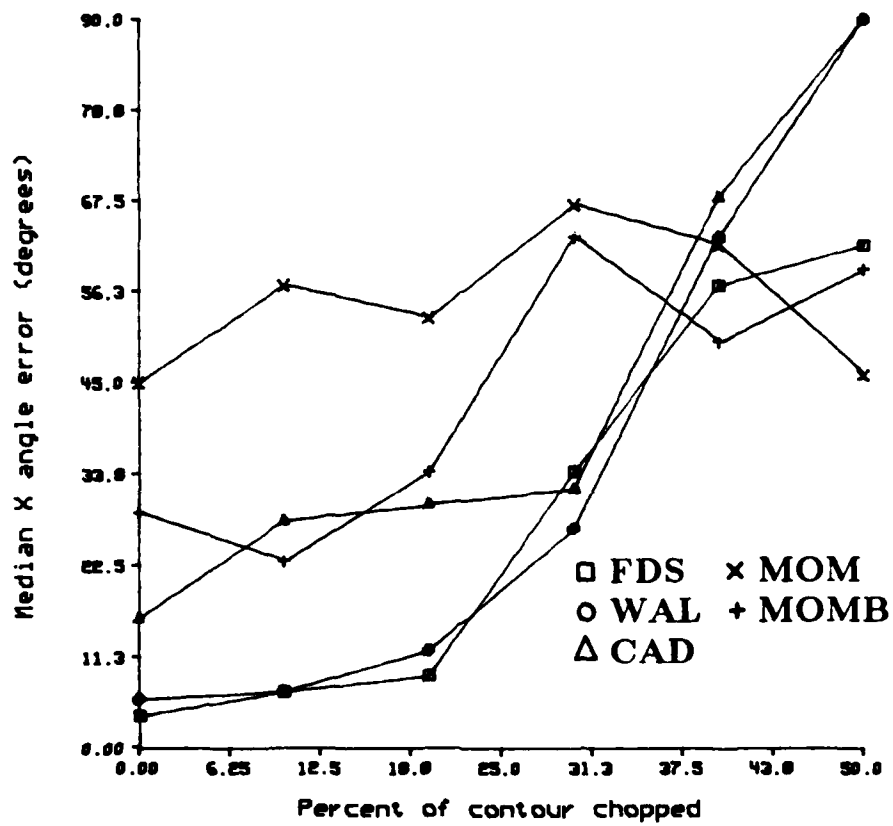


Figure 8.14 Combined results for the partial shape experiment: median x angle error.

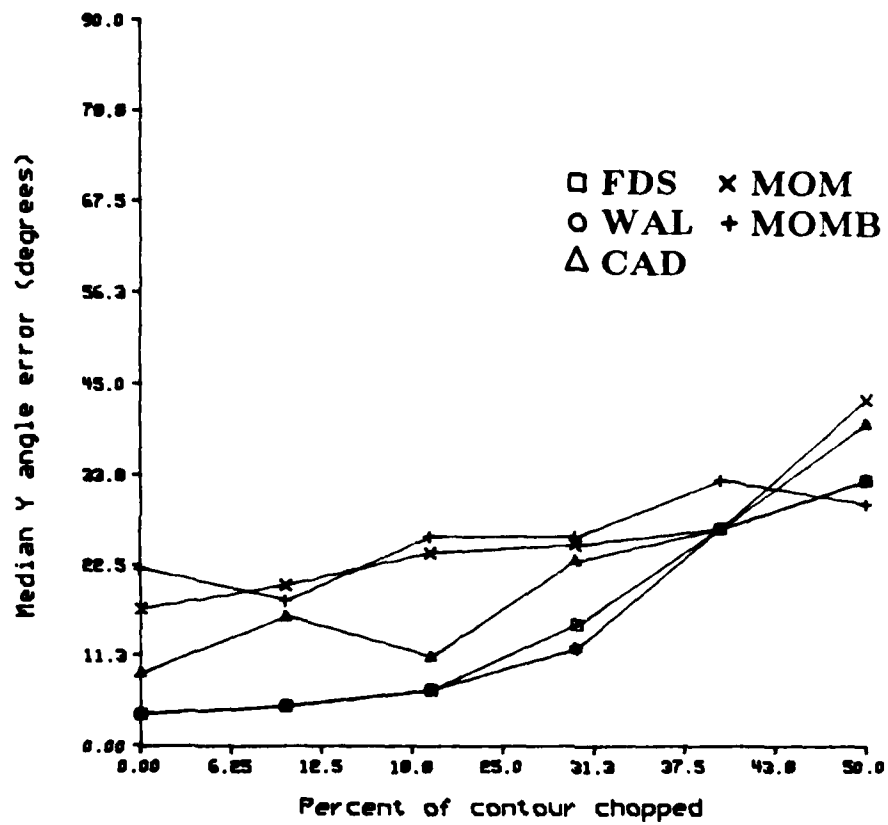


Figure 8.15 Combined results for the partial shape experiment: median y angle error.

CHAPTER 9

FOURIER-MELLIN TRANSFORM TECHNIQUE FOR PARTIAL SHAPE RECOGNITION

9.1 Introduction

Often the shapes to be recognized are only partially correct. This can come about as a result of noise, or of errors in segmenting the object, or if the view of the object is in some way obscured. If this happens then some part of the object's contour is either missing or distorted. Methods of complete shape analysis do not deal with this problem. Global, complete shape analysis methods rely on the fact that shape information distributed over the entire shape effects each component of the characterization. So, when part of the shape is in error, all of the features are degraded. Local, complete shape methods, even though using local information, usually utilize some global information to normalize the local features. Sometimes this global information needed can be tacitly assumed. For instance, the method used by Gifford [GIFF82] uses the distances from a group of features to normalize all the other distance features. Partial shape analysis methods attempt to overcome this difficulty. Partial methods extract local features invariant to scale, rotation, etc. and then attempt to combine features when necessary to attempt a match to the corresponding features of a prototype.

A great deal of effort has been exerted to apply syntactic pattern recognition methods to partial shape recognition [FU82]. Grammars are

developed to implement rules for the shape allowing for missing parts. However, it becomes very difficult to generate such rules when there are a large number of objects and possibly many views of each object. Algorithms for automatic generation of the necessary rules from test data have not been developed. As an alternative, methods that try to dynamically "warp" the features to provide a match to the library shape have been developed [WALL81, GIFF82]. These methods have had some success, but the need to test many possible combinations of features causes them to be very time consuming. Also, there are a number of thresholds and others parameters that must be determined a priori.

The method of partial shape recognition presented here attempts to overcome these difficulties by using transform methods [GROG83]. The essential character of the method is that it uses global information to determine how the features are to be aligned. Then a local comparison to determine the degree of match is performed. This global information is obtained by the use of a shift and scale invariant Fourier-Mellin correlation.

9.2 Curvature Function

Because of the physiological and psychological evidence, most partial shape methods use the curvature or the angles associated with the boundary of an object. If the boundary function, $\gamma(t) = x(t) + iy(t)$, is twice differentiable and $\gamma'(t) \neq 0$, $t \in [0, L]$, then the curvature function, $\kappa(t)$, is

$$\kappa(t) = \frac{\partial}{\partial t} \tan^{-1} \frac{\dot{y}(t)}{\dot{x}(t)},$$

where $\dot{x}(t)$ and $\dot{y}(t)$ are the derivatives of x and y with respect to t . Since the boundary functions for discrete images are polygons, a discrete approximation

of the curvature [BENN75] is

$$\kappa_i = \tan^{-1} \frac{y_i - y_{i-1}}{x_i - x_{i-1}} - \tan^{-1} \frac{y_{i-1} - y_{i-2}}{x_{i-1} - x_{i-2}}, \quad i = 0, \dots, K-1.$$

The discrete nature of the image and the use of the Freeman chain code causes the curvature to be quantized. Filtering is used to smooth the curvature function.

The curvature uniquely specifies a curve independent of translation and rotation. The scale of the shape is normalized by scaling the period of the curvature function. So, the curvature function sequence is resampled to a fixed number of points. For a simple closed curve, the integral of the curvature over a period is $\pm 2\pi$ (360°). So, the resampling is performed in such a way to preserve the total curvature.

If the shape was complete then only a starting point normalization would be necessary to compare an unknown curvature function to a prototype. However, when the contour is only partially correct, then the contour length of corresponding curvature features is also likely to be different. So, a scaling and shift of the time axis of the curvature needs to be determined. The Mellin transform has a time scale property [ALTE78,BAUD73,ROBB72] that is useful in determining the rescaling of the time axis of the unknown to match the prototype.

9.3 Mellin Transform

The Mellin transform has been used before to perform scale invariant correlation [CASA76a,CASA76b,CASA77,CASA78]. The Mellin transform is defined as

$$M_F(s) = \int_0^{\infty} F(x) x^{s-1} dx .$$

Evaluating the transform along the imaginary axis,

$$M_F(-iu) = \int_0^{\infty} F(x) x^{-iu-1} dx .$$

The Mellin has the following time scale property. If $G(x) = F(ax)$, then $M_G(-iu) = M_F(-iu) a^{-iu} = M_F(-iu) e^{-iu \ln a}$. Thus, $|M_G(-iu)| = |M_F(-iu)|$. A scaling of the independent variable introduces a simple linear phase shift in the transform domain.

Since the curvature functions for the boundaries are discrete, a discrete Mellin transform is desired. The Mellin transform is related to the Fourier transform. If $x = T e^t$, then the Mellin transform becomes

$$M_F(-iu) = T^{-iu} \int_{-\infty}^{\infty} F(T e^t) e^{-iut} dt .$$

$$M\{F(x)\} \equiv T^{-iu} \mathcal{F}\{F(T e^t)\} .$$

For a discrete signal, once the exponential sampling has been completed, then the Fast Fourier transform (FFT) is used to compute the Fourier transform. There are some difficulties when the sequence is exponentially sampled. Resampling of the function near the origin creates problems. Also, it is not clear what kind of resampling should be done. To circumvent these problems a

direct calculation of the Mellin transform has been developed [ZWIC83]. The discrete signal is assumed to be constant over each sample interval and the first sample in the sequence is zero. The Direct Mellin Transform (DMT) is

$$M_F(-iu_j) = \sum_{k=0}^{K-1} \frac{\Delta_k}{-iu_j} e^{-iu_j \cdot \ln k},$$

where

$$\Delta_k = F(x_k) - F(x_{k+1}) \quad \text{and} \quad u_j = \frac{2\pi}{N} j, \quad j = 1, 2, \dots, N-1.$$

9.4 Fourier-Mellin Technique

In order to match an unknown curvature function for a partial shape to a prototype, it is first necessary to eliminate the effects of the unknown starting point shift. This is accomplished by taking the magnitude of the Fourier transform. The Fourier transform has a time shift property analogous to the scale property of the Mellin transform. If $g(t) = f(t-t_0)$ and $a > 0$, then $\mathcal{F}\{g(t)\} = e^{-i\omega t_0} \mathcal{F}\{f(t)\}$. So, $|\mathcal{F}\{g(t)\}| = |\mathcal{F}\{f(t)\}|$. Taking the magnitude of the Fourier transform to obtain time shift invariance may cause some problems. The loss of information in the phase may cause ambiguity that will place a number of curvature functions into the same class [WALT63]. For instance, a signal and its Hilbert transform have the same Fourier transform magnitude.

The Mellin transform is used to obtain a scale estimate. Let $g(t) = f(at-t_0)$, ($a > 0$). Then

$$|\mathcal{F}\{g(t)\}| = \left| \frac{1}{a} F\left(\frac{\omega}{a}\right) \right| ,$$

where $F(\omega) = \mathcal{F}\{f(t)\}$. Let

$$F^+(\omega) = \begin{cases} |F(\omega)|, & \omega > 0, \\ 0, & \omega \leq 0 \end{cases} \quad G^+(\omega) = \begin{cases} |\mathcal{F}\{g(t)\}|, & \omega > 0, \\ 0, & \omega \leq 0 \end{cases} .$$

Now, taking the Mellin transform

$$M_G(-iu) = \frac{1}{a} a^{iu} M_F(-iu) , \quad \text{where}$$

$$M_F(-iu) = M\{F^+(\omega)\} \quad M_G(-iu) = M\{G^+(\omega)\} .$$

So, the Fourier transform of $M_F \times M_G^*$ is a cross correlation whose peak is shifted from the origin by $\ln(a)$, i.e.

$$\mathcal{F}\{M_F \times M_G^*\} = C_{FF}(\tau - \ln a) ,$$

where

$$C_{FF}(\tau) = \mathcal{F}\{M_F \times M_F^*\} .$$

Once this scale is determined, the unknown is resampled. The peak of a circular correlation of the prototype curvature sequence and the resampled unknown curvature sequence gives the shift necessary to align the two curvature functions. Now a local comparison can be made. The local comparison used in this report is simply the pointwise difference between the two curvature functions weighted by $1/\cosh(\tilde{\kappa}_i^P)$, where $\tilde{\kappa}_i^P$ is a smoothed version of the prototype curvature function, i.e.

$$d_i' = \frac{\kappa_i^U - \kappa_i^P}{\cosh(\tilde{\kappa}_i^P)} .$$

The weighting function was used to make it easier for a match to be made even when a curvature peak is slightly misaligned. The weighted difference sequence is then thresholded to determine the segments of the shape that match the prototype, i.e.

$d_i' \geq t$, then the i^{th} segment doesn't match.

$d_i' < t$, then the i^{th} segment matches.

An overall measure of the degree that the unknown matches the prototype, similar to the distance used by Gifford [GIF82], is defined as

$$d = d' + \frac{d'}{f} \left(\frac{1}{f} - 1 \right) ,$$

where

$$d' = \sum_{i \in M} d_i' ,$$

and M = set of indices, i , where κ_i^U matches κ_i^P and f = the fraction of the total number of points that match. It is this distance measure that is used to classify the unknown shapes in the experiments described later.

9.5 Experimental Results

First, the steps of the algorithm will be followed for a single unknown shape. Figure 9.1 depicts a F104 prototype contour and an unknown contour. This is the prototype with 10% of the original contour deleted. Also notice that the starting point has been changed. The deleted contour segment is



Figure 9.1 Complete F104 contour prototype (top) and 10% chopped unknown (bottom) contours.

replaced with a right angle segment which, in this case, is longer than the original piece of contour it replaces. Figure 9.2 is the curvature function for the prototype contour before filtering. A 30 point rectangular filter is then applied. The resulting sequence is shown in Figure 9.3. This sequence is then resampled from 624 samples to 256 samples. (See Figure 9.4.) The unknown curvature function is also computed, smoothed, and resampled to 256 samples. It is shown in Figure 9.5.

For comparative purposes, the unknown and prototype are correlated. (See Figure 9.6.) Then the unknown is shifted by the amount corresponding to the peak in the correlation. From Figure 9.7 it is easy to see the effect the scale difference might have on the matching process. The simple difference function for this shift of the unknown curvature and the prototype is shown in Figure 9.8. This difference is large even over segments where the two shapes should correspond. This illustrates the necessity of determining the proper time scale.

Now the FFT is used to compute Fourier transforms. Then DC component is set to zero and the magnitude computed. The Fourier transform magnitudes are shown in Figure 9.9 and Figure 9.10. The Mellin transform is computed using the direct method (DMT). The magnitude of the Mellin transform for both the unknown and prototype Fourier magnitudes are shown together in Figure 9.11. It is clear that the Mellin transform magnitudes of the two Fourier magnitudes are very similar. Now the Mellin transforms are multiplied. The cross correlation is obtained by taking a DFT of the resulting sequence. The magnitude of this cross correlation is shown in Figure 9.12. Notice that the peak has shifted from the origin corresponding to the logarithm of the scale change. This scale is used to resample the unknown. The scaled

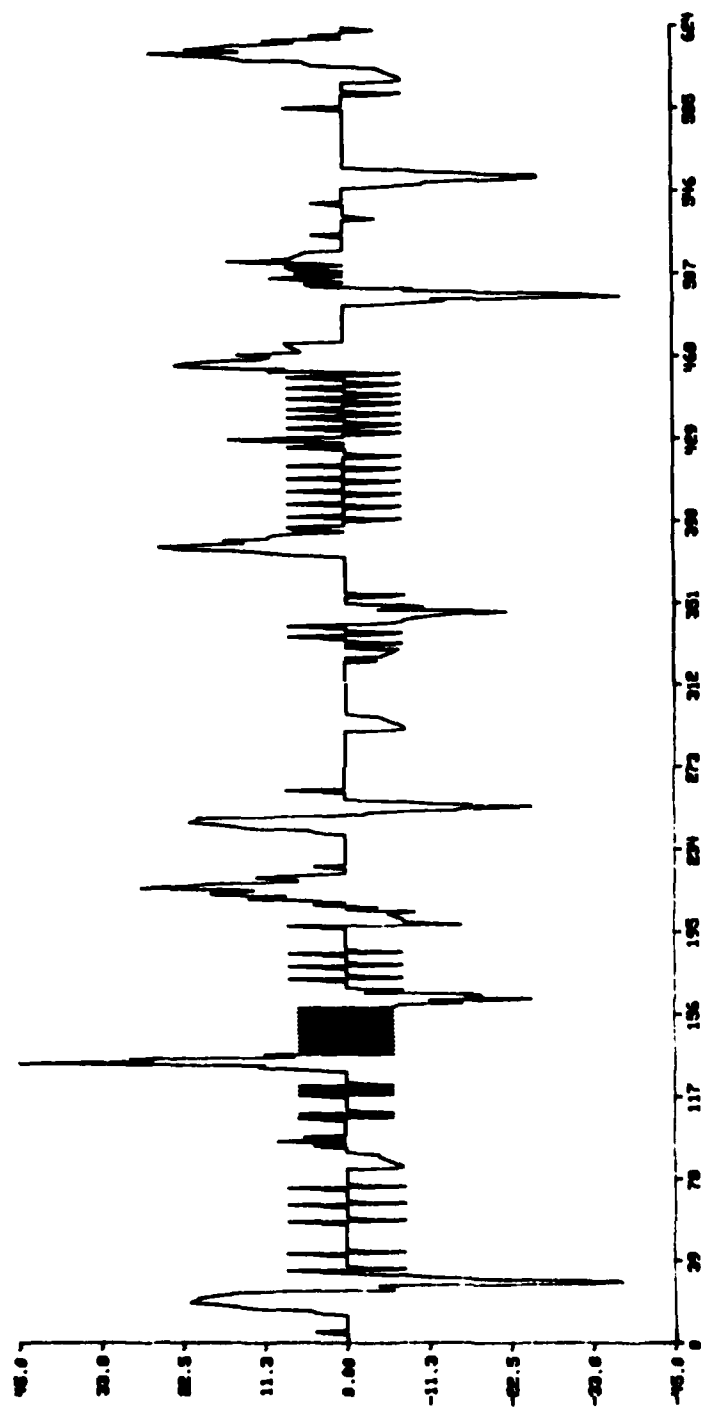


Figure 9.2 Original F104 prototype curvature function, no filtering.

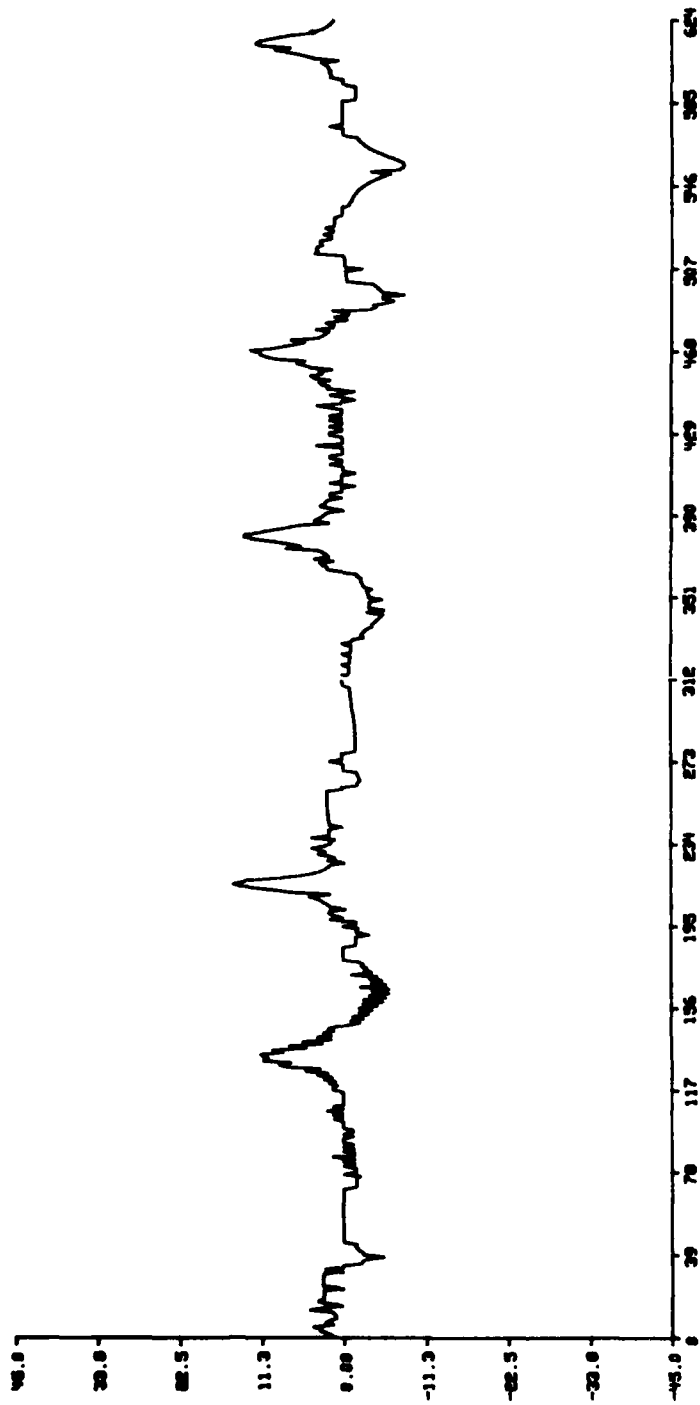


Figure 9.3 F104 prototype curvature function, 30 sample filter.

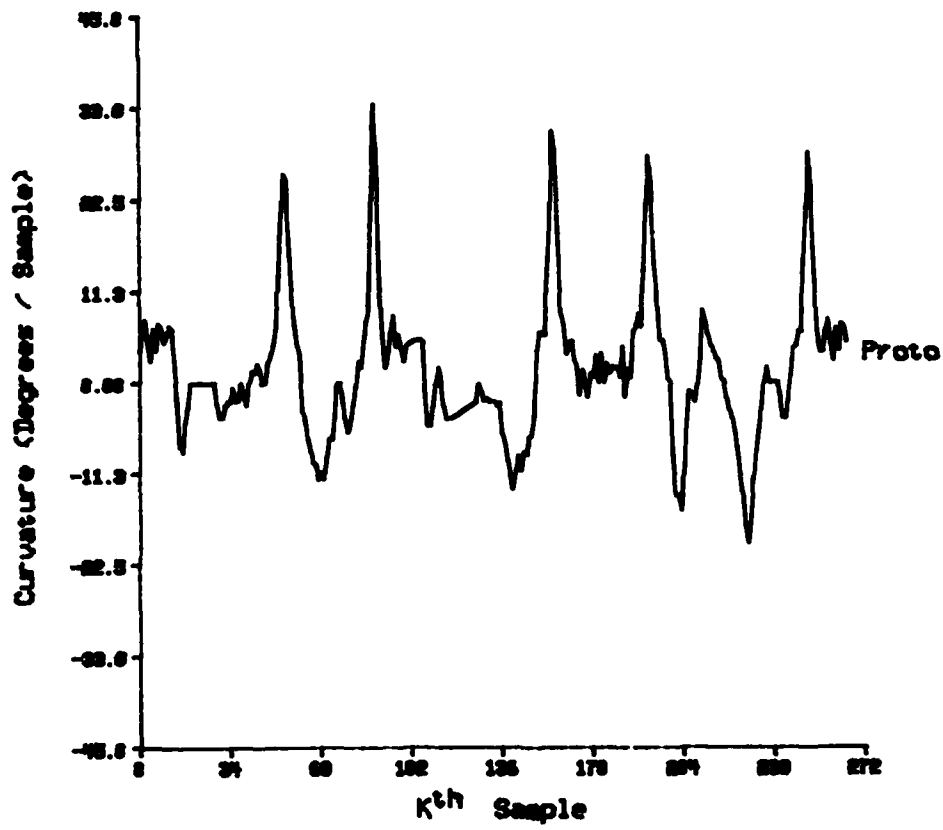


Figure 9.4 F104 prototype curvature function filtered and resampled to 256 samples.

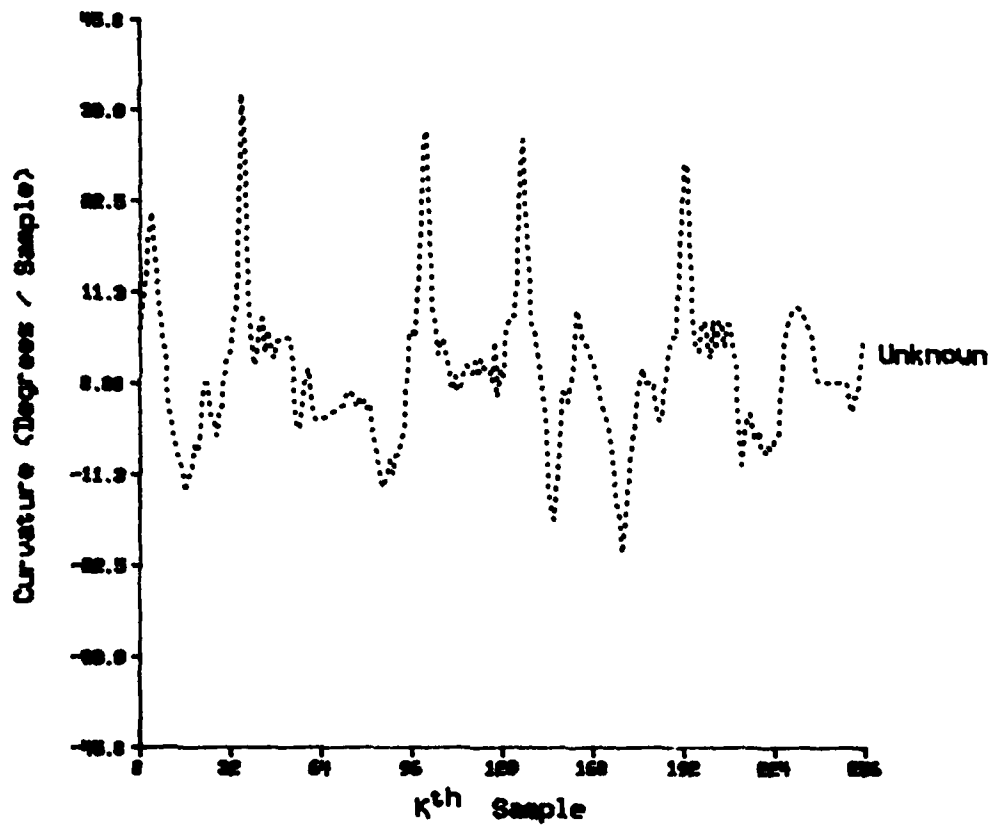


Figure 9.5 10% chopped F104 shifted curvature function (filtered and resampled).

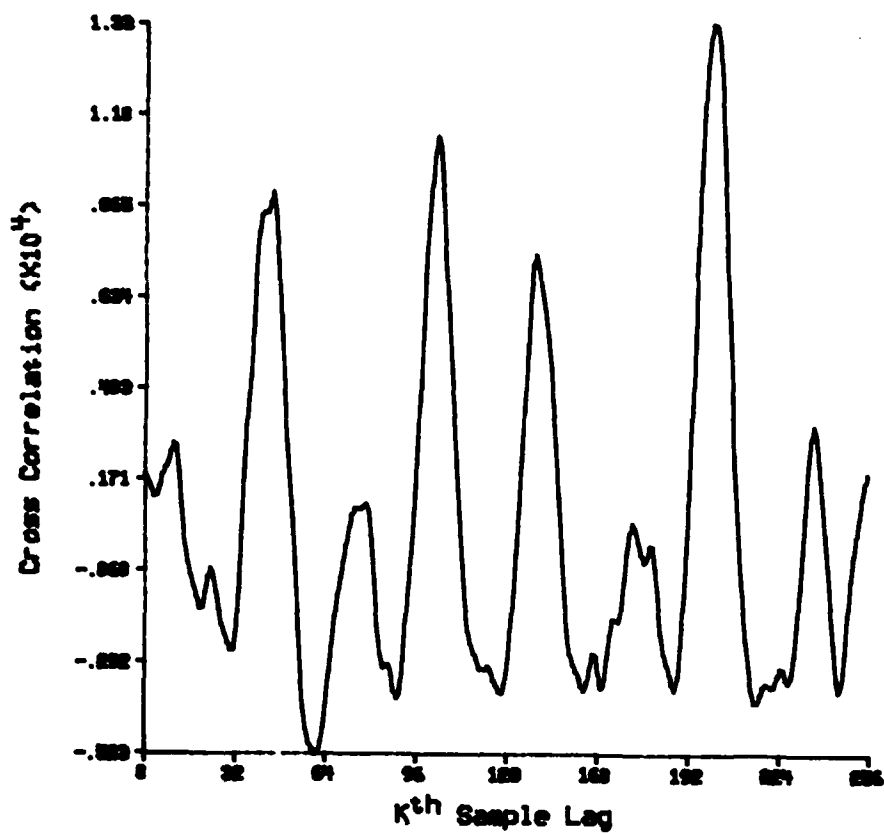


Figure 9.6 Correlation function for prototype and unknown curvature functions.

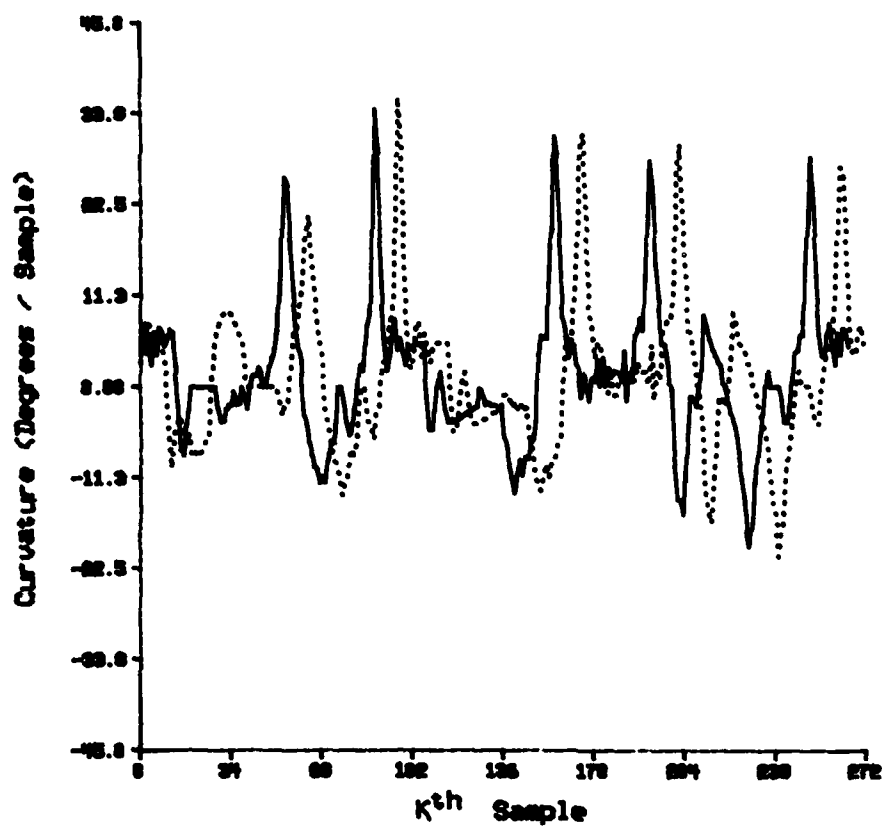


Figure 9.7 Prototype (solid) and shifted unknown (dotted) curvature functions.

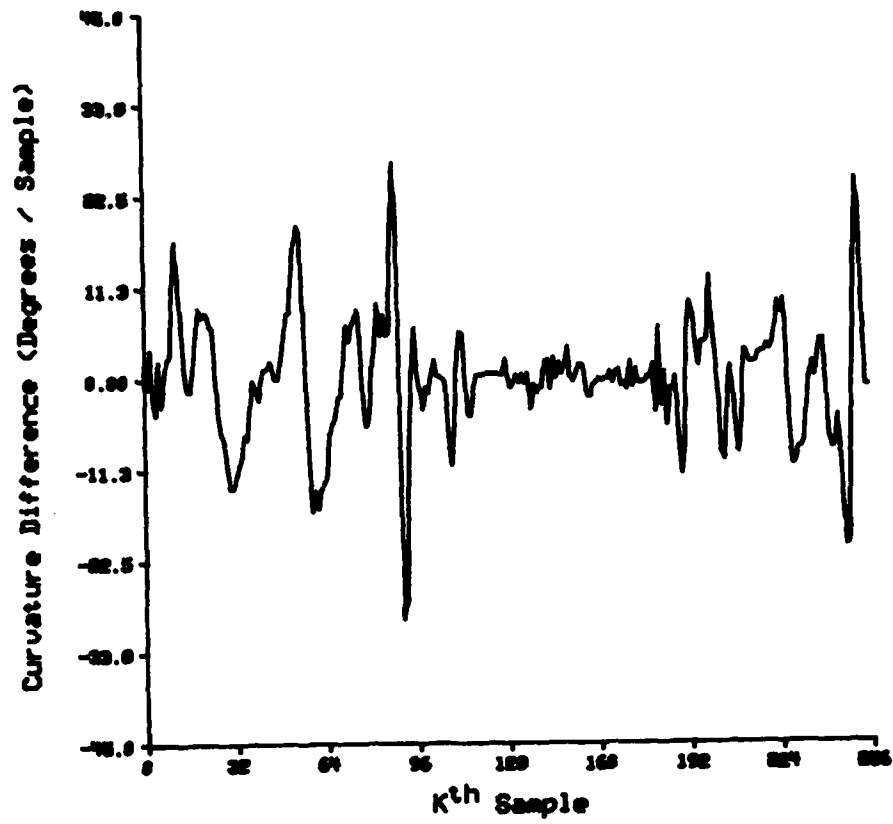


Figure 9.8 Difference function for prototype and shifted unknown curvature functions.

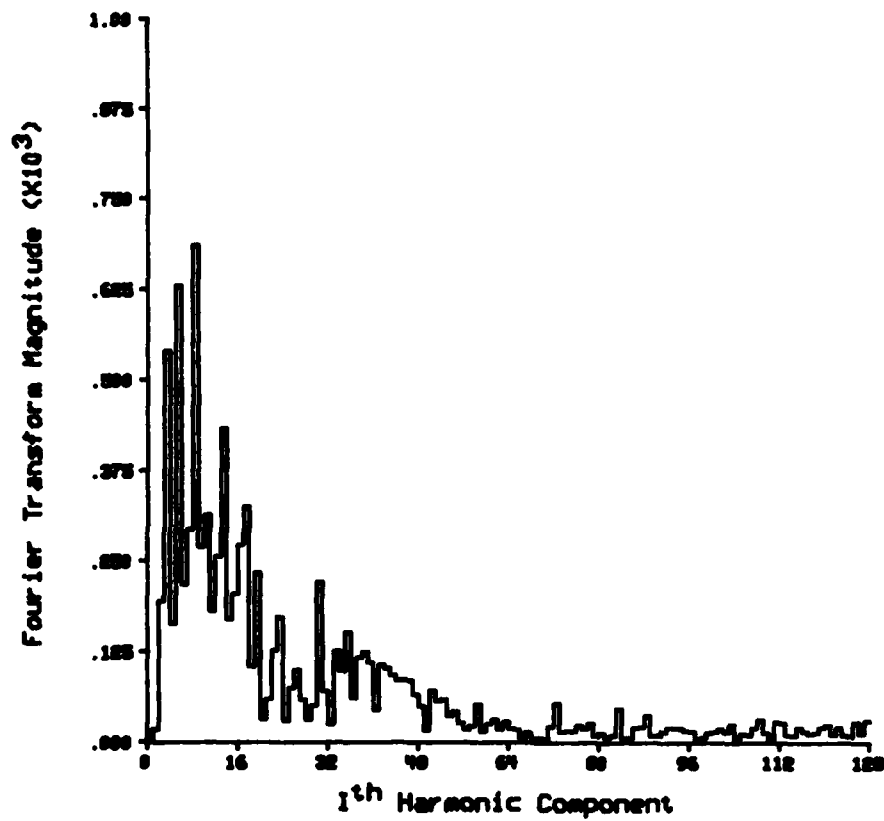


Figure 9.9 Magnitude of the FFT for the prototype curvature function.

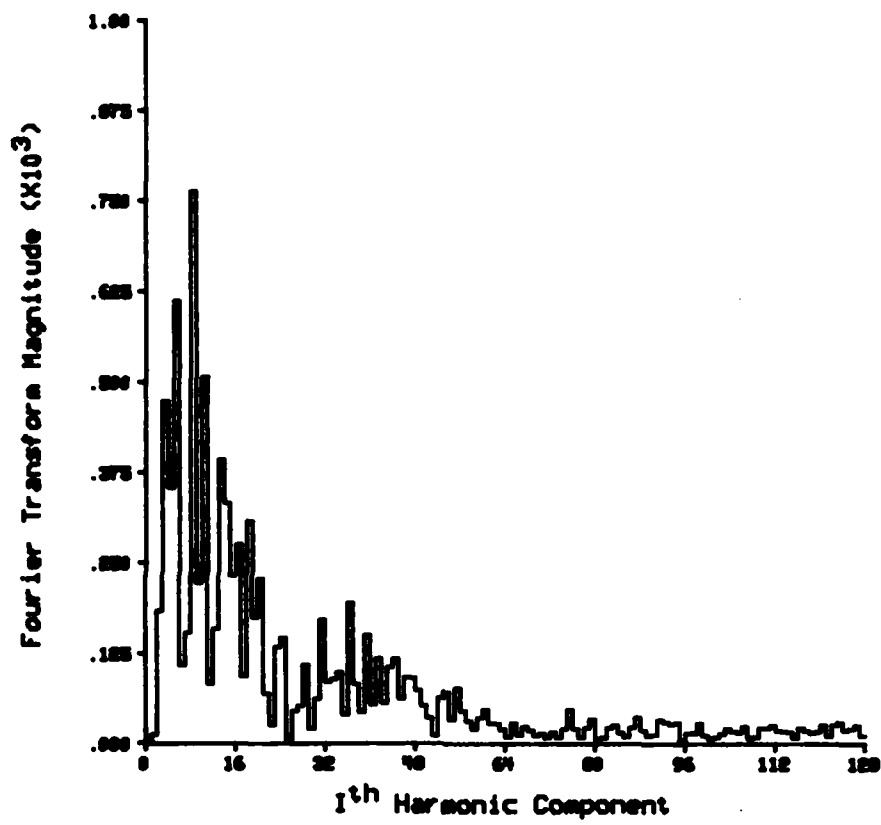


Figure 9.10 Magnitude of the FFT for the unknown curvature function.

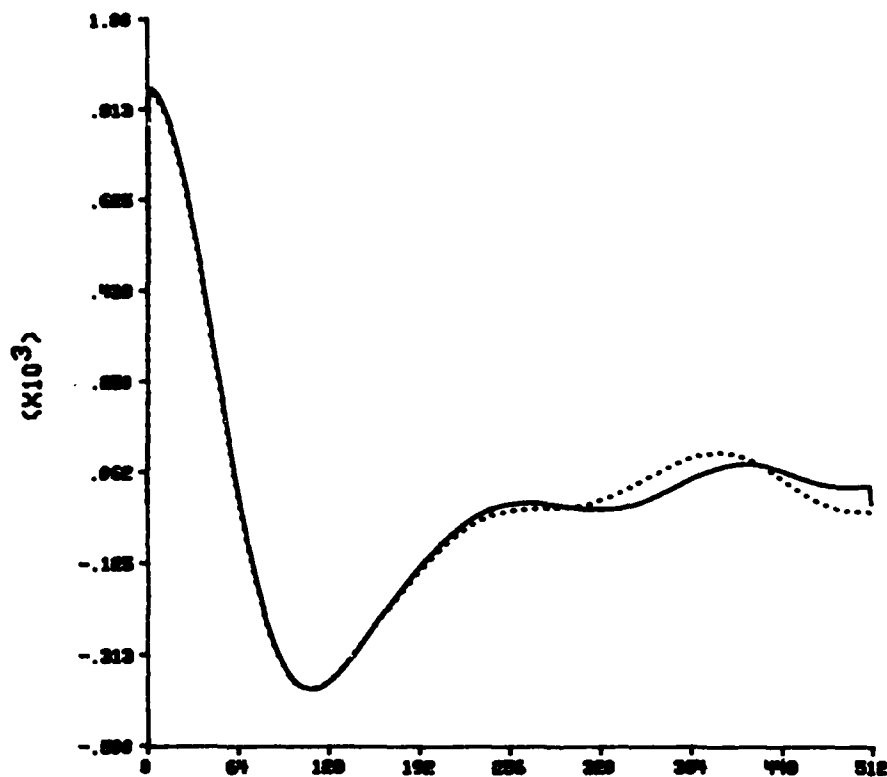


Figure 9.11 Mellin transform (DMT) magnitude of the FFT magnitude, prototype (solid) and unknown (dotted).

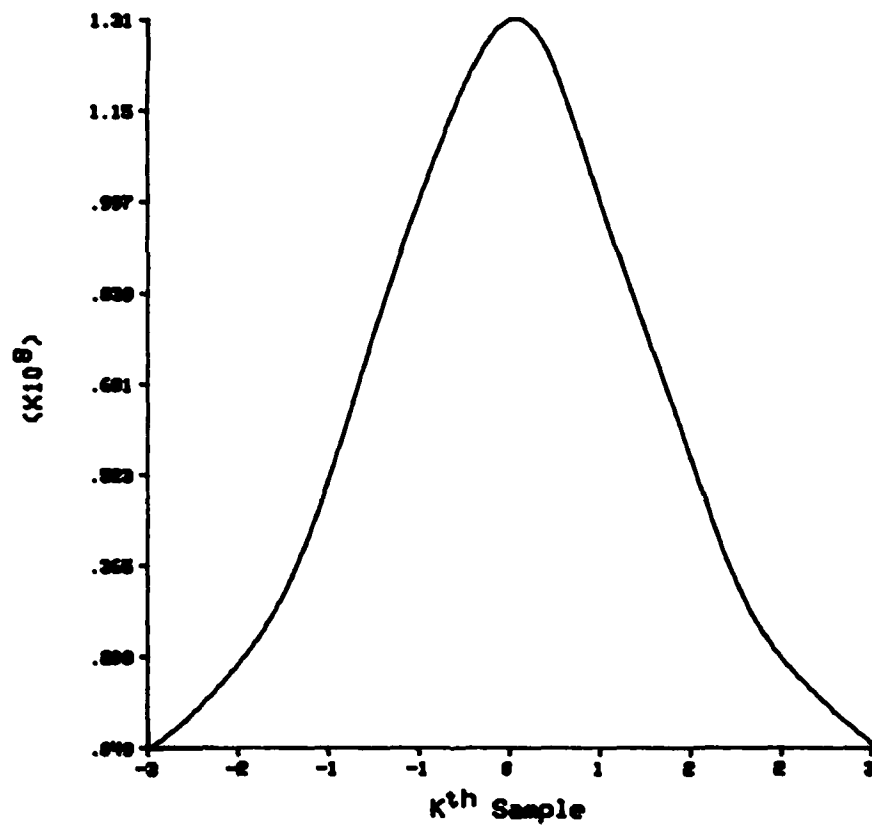


Figure 9.12 Fourier-Mellin cross correlation function.

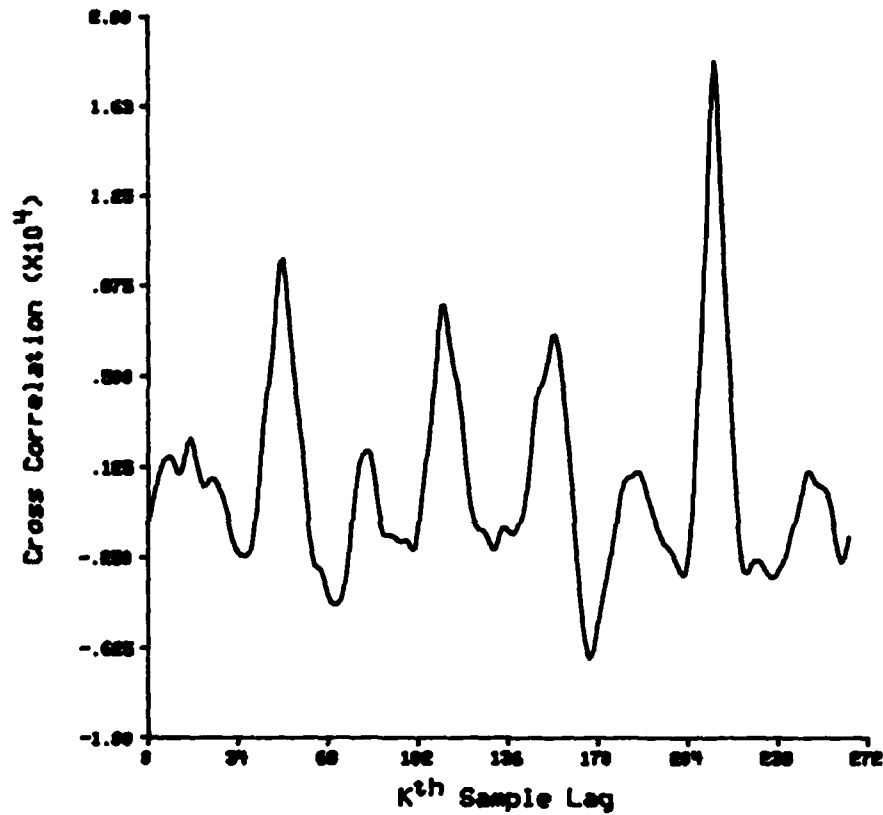
*P. 229
Blank*

Figure 9.13 Correlation function for prototype and scaled unknown curvature functions.

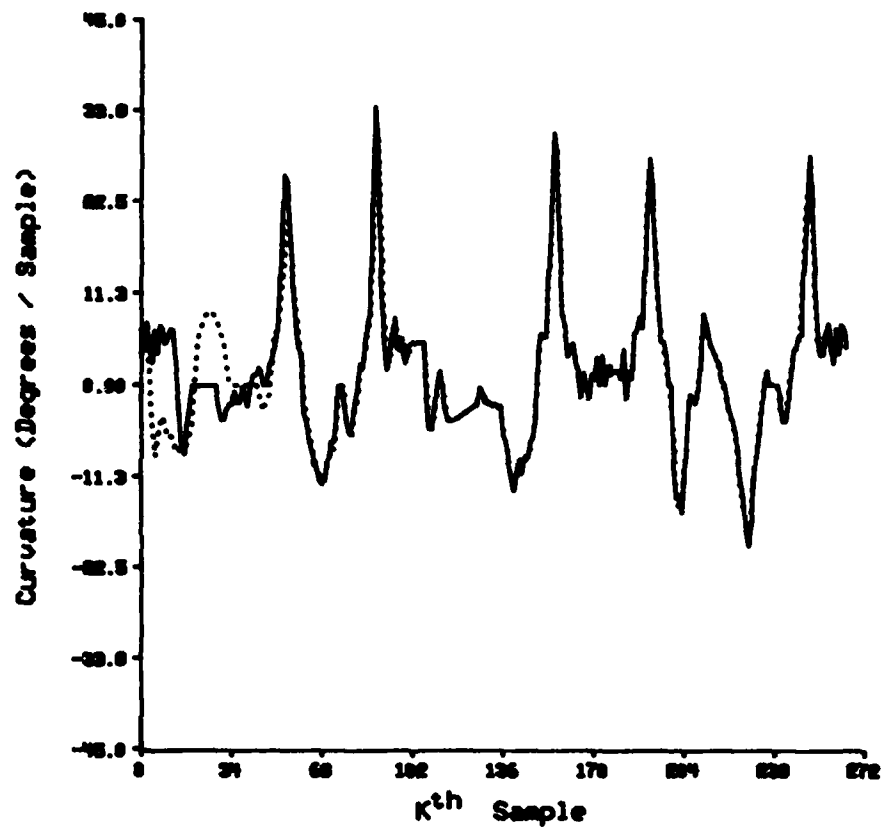


Figure 9.14 Prototype (solid) and scaled & shifted unknown (dotted) curvature functions.

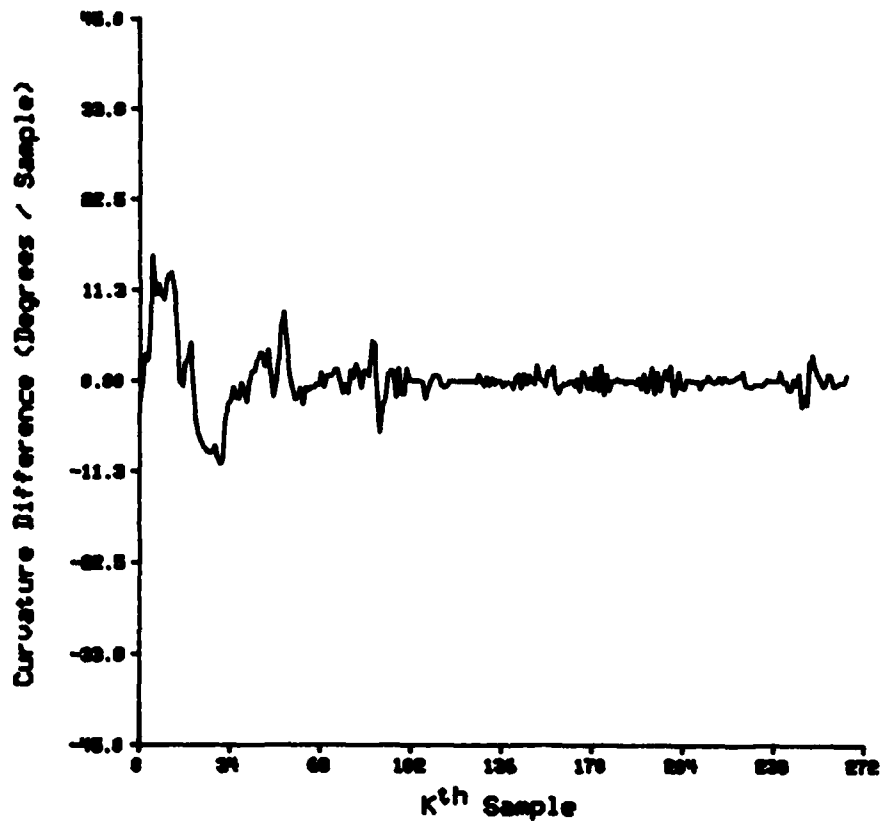


Figure 9.15 Difference function for prototype and scaled & shifted unknown curvature functions.

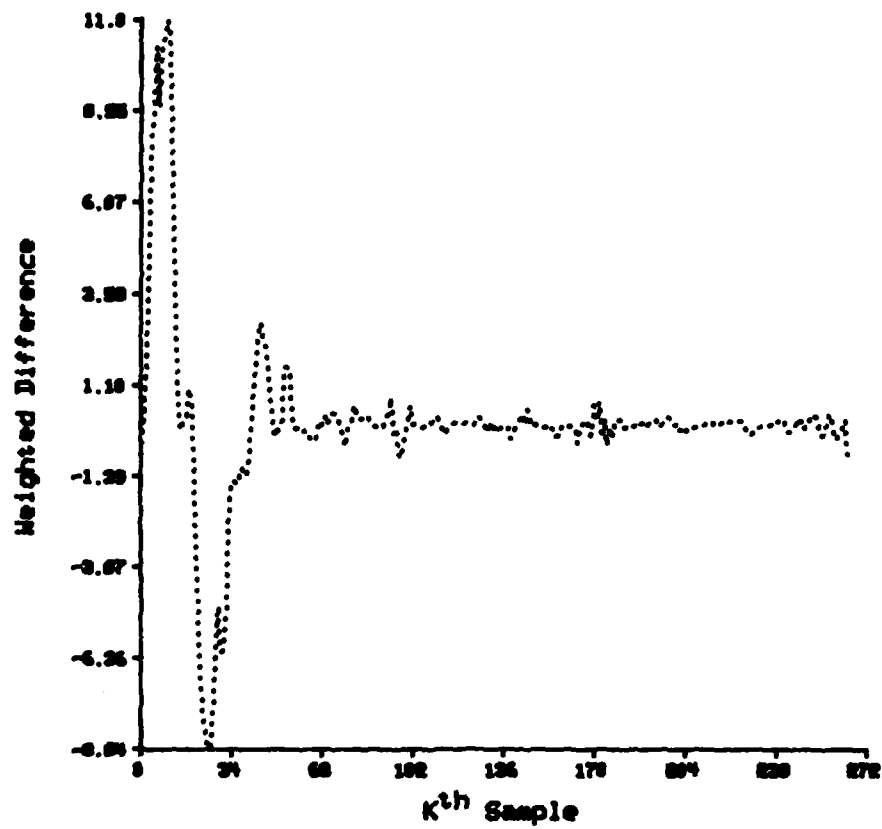


Figure 9.16 Weighted difference function.

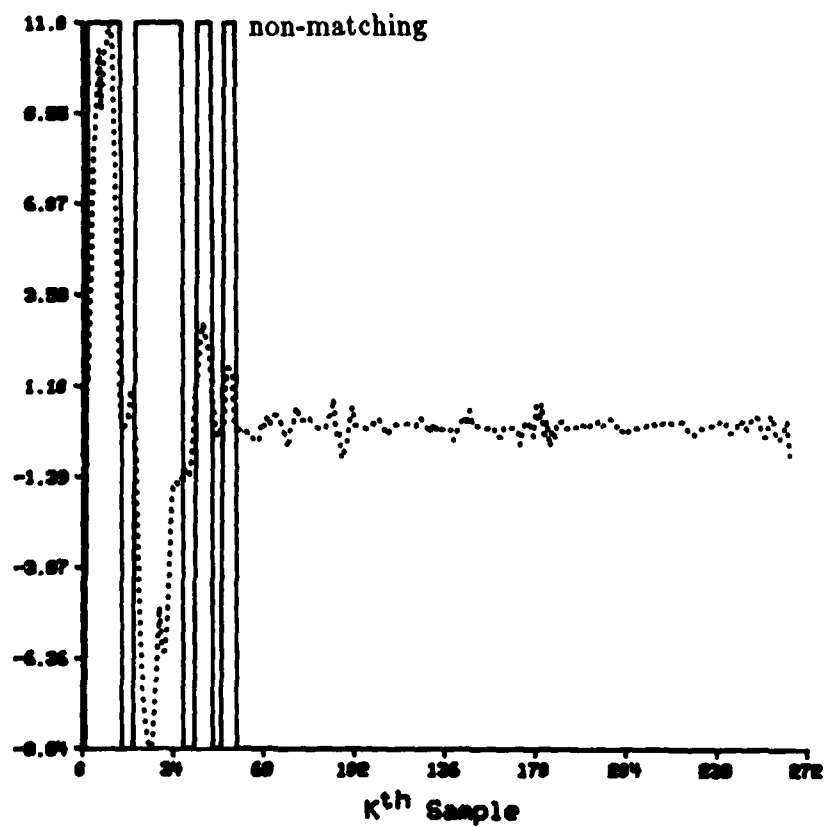


Figure 9.17 Weighted difference function (dotted) with segmentation (solid).

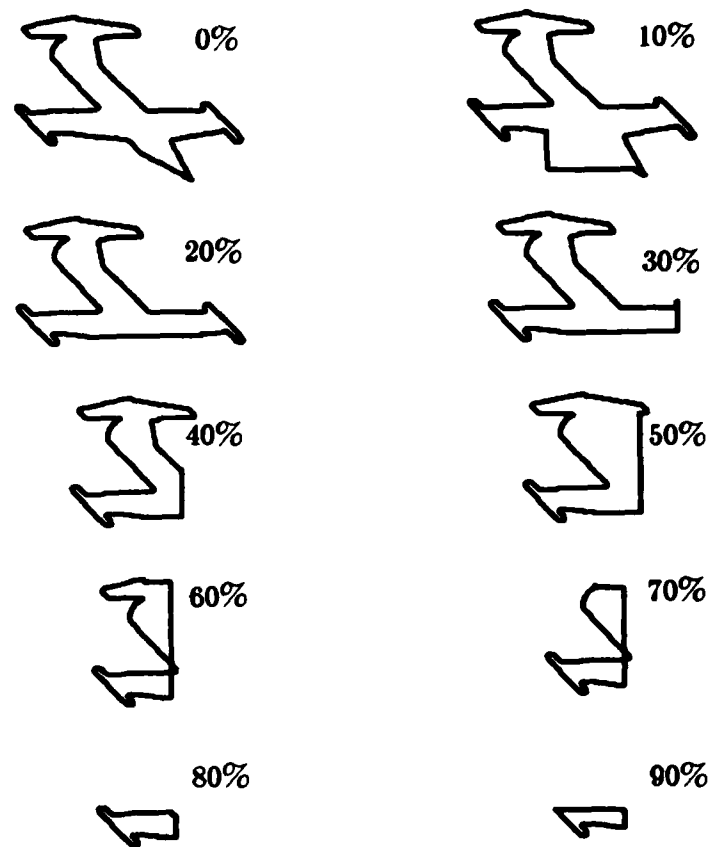


Figure 9.18 Side view of F104 chopped 0, 10, 20, 30, 40, 50, 60, 70, 80, and 90%.

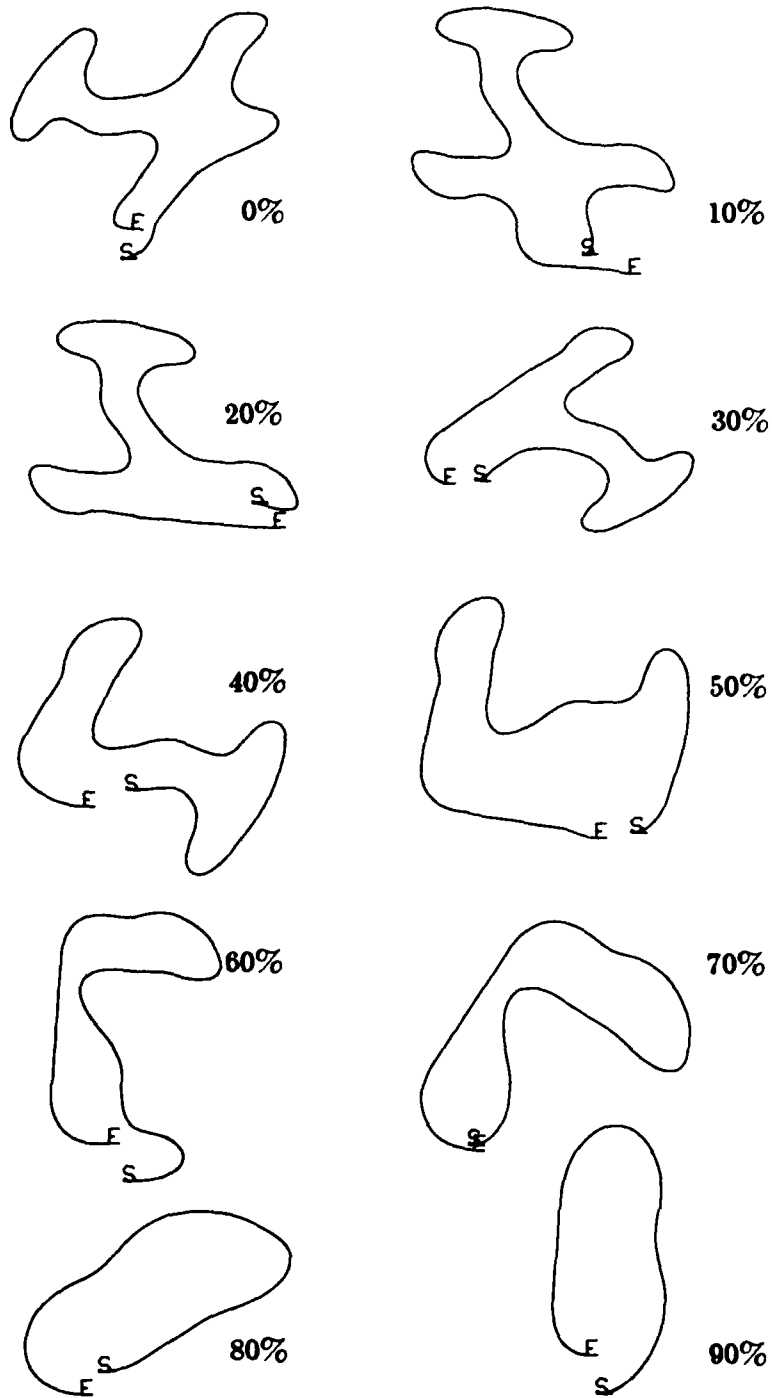


Figure 9.19 Contours reconstructed from filtered & resampled curvature functions (0 - 90 % chopped).

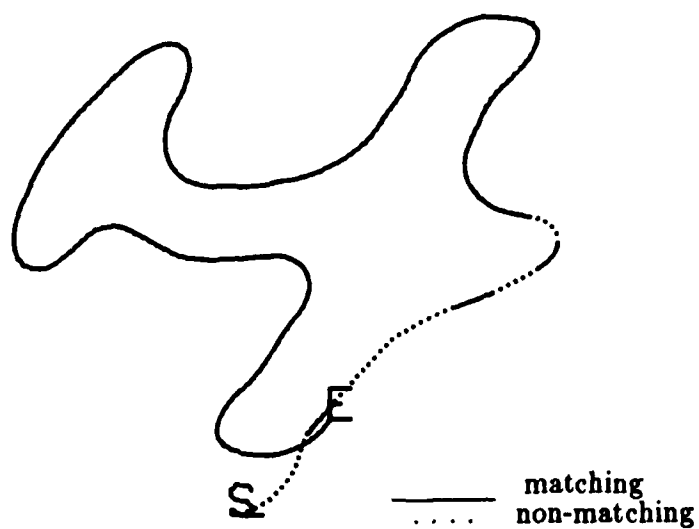


Figure 9.20 Segmented reconstructed contour, 10% chopped.

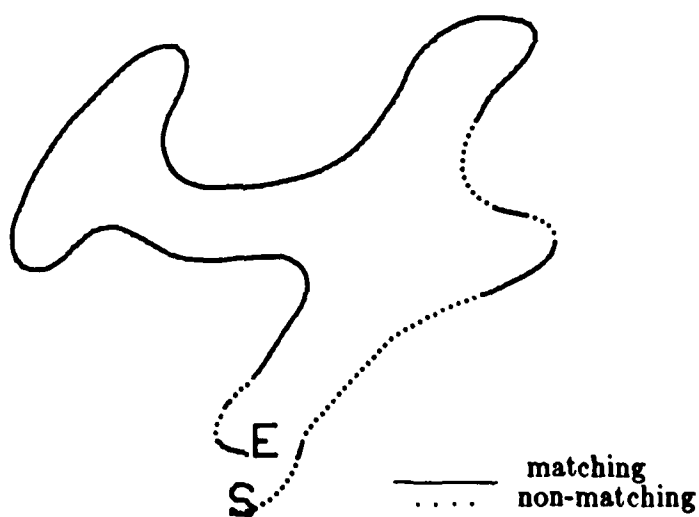


Figure 9.21 Segmented reconstructed contour, 20% chopped.

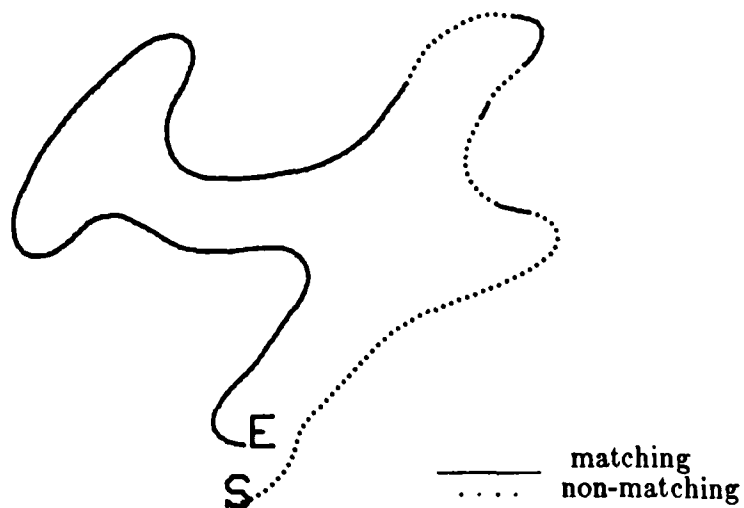


Figure 9.22 Segmented reconstructed contour, 30% chopped.

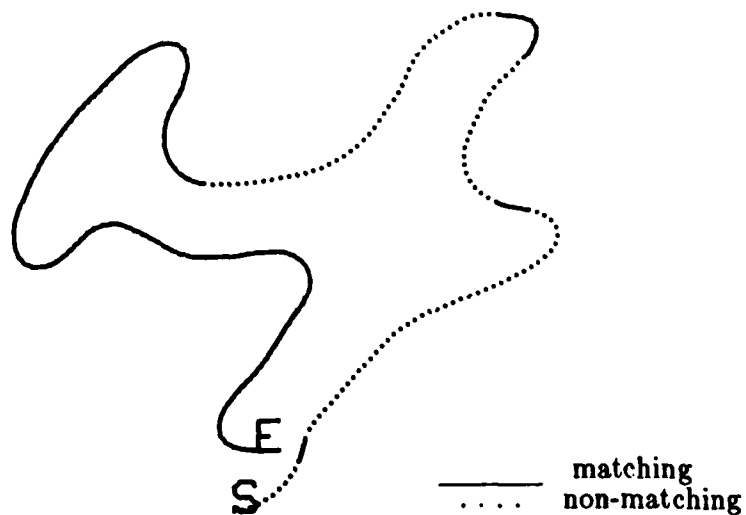


Figure 9.23 Segmented reconstructed contour, 40% chopped.

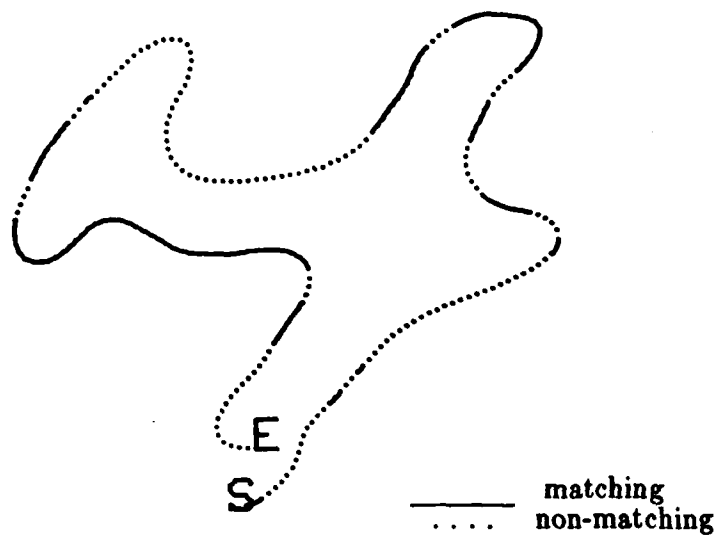


Figure 9.24 Segmented reconstructed contour, 50% chopped.

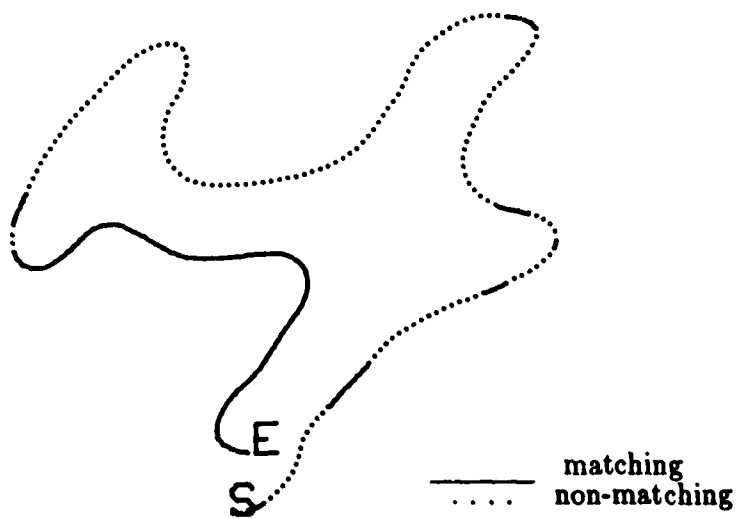


Figure 9.25 Segmented reconstructed contour, 60% chopped.

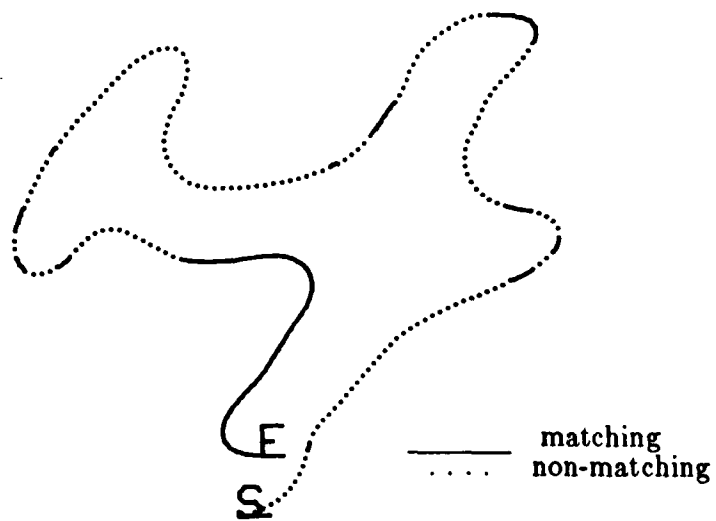


Figure 9.26 Segmented reconstructed contour, 70% chopped.

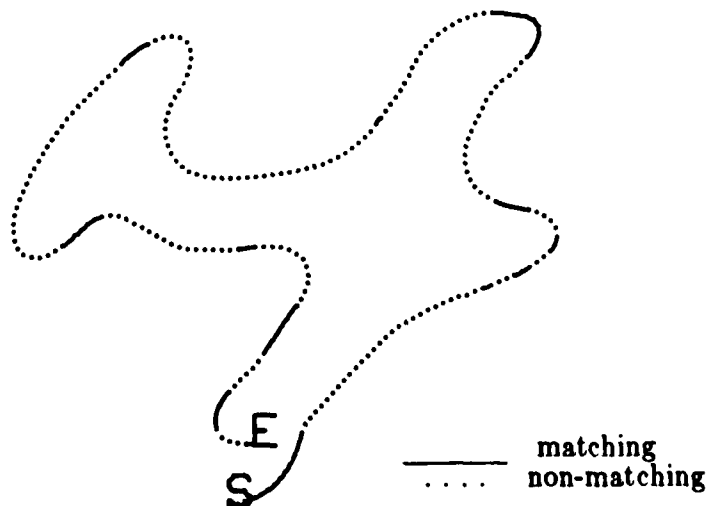


Figure 9.27 Segmented reconstructed contour, 80% chopped.

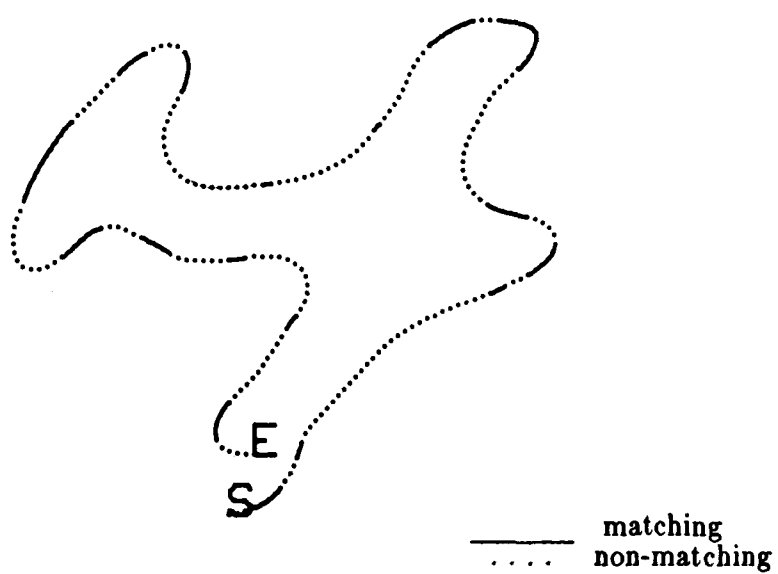


Figure 9.28 Segmented reconstructed contour, 90% chopped.

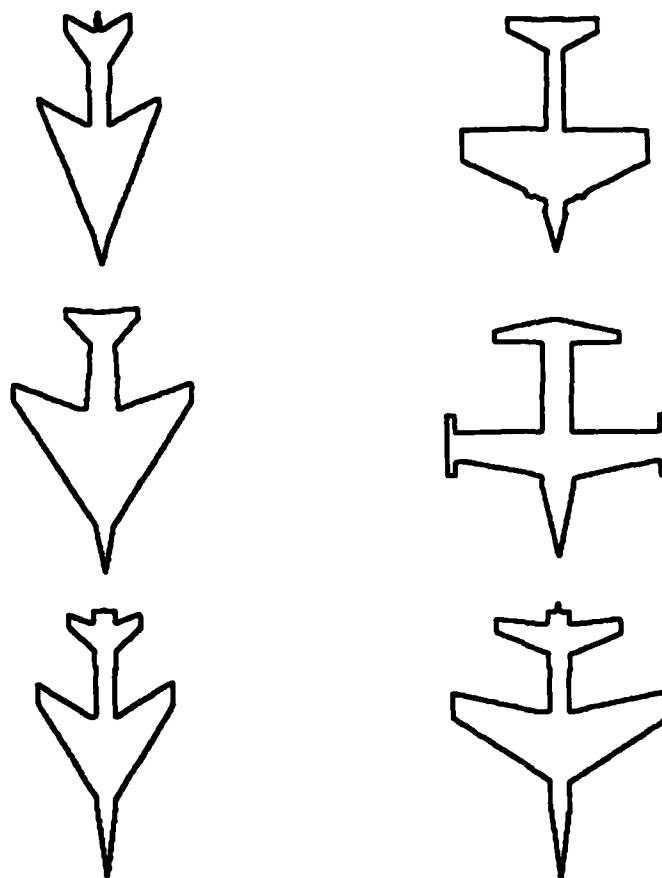


Figure 9.29 Overhead view of six aircraft.

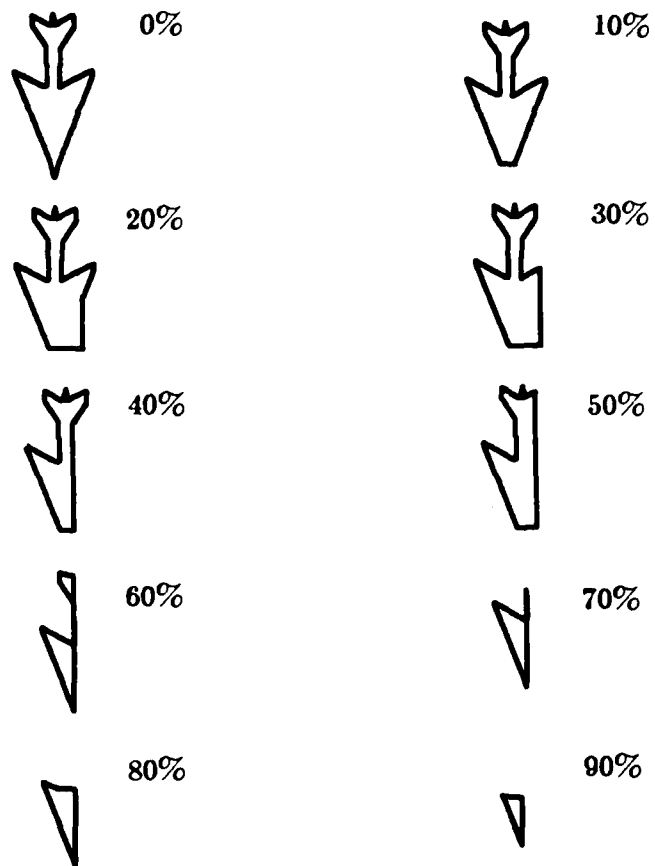


Figure 9.30 MIG overhead view chopped 0% to 90%.

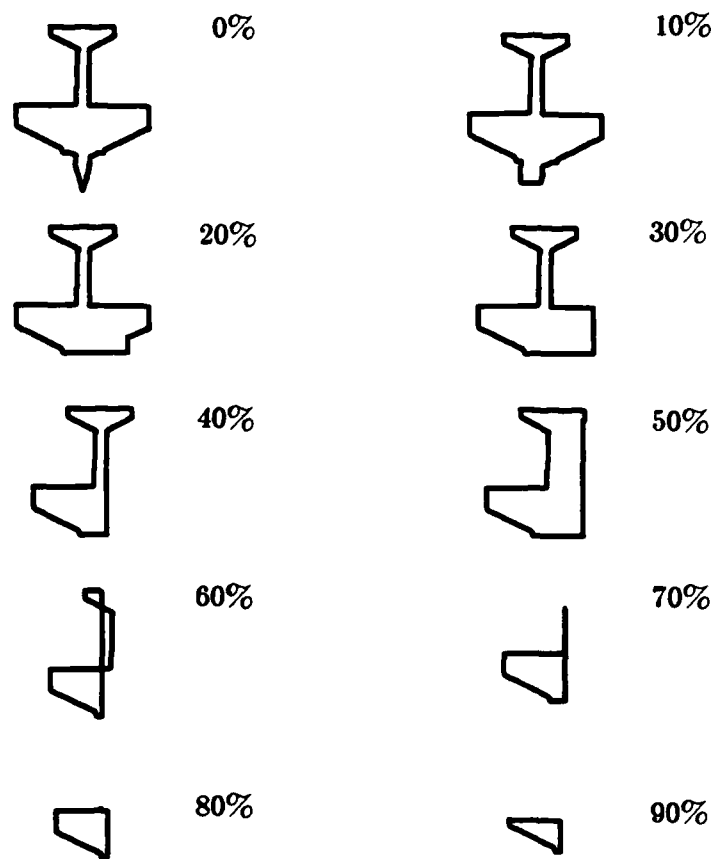


Figure 9.31 B57 overhead view chopped 0% to 90%.

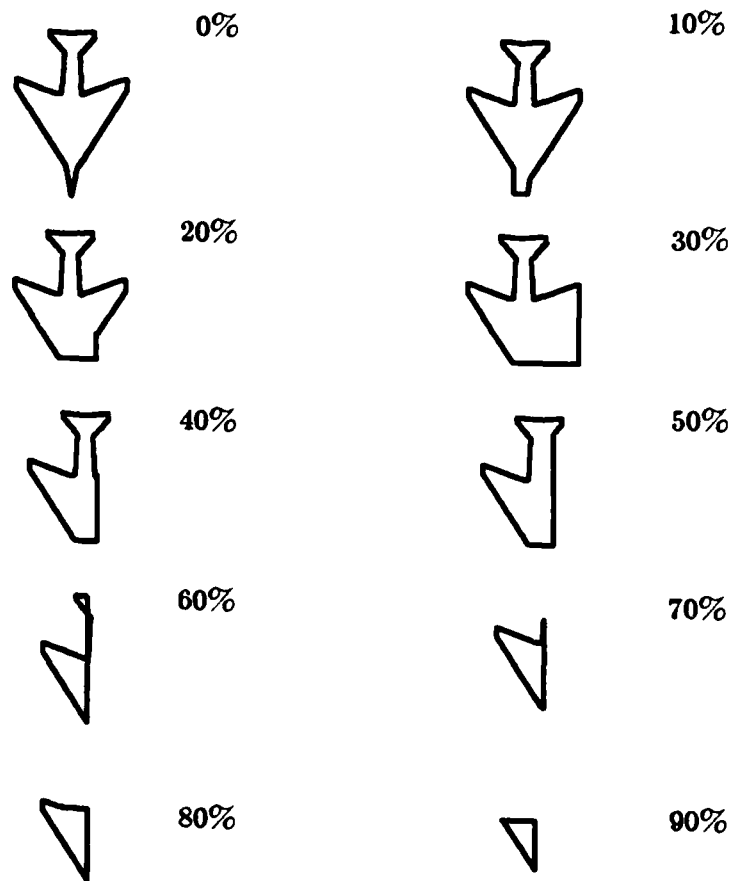


Figure 9.32 Phantom overhead view chopped 0% to 90%.

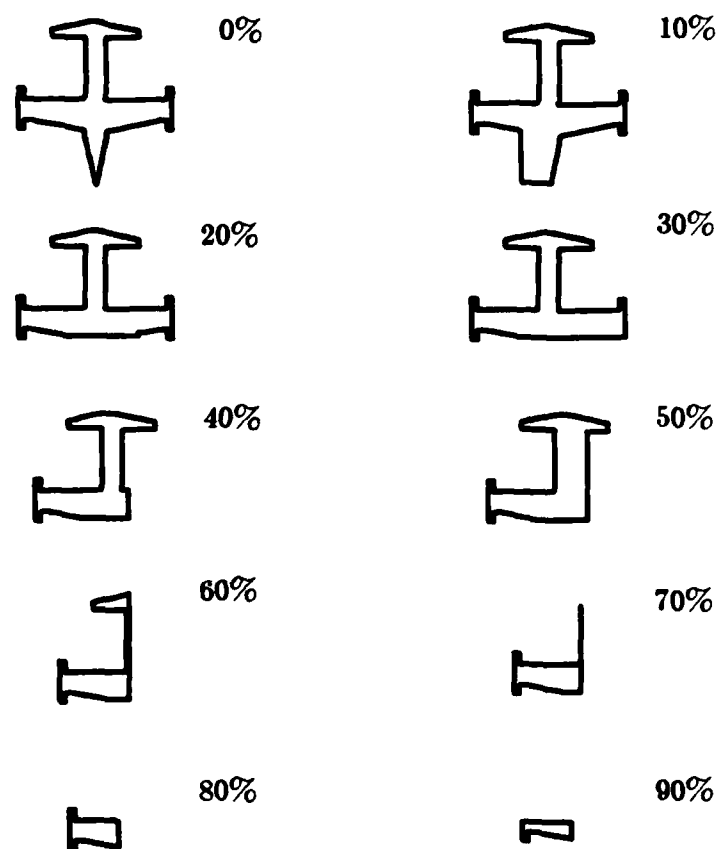


Figure 9.33 F104 overhead view chopped 0% to 90%.

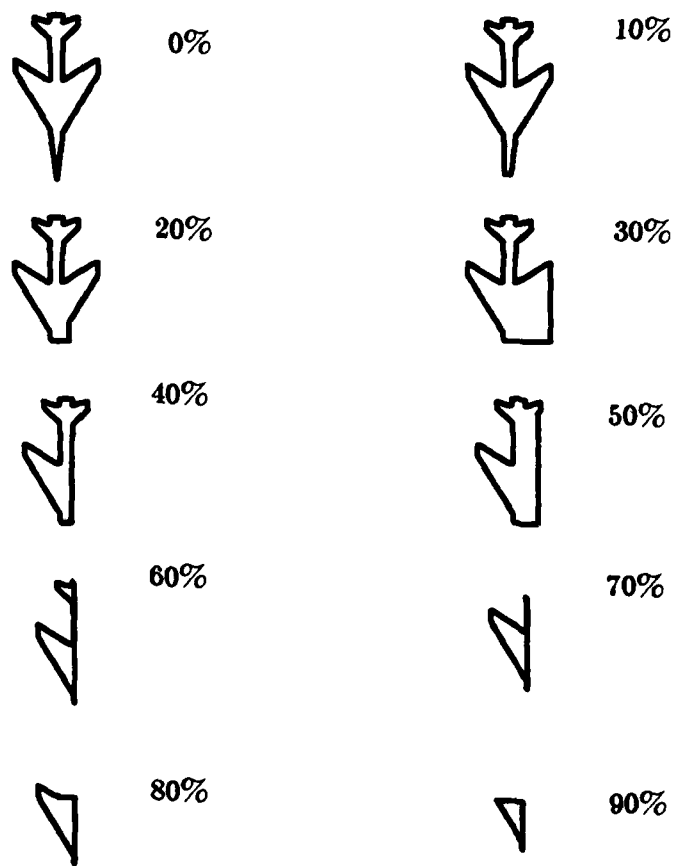


Figure 9.34 F105 overhead view chopped 0% to 90%.

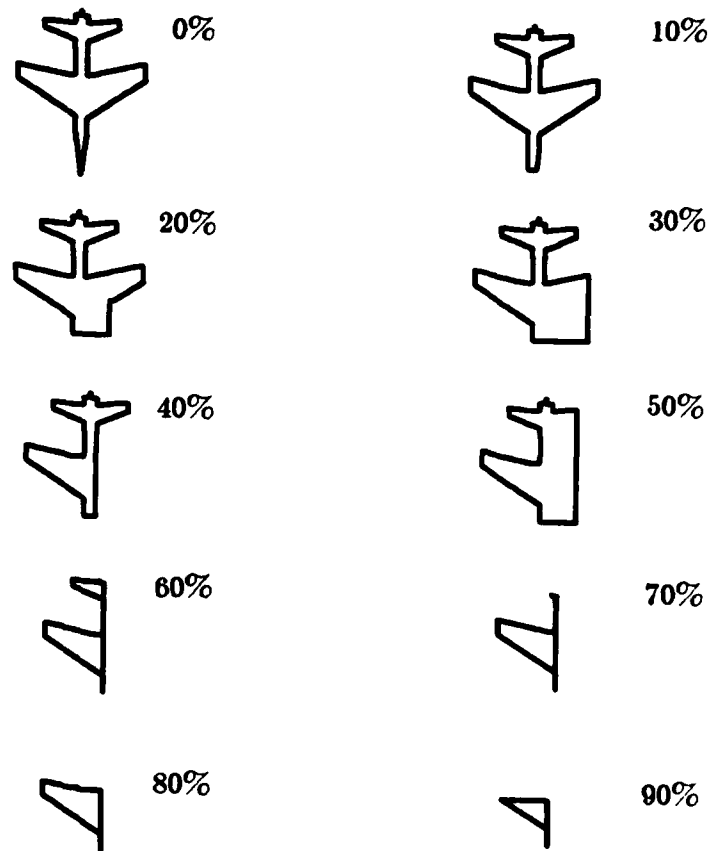


Figure 9.35 Mirage overhead view chopped 0% to 90%.

Table 9.1 **Fourier-Mellin partial shape method classification results for overhead views.**

Percent Chopped	Number Correctly Classified (out of six)
0	6
10	6
20	5
30	4
40	5
50	6
60	2
70	1
80	2
90	2

9.6 Conclusions

This transform method attempts to overcome the many problems previously encountered in performing partial shape recognition. A large classification experiment would be needed to verify its capabilities. The method appears to perform well with up to 50% of the contour chopped. There are several important questions concerning the use of this method that have yet to be addressed. Some of these are a) How many Fourier and Mellin components are necessary to still obtain adequate results? b) What are the computation and storage tradeoffs with respect to the other methods available? c) What information is lost in taking the modulus of the Fourier transform of the curvature function? d) What type of error norm should be used to determine a match? Once the scale and starting point alignment has taken place using the Fourier-Mellin transform technique, then some of the other matching techniques, such as dynamic time warping, might be used to improve the classification results. The introduction of this method provides the possibility of using transform methods to partial shape recognition.

CHAPTER 10

CONCLUSIONS AND RECOMMENDATIONS

The global shape methods using functional approximation have been investigated in several ways. The accumulation of the various properties related to the geometry of a shape have helped to indicate how the methods construct a characterization of the shape. The set of shape recognition experiments provided empirical results useful in determining the effectiveness of the various shape methods in performing shape recognition.

The Fourier descriptors of the boundary performed better than any of the other four methods. Walsh points of the boundary performed almost as well and can be computed quickly. The Walsh points, however, do not have the nice properties of the Fourier descriptors of the boundary. The performance of the Fourier descriptors of the cumulative angular deviant was significantly below that of the Fourier descriptors of the boundary. The performances of the two moment methods were rather poor, far below that of the other three methods just mentioned. There is a numerical precision or dynamic range problem in using the moments methods that has yet to be overcome. Obtaining the moments of the silhouette is really a two-dimensional transform. Much of the information contained in the moment set is necessary to approximate the binary nature of the silhouette rather than extracting the shape characteristics. The moments, however, can be used when there is more than one connected component to consider as a single shape.

Another important issue is the representation problem for a three-dimensional rigid body. The questions of how many and which views are adequate to represent the shape have yet to be answered. Some empirical data for some complex objects (aircraft) has been presented by Charpentier and Glenn [CHAR81,GLEN82]. In general, the theory necessary to solve these problems is not available. However, it has been determined that two non-parallel views are adequate to represent a quadric surface [GROS83]. A related problem is the following: How many different directions are necessary to illuminate the entire boundary of a convex body with a bundle of (parallel) rays? It has been asserted that 8 or less different views are needed to illuminate the surface of a three-dimensional convex body [BOLT80]. This illumination of the surface is similar to the operation of producing a two-dimensional range image of a three-dimensional convex object. It is also known that the intersection of the back projections of the simply connected silhouettes forms a convex hull for the object.

Another possible way to formulate this problem is with the use of the Radon transform [GELF66]. Let $I_{\theta\phi}(x',y')$ be the Radon transform of the three-dimensional rigid body, O , with a constant density, ρ , for the parallel bundle of rays with a direction (θ,ϕ) . The silhouette is then the thresholded image,

$$I_{\theta\phi}'(x',y') = g(I_{\theta\phi}(x',y'))$$

where $g(t) = 1$, if $t \geq 0$ and $g(t) = 0$, if $t < 0$. This thresholded image is then used to reconstruct the object. If the thresholding had not taken place, this would be ordinary reconstruction from projections. So, the question becomes this: Under what conditions can the function $f(x,y,z)$ be approximated to an

acceptable degree of accuracy when reconstructed from a limited number of thresholded views given the additional constraints that the object is of bounded support and is of constant density?

The representation problem is important for practical reasons, since having the smallest possible number of library views would reduce the storage requirements and also reduce the time necessary to determine a match.

The Fourier-Mellin technique introduced has the potential of solving some of the problems that make it difficult to recognize partial shapes. This method was shown to be capable of recognizing shapes with up to 50% of the contour in error. A larger partial shape recognition experiment would help verify the method's usefulness. The technique of time shift and scale invariant correlation should have applications in many areas where part of a signal is seriously degraded but a match to a prototypical signal is desired. More research is also needed to determine the best segmentation procedure and distance function for the matching process. More investigations are also needed to determine how many Fourier and Mellin components are actually necessary to obtain an adequate level of performance.

REFERENCES

- [ALTE78] Richard A. Altes, "The Fourier-Mellin transform and mammilian hearing," *Journal Acoustical Society of America*, Vol. 63, No. 1, pp. 174-183, January, 1978.
- [ATTN66] H. W. Hake, "Form discrimination and the invariance of form," *Pattern Recognition*, Uhr(ed.), pp. 123-141, Wiley, 1966.
- [BAUD73] Patrick Baudelaire, "Linear Stretch-Invariant Systems," *Proceedings IEEE*, Vol. 61, No. 4, pp. 467-468, April, 1973.
- [BENN75] John R. Bennett and John S. Mac Donald, "On the Measurement of Curvature in a Quantized Environment," *IEEE Trans. on Computers*, Vol. C-24, No. 8, pp. 803-820, August, 1975.
- [BOLT80] Vladimir Grigor'evich Boltyanskii and Izrail' Tsudikovich Gohberg, *The Decomposition of Figures into Smaller Parts*, University of Chicago Press, Chicago, 1980.
- [CASA76a] David Casasent and Demetri Psaltis, "Scale Invariant Optical Transform," *Optical Engineering*, Vol. 15, No. 3, pp. 258-261, May-June, 1976.
- [CASA76b] David Casasent and Demetri Psaltis, "Position, rotation, and scale invariant optical correlation," *Applied Optics*, Vol. 15, No. 7, pp. 1795-1799, July, 1976.
- [CASA77] David Casasent and Demetri Psaltis, "New Optical Transforms for Pattern Recognition," *Proceedings IEEE*, Vol. 65, No. 1, pp. 77-84, January, 1977.
- [CASA78] David Casasent and Mark Kraus, "Polar Camera for Space-variant Pattern Recognition," *Applied Optics*, Vol. 17, No. 10, pp. 1559-1561, 15 May, 1978.

- [CHAR81] Didier Jean Charpentier, "Three Dimensional Shape Recognition Using Fourier Descriptors and Multi-Dimensional Chain-code," *M.S.E.E. Thesis*, Purdue University, August, 1981.
- [CRIM82] Thomas R. Crimmins, "A Complete Set of Fourier Descriptors for Two-Dimensional Shapes," *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. SMC-12, No. 6, November/December, 1982.
- [DUDA77] S. A. Dudani, "Aircraft identification by moment invariants," *IEEE Trans. Computers*, vol. C-26, pp. 39-46,
- [FU82] K. S. Fu, *Syntactic Pattern Recognition and Applications*, Englewood Cliffs, NJ, Prentice-Hall, 1982.
- [GELF64] I. M. Gel'fand and G. E. Shilov, *Generalized Functions*, vol. 1, Acedemic Press, N.Y., 1964.
- [GELF66] I. M. Gel'fand, M.I. Graev, and N.Ya. Vilenkin, *Generalized Functions*, vol. 5, Acedemic Press, N.Y., 1966.
- [GIAR77] Charles R. Giardina and Frank P. Kuhl, "Accuracy of Curve Approximation by Harmonically Related Vectors with Elliptical Loci," *Computer Graphics and Image Processing* 6, pp. 277-285, 1977.
- [GIAR78] Charles R. Giardina, "Bounds on the Truncation Error for Walsh Expansions," *Notices of the American Math. Society*, Vol. 25, no. 2, pp. A-311, notice no. 754-B23, February 1978.
- [GIAR83] Charles R. Giardina, Frank P. Kuhl, T. A. Grogan, and O. Robert Mitchell, "A Spatial Domain Walsh Feature Set," *Trans. of the 28th Conference of Army Mathematicians*, ARO Report 83-1, pp. 397-408, February, 1983.
- [GIFF82] Jonathan Paul Gifford, "Classification of Three-Dimensional Partial Shape Using Local Shape Descriptors," *Master's Thesis*, Purdue University, May, 1982.
- [GLEN82] Marcus Elgin Glenn, "Fourier Descriptors of Two and Three Dimensional Objects," *M.S.E.E. Thesis*, Purdue University, August, 1982.

- [GRAD80] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, New York, 1980.
- [GRAN72] G. H. Granlund, "Fourier preprocessing for hand print character recognition," *IEEE Trans. Comput.*, vol. C-21, pp. 195-201, Feb. 1972.
- [GROG83] Timothy A. Grogan and O. Robert Mitchell, "Partial Shape Recognition using Fourier-Mellin Transform Methods," *Optical Society of America Winter '83 Topical Meeting on Signal Recovery and Synthesis with Incomplete Information and Partial Constraints*, pp. ThA19-1 - ThA19-4, January 12-14, 1983 Incline Village, Nevada.
- [GROS83] Charles B. Grosch, "Determination of quadric surfaces from two projected views," *Photogrammetric Engineering and Remote Sensing*, Vol.49, No. 5, pp. 623-628, May, 1983.
- [HOFF78] W. C. Hoffman, "The Lie transformation group approach to visual neuropsychology," *Formal Theories of Visual Perception*, Leeuwenberg and Buffart (eds.), Wiley, N.Y., 1978.
- [HU62] M. K. Hu, "Visual pattern recognition by moment invariants," *IEEE Trans. Inform. Theory*, vol. 8, pp. 179-187, Feb. 1962.
- [KASH81] R. L. Kashyap and R. Chellappa, "Stochastic models for closed boundary analysis: Representation and reconstruction," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 5, Sept. 1981.
- [KOLM75] A. N. Kolmogorov and S. V. Fomin, *Introductory Real Analysis*, English edition translated and edited by Richard A. Silverman, Dover Publications, Inc., New York, 1975.
- [KUHL82] Frank P. Kuhl and Charles R. Giardina, "Elliptic Fourier features of a closed contour," *Computer Graphics and Image Processing*, vol. 18, pp. 236-258, 1982.
- [LAY82] Steven R. Lay, *Convex Sets and their Applications*, John Wiley & Sons, New York, 1982.

- [LAWR72] J. Dennis Lawrence, *A Catalog of Special Plane Curves*, Dover Publications, New York, 1972.
- [MITC81] O. R. Mitchell, A. P. Reeves, and K. S. Fu, "Information Extraction from Aerial Photographs," Final Report, *Subcontract No. SCEE-PDP/80/1*, May, 1981.
- [MITC82a] O. Robert Mitchell, Frank P. Kuhl, Timothy A. Grogan, and Didier Charpentier, "A Shape Extraction and Recognition System," *Southcon/82*, March 23-25, Orlando, Florida, pp.(4/1) 1-4, 1982.
- [MITC82b] O. Robert Mitchell, Anthony P. Reeves, and Timothy A. Grogan, "Algorithms and Architectures for Global Shape Analysis in Time-Varying Imagery," *SPIE Proceedings 360, Robotics and Industrial Inspection*, San Diego, Aug.24-27, 1982.
- [MOEL82] Harold Moellering and John N. Rayner, "The Harmonic Analysis of Spatial Shapes Using Dual Axis Fourier Shape Analysis (DAFSA)," *Geographical Analysis*, Vol. 13, No. 1, pp. 64-77, January, 1981.
- [PAVL78] T. Pavlidis, "A review of algorithms for shape analysis," *Computer Graphics and Image Processing*, vol. 7, pp. 243-258, 1978.
- [PERS77] E. Persoon and K. S. Fu, "Shape discrimination using Fourier descriptors," *IEEE Trans. Syst., Man, Cybern.* vol. SMC-7, pp. 170-179, Mar. 1977.
- [PITT47] Walter Pitts and Warren S. McCulloch, "How we know universals: the perception of auditory and visual forms," *Bulletin of Mathematical Biophysics*, Vol.9, pp. 127-147, 1947.
- [PROF82] D. Proffitt, "Normalization of Discrete Planar Objects," *Pattern Recognition*, Vol. 15, No. 3, pp. 137-143, 1982.
- [REEV81a] Anthony P. Reeves, "The general theory of moments for image shape analysis," *School Elect. Engr. Purdue Univ.*, Technical Report TR-EE 81-37, 1981.

- [REEV81b] Anthony P. Reeves and Abdolrahim Rostampour, "Shape Analysis of Segmented Objects Using Moments," *Proceedings of the IEEE Computer Society Conference on Pattern Recognition and Image Processing*, pp. 171-173, August, 1981.
- [RICH74] C. W. Richard, Jr. and H. Hemani, "Identification of three-dimensional objects using Fourier descriptors of the boundary curve," *IEEE Trans. Syst., Man, Cybern.* vol. SMC-4, pp. 371-378, July 1974.
- [ROBB72] Gregory M. Robbins and Thomas S. Huang, "Inverse Filtering for Linear Shift-Variant Imaging Systems," *Proceedings of the IEEE*, Vol. 60, No. 7, pp. 862-872, July, 1972.
- [ROSE70] Azriel Rosenfeld, "Connectivity in digital pictures," *Journal of the Assoc. for Comput. Mach.*, vol. 17, No. 1, p. 146, January, 1970.
- [ROSE76] Azriel Rosenfeld and Avinash C. Kak, *Digital Picture Processing*, Academic Press, New York, 1976.
- [SANT76] Luis Antonio Santalo Sors, *Integral Geometry and Geometric Probability*, Addison-Wesley, 1976.
- [TANG82] Gregory Y. Tang, "A Discrete Version of Green's Theorem," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-4, No. 3, pp. 242-249, May, 1982.
- [TEAG80] M. R. Teague, "Image analysis via the general theory of moments," *Journ. Optical Soc. Am.*, vol. 70, no. 8, pp. 920-930, August, 1980.
- [THOM75] R. Thom, *Structural Stability and Morphogenesis*, W.A. Benjamin, Inc., 1975.
- [TUOM83] David Lee Tuomenoksa, George B. Adams, Howard Jay Siegel, and O. Robert Mitchell, "A Parallel Algorithm for Contour Extraction - Advantages and Architectural Implications," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Arlington, Virginia, June 15-23, 1983.

- [WALL80a] Timothy P. Wallace and P.A. Wintz, "An efficient three-dimensional aircraft recognition algorithm using normalized Fourier descriptors," *Computer Graphics and Image Processing*, vol. 13, pp. 99-126, 1980.
- [WALL80b] Timothy P. Wallace and Owen R. Mitchell, "Analysis of Three-Dimensional Movement Using Fourier Descriptors," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, No. 6, pp. 583-588, November, 1980.
- [WALL81] Timothy P. Wallace, O. Robert Mitchell, and Keinosuke Fukunaga, "Three Dimensional Shape Analysis Using Local Shape Descriptors," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-3, No. 3, pp. 310-323, May, 1981.
- [WALT63] Adriaan Walther, "The Question of Phase Retrieval in Optics," *Optica Acta*, pp. 41-49, Vol. 10, No. 1, January, 1963.
- [YATE47] Robert C. Yates, *Curves and their Properties*, Edwards Brothers, Ann Arbor, 1947.
- [ZAHN72] Charles T. Zahn and Ralph Z. Roskies, "Fourier descriptors for plane closed curves," *IEEE Trans. Computers*, vol. C-21, no.3, Mar. 1972.
- [ZWIC83] Philip E. Zwicke and Imre Kiss, Jr., "A New Implementation of the Mellin Transform and its Application to Radar Classification of Ships," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-5, No. 2, pp. 191-199, March, 1983.