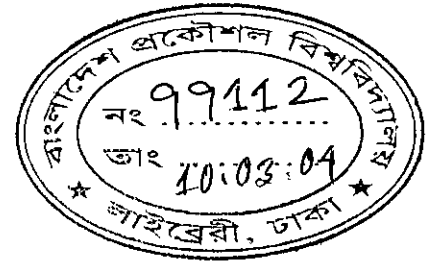


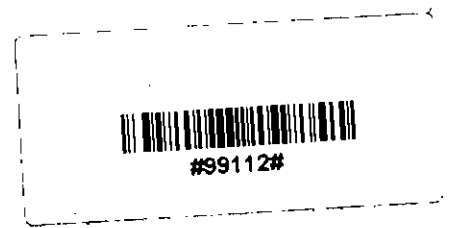
**THREE DIMENSIONAL DISTORTION INVARIANT OPTICAL
PATTERN RECOGNITION USING PROJECTION-SLICE
SYNTHETIC DISCRIMINANT FUNCTION**

by

Mohammad Rafiqul Haider





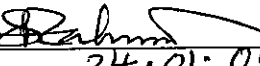
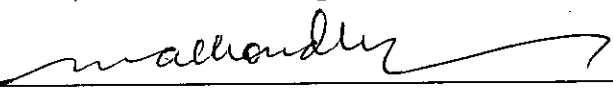

MASTER OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING



DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY
2004

The thesis titled "Three Dimensional Distortion Invariant Optical Pattern Recognition using Projection-slice Synthetic Discriminant Function", submitted by Mohammad Rafiqul Haider, Roll No. 040206250P, Session April 2002, has been accepted as satisfactory in partial fulfillment of the requirement for the degree of MASTER OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING on 24 January 2004.

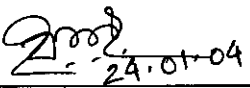
BOARD OF EXAMINERS

1.  24.01.2004
Dr. M. Nazrul Islam
Associate Professor
Department of Electrical and Electronic Engineering
BUET, Dhaka-1000, Bangladesh
Chairman
(Supervisor)
2.  24.1.04
Dr. Md. Abdul Matin
Professor
Department of Electrical and Electronic Engineering
BUET, Dhaka-1000, Bangladesh
Member
3.  24.01.04
Dr. Md. Saifur Rahman
Professor
Department of Electrical and Electronic Engineering
BUET, Dhaka-1000, Bangladesh
Member
4. 
Dr. Mohammad Ali Choudhury
Professor and Head
Department of Electrical and Electronic Engineering
BUET, Dhaka-1000, Bangladesh
Member
(Ex-officio)
5.  24.1.04
Dr. Farruk Ahmed
Professor
Department of Applied Physics and Electronics
University of Dhaka, Dhaka-1000, Bangladesh
Member
(External)

DECLARATION

It is hereby declared that this thesis or any part of it has not been submitted elsewhere for the award of any degree or diploma.

Signature of the Candidate



Mohammad Rafiqul Haider

DEDICATION

To my beloved parents.

CONTENTS

CERTIFICATION	ii
DECLARATION	iii
DEDICATION	iv
TABLE OF CONTENTS	v
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ABBREVIATIONS	xvi
LIST OF SYMBOLS	xvii
ACKNOWLEDGEMENTS	xviii
ABSTRACT	xix
1. INTRODUCTION	1
1.1 Introduction	1
1.2 Pattern Recognition tasks	2
1.2.1 Pattern Association	2
1.2.2 Pattern Classification	3
1.2.3 Pattern Mapping	3
1.2.4 Pattern Grouping	3
1.2.5 Feature Mapping	4
1.2.6 Pattern Variability	4
1.2.7 Temporal Patterns	4
1.2.8 Stability-plasticity Dilemma	5
1.3 Methods of Pattern Recognition Tasks	5
1.3.1 Statistical Pattern Recognition	5
1.3.2 Syntactic Pattern Recognition	6
1.3.3 Neural Pattern Recognition	6
1.4 Implementation Techniques of Pattern Recognition	6

1.5 Optical Pattern Recognition	7
1.5.1 VanderLugt Correlator (VLC)	8
1.5.2 Joint Transform Correlator (JTC)	9
1.6 Distortion Invariant Pattern Recognition	13
1.7 Class Associative Pattern Recognition	14
1.8 Objectives of the Thesis	14
1.9 Thesis Outline	15
2. JOINT TRANSFORM CORRELATION TECHNIQUE	16
2.1. Introduction	16
2.2. Classical Joint Transform Correlation Technique	16
2.2.1 Single Target Detection	18
2.2.1.1. Analysis	18
2.2.1.2. Simulation Results	19
2.2.2. Multiple Target Detection	22
2.2.2.1. Analysis	22
2.2.2.2. Simulation Results	23
2.3. Classical Joint Transform Correlation with Image Subtraction	28
2.3.1. Analysis	28
2.3.2. Simulation Results	29
2.4. Fringe-adjusted Joint Transform Correlation with Image Subtraction	31
2.4.1. Analysis	31
2.4.2. Simulation Results	32
2.5.1 Fractional Power Fringe-adjusted Joint Transform Correlation	36
2.5.2 Simulation Results	37
2.6 Phase-encoded Fringe adjusted Joint Transform Correlation	40
2.6.1 Analysis	40
2.6.2 Simulation Results	44
2.7 Conclusion	46

3 EFFICIENT MULTIPLE TARGET DETECTION

ALGORITHM	48
3.1 Introduction	48
3.2 Multi-target Detection	49
3.2.1 Theoretical Analysis	49
3.2.2 Edge Extraction of Images	52
3.2.3 Simulation Results	53
3.3 Class-associative Target Detection	73
3.3.1. Theoretical Analysis	73
3.3.1.1. Two Objects in a Class	73
3.3.1.2. Three Objects in a Class	76
3.3.1.3. Multiple Objects in a Class	77
3.3.2 Simulation Results	77
3.4. Conclusion	89
4. PROJECTION SLICE SYNTHETIC DISCRIMINANT FUNCTION- BASED JOINT TRANSFORM CORRELATION	
	90
4.1. Introduction	90
4.2. Synthetic Discriminant Function	91
4.3. Projection Slice SDF	91
4.4. Calculation of Weighting Factor	93
4.5. Simulation Results	94
4.5.1. Composite Image Formation using Two Training Images	99
4.5.2. Composite Image Generation for In-plane Rotation	103
4.5.3. Performance Evaluation of In-plane Rotated Composite Image	108
4.5.4. Composite Image Formation for Out-of-plane Rotation	111
4.5.4.1. Composite Image for Azimuthally Rotated Objects	112
4.5.4.2. Performance Evaluation of the Composite Image	115

for Azimuthally Rotated Objects	
4.5.4.3. Composite Image for Change of Elevation Angle of Objects	122
4.5.4.4. Performance Evaluation of the Composite Image for Change of Elevation Angle of Objects	125
4.5.4.5. Performance Evaluation of the Composite Images for 3-D Distortions	129
4.6. Conclusion	140
5. CONCLUSION	141
5.1. Conclusion	141
5.2. Future Works	143
REFERENCES	145

LIST OF TABLES

Table 3.1	Performance comparison for noise free input scene	72
Table 3.2	Performance comparison for noisy input scene	72
Table 4.1	Peak intensity and weight factor after 100 iterations	108
Table 4.2	Peak intensity and weight factor after 100 iterations	115
Table 4.3	Peak intensity and weight factor after 50 iterations	122
Table 4.4	Detection performance for 3D distorted images	131
Table 4.5	Detection performance of 3D distorted objects	134
Table 4.6	Detection performance of 3-D distorted objects in noisy conditions	138
Table 4.7	Detection performance of 3-D distorted objects in noisy conditions after post-processing technique	138

LIST OF FIGURES

Fig 1.1	VanderLugt correlator	8
Fig 1.2	Joint transform correlator	10
Fig 2.1	Classical JTC architecture	17
Fig. 2.2 (a)	Input joint image of identical target object	20
Fig. 2.2 (b)	Classical JTC output of Fig. 2.2 (a)	20
Fig. 2.3 (a)	Input joint image of non-target object	21
Fig. 2.3 (b)	Classical JTC output of Fig. 2.3 (a)	21
Fig. 2.4 (a)	Input joint image with identical targets	24
Fig. 2.4 (b)	Classical JTC output of Fig. 2.4 (a)	24
Fig. 2.5 (a)	Input joint image with one target and one non-target	26
Fig. 2.5 (b)	Classical JTC output of Fig. 2.5 (a)	26
Fig. 2.6 (a)	Input joint image with multiple target and non-target objects	27
Fig. 2.6 (b)	Classical JTC output of Fig. 2.6 (a)	27
Fig. 2.7 (a)	Input joint image with multiple identical targets and non-target objects	30
Fig. 2.7 (b)	Classical JTC output after image subtraction of Fig. 2.7 (a)	30
Fig. 2.8 (a)	Input joint image with multiple identical targets and multiple non-target objects	33
Fig. 2.8 (b)	FJTC output of fig. 2.8 (a) with $C=1$ and $D=1e-1$	34
Fig. 2.8 (c)	FJTC output fig. 2.8 (a) with $C=1$ and $D=1e-4$	34
Fig. 2.8 (d)	FJTC output fig. 2.8 (a) with $C=1$ and $D=1e-9$	35
Fig. 2.8 (e)	FJTC output fig. 2.8 (a) with $C=1$ and $D=1e-12$	35
Fig. 2.9 (a)	Binary joint image	38
Fig. 2.9 (b)	Performance of GFAF with $m=0$	38
Fig. 2.9 (c)	Performance of GFAF with $m=1$	39
Fig. 2.9 (d)	Performance of GFAF with $m=2$	39
Fig. 2.10	Block diagram of PFJTC Technique	41

Fig. 2.11 (a)	Input joint image with multiple identical targets and multiple non-target objects	45
Fig. 2.11 (b)	PFJTC output of fig. 2.11 (a) with $C=1$ and $D=1e-4$	45
Fig. 3.1	Architecture of the proposed PSPFJTC scheme	50
Fig. 3.2 (a)	Binary joint input image	54
Fig. 3.2 (b)	Classical JTC output of Fig. 3.2 (a)	54
Fig. 3.2 (c)	JTC output of fig. 3.2 (a) after image subtraction	55
Fig. 3.2 (d)	FJTC output of fig. 3.2 (a) with image subtraction	55
Fig. 3.2 (e)	PFJTC output of fig. 3.2 (a)	56
Fig. 3.2 (f)	PS PFJTC output of fig. 3.2 (a)	56
Fig. 3.3 (a)	Gray level joint input image	57
Fig. 3.3 (b)	Classical JTC output of fig. 3.3 (a)	57
Fig. 3.3(c)	JTC output after image subtraction of fig. 3.3 (a)	58
Fig. 3.3 (d)	FJTC output with image subtraction of fig. 3.3 (a)	58
Fig. 3.3 (e)	PFJTC output of fig. 3.3 (a)	59
Fig. 3.3 (f)	Phase shifted PFJTC output of fig. 3.3 (a)	59
Fig. 3.4 (a)	Gray level joint input scene of 5 tanks as target, without noise	61
Fig. 3.4 (b)	PFJTC output of fig. 3.4 (a) with $D=10^{-4}$	62
Fig. 3.4 (c)	PSPFJTC output of fig. 3.4 (a) with $D=10^{-4}$	62
Fig. 3.5 (a)	Gray level joint input scene without noise of 2 tanks as target and 3 trucks as non-target	63
Fig. 3.5 (b)	PFJTC output of fig. 3.5 (a) with $D=10^{-4}$	64
Fig. 3.5 (c)	PS PFJTC output of fig. 3.5 (a) with $D=10^{-4}$	64
Fig. 3.6 (a)	Gray level joint input scene with noise of 5 tanks as target	66
Fig. 3.6 (b)	PFJTC output of fig. 3.6 (a) with $D=10^{-4}$	67
Fig. 3.6 (c)	PS PFJTC output of fig. 3.6 (a) with $D=10^{-4}$	67
Fig. 3.6 (d)	PFJTC output of fig. 3.6 (a) with $D=10^{-3}$	68
Fig. 3.6 (e)	PSPFJTC output of fig. 3.6 (a) with $D=10^{-3}$	68
Fig. 3.7 (a)	Gray level joint input scene with noise of 2 tanks as target and	69

	3 trucks as non-target	
Fig. 3.7 (b)	PFJTC output of fig. 3.7 (a) with $D=10^{-4}$	70
Fig. 3.7 (c)	PS PFJTC output with $D=10^{-4}$	70
Fig. 3.7 (d)	PFJTC output of fig. 3.7 (a) with $D=10^{-3}$	71
Fig. 3.7 (e)	PSPFJTC output of fig. 3.7 (a) with $D=10^{-3}$	71
Fig. 3.8 (a)	Binary joint image with two objects in the class and in the input scene	78
Fig. 3.8 (b)	Correlation output of fig. 3.8 (a) with equal values of power spectra parameter	79
Fig. 3.8 (c)	Side view of Fig. 3.8 (b)	79
Fig. 3.8 (d)	Correlation output of Fig. 3.8 (a) with unequal values of power spectra parameter	80
Fig. 3.8 (e)	Side view of Fig. 3.8 (d)	80
Fig. 3.9 (a)	Binary joint image with 2 objects in the class and multiple objects in the input scene	81
Fig. 3.9 (b)	Correlation output of Fig. 3.9 (a)	82
Fig. 3.9 (c)	Side view of Fig. 3.9 (b)	82
Fig. 3.10 (a)	Binary joint image with three objects in the class and in the input scene	84
Fig. 3.10 (b)	Correlation output of Fig. 3.10 (a) with equal values of α , β and γ	85
Fig. 3.10 (c)	Side view of Fig. 3.10 (b)	85
Fig. 3.10 (d)	Correlation output with unequal α , β and γ	86
Fig. 3.10 (e)	Side view of fig. 3.10 (d)	86
Fig. 3.11 (a)	Binary joint image with 3 objects in the class and multiple objects in the input scene	87
Fig. 3.11 (b)	Correlation output of Fig. 3.11(a)	88
Fig. 3.11 (c)	Side view of Fig. 3.11 (b)	88
Fig 4.1 (a)	Original binary image of 'E' character	95
Fig 4.1 (b)	10 projections of fig. 4.1 (a)	96

Fig 4.1 (c)	Reconstructed image of fig. 4.1 (a)	96
Fig 4.1 (d)	60 projections of fig. 4.1 (a)	97
Fig 4.1 (e)	Reconstructed image of fig. 4.1 (a)	97
Fig 4.1 (f)	180 projections of fig. 4.1 (a)	98
Fig 4.1 (g)	Reconstructed image of fig. 4.1 (a)	98
Fig 4.2(a)	2 training images	99
Fig 4.2 (b)	90 projections from each of the training image of fig. 4.2 (a)	100
Fig 4.2 (c)	Composite image using 2 training images in fig. 4.2 (a)	100
Fig 4.2 (d)	60 projections from each of the training image of fig. 4.2 (a)	101
Fig 4.2 (e)	Composite image using 2 training images in fig. 4.2 (a)	101
Fig 4.2 (f)	120 projections from each of the training image of fig. 4.2 (a)	102
Fig 4.2 (g)	Composite image using 2 training images in fig. 4.2 (a)	102
Fig. 4.3	Training images	104
Fig 4.4 (a)	Composite image using equal (=unity) weight factor	105
Fig 4.4 (b)	PSPFJTC output of all training images in fig. 4.3	105
Fig 4.5 (a)	Composite image after 100 iterations	106
Fig 4.5 (b)	PSPFJTC output of all training images in fig. 4.3	107
Fig 4.5 (c)	Side-view of Fig 4.5 (b)	107
Fig 4.6 (a)	Input binary joint image of characters containing in-plane rotated target and non-target images	109
Fig 4.6 (b)	CJTC output of fig. 4.6 (a) after image subtraction	109
Fig 4.6 (c)	POJTC output of fig. 4.6 (a)	110
Fig 4.6 (d)	PSPFJTC output of fig. 4.6 (a)	110
Fig. 4.7 (a)	Composite gray-level car image for azimuth change	113
Fig. 4.7 (b)	Training images	113
Fig 4.7 (c)	PSPFJTC output of all training images in fig. 4.7 (b)	114
Fig. 4.7 (d)	Side view of Fig. 4.7 (c)	114
Fig 4.8 (a)	Joint image containing all target images	116
Fig. 4.8 (b)	POJTC output of Fig. 4.8 (a)	117
Fig 4.8 (c)	PSPFJTC output of Fig. 4.8 (a)	117

Fig 4.9 (a)	Joint image containing two target and two non-target images	118
Fig. 4.9 (b)	POJTC output of Fig. 4.9 (a)	119
Fig. 4.9 (c)	PSPFJTC output of Fig. 4.9 (a)	119
Fig. 4.10 (a)	Joint image containing one target and three non-target images	120
Fig. 4.10 (b)	POJTC output of Fig. 4.10 (a)	121
Fig. 4.10 (c)	PSPFJTC output of Fig. 4.10 (a)	121
Fig. 4.11 (a)	Composite image of gray-level car image for change in Elevation angle	123
Fig. 4.11 (b)	Training images	123
Fig. 4.11 (c)	PSPFJTC output of all training images in fig. 4.11 (b)	124
Fig. 4.11 (d)	Side view of Fig. 4.11 (c)	124
Fig. 4.12 (a)	Joint image containing three target and four non-target images	125
Fig. 4.12 (b)	POJTC output of Fig. 4.12 (a)	126
Fig. 4.12 (c)	PSPFJTC output of Fig. 4.12 (a)	126
Fig. 4.13 (a)	Joint image containing one target and six non-target images	127
Fig. 4.13 (b)	POJTC output of Fig. 4.13 (a)	128
Fig. 4.13 (c)	PSPFJTC output of Fig. 4.13 (a)	128
Fig. 4.14 (a)	Gray-level joint image containing out-of-plane rotated target images only	129
Fig. 4.14 (b)	PSPFJTC output of Fig. 4.14 (a)	130
Fig. 4.14 (c)	Side view of Fig. 4.14 (b)	130
Fig. 4.15 (a)	Gray-level joint image containing out-of-plane rotated target and non-target images	132
Fig 4.15 (b)	PSPFJTC output of Fig. 4.15 (a)	133
Fig. 4.15 (c)	Side view of Fig. 4.15 (b)	133
Fig. 4.16 (a)	Joint image with an additive white noise of 5 dB	134
Fig. 4.16 (b)	PSPFJTC output of Fig. 4.16 (a)	135
Fig. 4.16 (c)	Side view of Fig. 4.16 (b)	135
Fig. 4.17 (a)	Joint image with an additive white noise of 0 dB	136

Fig. 4.17 (b)	PSPFJTC output of Fig. 4.17 (a)	137
Fig. 4.17 (c)	Side view of Fig. 4.17 (b)	137
Fig. 4.18 (a)	Correlation output of fig. 4.16 (c) after post-processing	139
Fig. 4.18 (b)	Correlation output of fig. 4.17 (c) after post-processing	139

LIST OF ABBREVIATIONS

CCD	Charge coupled device
CFAF	Class-associative fringe-adjusted filter
CH	Circular harmonic filter
FAF	Fringe-adjusted filter
FJTC	Fringe-adjusted joint transform correlator
FPFJTC	Fractional power fringe-adjusted joint transform correlator
JPS	Joint power spectrum
JTC	Joint transform correlator
LCLV	Liquid crystal light valve
MACE	Minimum average correlation energy filter
MTDA	Multi-target detection algorithm
MVSDF	Minimum variance synthetic discriminant function
POF	Phase only filter
POJTC	Phase only joint transform correlation
PSDF	Projection-slice synthetic discriminant function
PSPFJTC	Phase-shifted phase-encoded fringe-adjusted joint transform correlator
SDF	Synthetic discriminant function
SLM	Spatial light modulator
VLC	VanderLugt Correlator

LIST OF SYMBOLS

α	Power spectra parameter of reference 1 in class detection
β	Power spectra parameter of reference 2 in class detection
γ	Power spectra parameter of reference 3 in class detection
λ	Wavelength of collimated light in JTC scheme
ξ	Normalized acceptable error limit in PSDF generation
\mathfrak{F}	Fourier transform operation
\mathfrak{F}^{-1}	Inverse Fourier transform operation
\otimes	Correlation operation
$*$	Complex conjugate operation
f	Focal length of lenses
\times	Multiplication operation
x, y, x', y'	Space domain variable
u, v	Fourier domain variable
η	Relaxation factor in PSDF generation
m, C, D	GFAF filter parameters
H_{faf}	Fringe-adjusted filter function in JTC
H_{gfaf}	Generalized fringe-adjusted filter function in JTC
H_{gfaf}^c	Classical filter function of GFAF in JTC
H_{gfaf}^f	Fringe-adjusted filter function of GFAF in JTC
H_{gfaf}^p	Phase-only filter function of GFAF in JTC
H_{gfaf}^{sdf}	GFAF filter function with SDF or PSDF as reference in JTC
r_{sdf}	SDF reference image

ACKNOWLEDGMENT

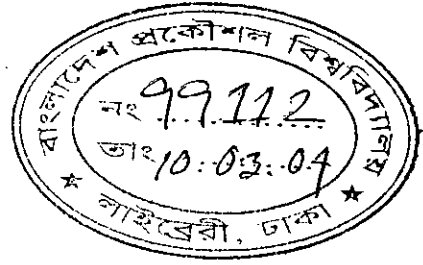
The author would like to express his deep gratitude to his supervisor Dr. M. Nazrul Islam, Associate Professor, Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology (BUET), Dhaka-1000, for his guidance in this research work. His ideas, insights, suggestions, questions and enthusiasm were great help in stimulating the thought and in bringing this research forward. The whole-hearted co-operation of Dr. Nazrul Islam made this work completed timely and successfully.

The author is grateful to Bangladesh University of Engineering and Technology, Dhaka-1000, for providing various resources required for this work. The author would also like to thank all others who extended their hands directly and indirectly to complete this work. The author would also like to express profound gratitude to his family members and relatives for their unwavering support and encouragement during the graduate studies.

ABSTRACT

This thesis is concerned about the detection or pattern recognition of three dimensionally distorted multiple target objects with the application of projection-slice synthetic discriminant function and optical correlation technique. Though distortion invariant recognition has been made for single and multiple target of same object, still no work has been reported on 3-D distortion invariant optical pattern recognition with better utilization of space bandwidth product. The previous method like phase-encoded fringe-adjusted joint transform correlation technique is very much sensitive to the number of objects and presence of noise in the input scene. In this thesis work, a modified multi-target detection algorithm employing phase-shifting and phase-encoding principle, has been proposed to get more prominent detection signal such as better utilization of space bandwidth product of the input and output plane with same filter parameters, even in the midst of very much similar non-target and noise. In this work, the out-of-plane rotation of an object is considered as the 3-D distortion of the object. The projection-slice theorem has been taken to create synthetic discriminant function-based matched filters that are capable of discerning out-of-plane rotation of objects. According to projection-slice theorem, the 1-D Fourier transformation of the projection of an image is equivalent to a 1-D Fourier slice of that image taken on the line specified. The idea introduced in this thesis is to integrate 1-D Fourier slices from different images to form a 2-D image for the recognition of the object after it has undergone a distortion. To satisfy the equal peak criterion, the composite image becomes a weighted combination of all the slices. In the previous works like synthetic discriminant function-based techniques, the entire two-dimensional image is used in the design of filter response that does not match any of the training images. The incorporation of projection-slice theorem with synthetic discriminant function eliminates this problem, where at least certain number of slices must match with each of the training image. As a result, projection-slice synthetic discriminant function technique is well suited for detection of 3-D distorted objects. In projection-slice synthetic discriminant function

technique, the reduction of slice number with the increase of distortion range introduces severe loss of information in the composite image. Therefore, for the detection of 3-D distorted objects using projection-slice synthetic discriminant function, a modified class-associative target detection algorithm has been incorporated. Class-associative target detection employs a multi-target detection algorithm (MTDA) facilitating multiple reference images for which more than one dissimilar objects or images of a class can be detected. For detection of higher range of 3-D distorted images, different composite images are formed relating to different 3-D distortion and are used as the member of a class. An efficient and fast method for class-associative target detection and an enhanced version of Generalized fringe-adjusted filter have been suggested. Computer simulation shows satisfactory performance of the proposed scheme in getting distortion invariant detection of members of the class mainly with out-of-plane rotation. The proposed method also ensures almost equal target peaks while negligible peaks for non-target objects.



Chapter 1

INTRODUCTION

1.1 Introduction

Pattern recognition is an important task for both data processing and decision-making. Pattern recognition technique is mainly concerned with the automatic detection and classification of objects and events. It deals with the detection and identification of a desired pattern or target in an unknown input scene and determination of the spatial location of the targets. There is no doubt that pattern recognition is an important, useful and rapidly developing technology with cross-disciplinary interest and participation. Pattern recognition techniques overlap with other areas such as (Adaptive) signal processing and systems, artificial intelligence, Neural modeling, Optimization/estimation theory, Fuzzy sets etc. Some of the major applications of pattern recognition techniques include target detection, computer vision, radar signal classification, image and speech preprocessing, face recognition, biometric (fingerprint) identification etc.

Pattern recognition, naturally, is based on patterns. A pattern can be as basic as a set of measurements or observations, perhaps represented in vector or matrix notation. The use of measurements already presupposes some preprocessing and instrumentation system complexity. These measurements may be a two-dimensional image, a drawing, a waveform, a set of measurements, a temporal or spatial history (sequence) of events, the state of a system, the arrangement of a set of objects and so forth. On the other hand, the characteristic features of a pattern are any extractable measurement used. Examples of low level features are signal intensities. Features may also result from applying a feature extraction algorithm or operator to the input data. The key is to choose and extract features that are computationally feasible, lead to good pattern recognition system success and reduce the problem data without discarding valuable information. The

number of characteristic features is virtually always chosen to be fewer than the total necessary to describe the complete target and thus there is a loss of information. It is different from image processing where the input is an image and the output is also an image. Since pattern recognition is fundamentally an information reduction process, it is therefore not possible to reconstruct the original pattern but to make a precise decision [1-2].

The methodology how a human being and a machine classify objects is quite different. The main difference between human and machine intelligence comes from the fact that humans perceive everything as a pattern, whereas for a machine everything is data. If there is no pattern, then it is very difficult for a human being to remember and reproduce the data later. Thus the storage and recall operations in human beings and machines are performed by different mechanism. Again the major characteristic of a human being is the ability to learn continuously from examples, which is not understood well enough to implement in algorithmic fashion in machine [3].

1.2 Pattern Recognition Tasks

The inherent differences of information handling by human being and machine, have led to identify different pattern recognition tasks which human beings are able to perform very naturally and effortlessly. On the contrary, no simple algorithm has yet been reported to implement these tasks in a machine. A brief description of several pattern recognition tasks is given below.

1.2.1 Pattern Association

Pattern association problem involves storing a set of patterns or a set of input-output pattern pairs in such a way that when a test pattern is presented, the stored pattern or the pattern pair corresponding to the test pattern is recalled. It is desirable to recall the

correct pattern even though the test pattern is noisy or distorted. The problem of storage and recall of patterns is called auto-association task. On the other hand, the problem of storage and recall of pattern pairs is called hetero-association task [3].

1.2.2 Pattern Classification

The objective of pattern classification is to find out the implicit relation among the patterns of the same class, so that when a test pattern is given, the corresponding output class level is retrieved. Here the individual patterns of each class are not memorized. In this case, the test patterns belonging to a class are not the same as the patterns used in training. Speech spectra of steady vowels generated by a person or hand-printed characters could be considered as examples of patterns for pattern classification problems.

1.2.3 Pattern Mapping

In a pattern mapping, given a set of input pattern and the corresponding output patterns, the objective is to capture the implicit relationship between the input and output patterns, so that when a test input pattern is given, the pattern corresponding to the output of the generating system is retrieved. In this case system should perform some kind of generalization as opposed to memorizing the information.

1.2.4 Pattern Grouping

In this case, given a set of patterns, the problem is to identify the subset of patterns possessing similar distinctive features, and group them together. Since the number of groups and features of each group are not explicitly stated, this problem belongs to the category of unsupervised learning. In the pattern classification problem the patterns of

each group are given separately. In pattern grouping, patterns belonging to several groups are given and the system has to resolve them into different groups.

1.2.5 Feature Mapping

In several patterns the features are not unambiguous. In fact, the features vary over a continuum and hence it is difficult to form groups of patterns having some distinctive features. In this case, it is desirable to display the feature variations in the patterns. In this case the main purpose is to generate the feature map of a pattern and not to identify the group or class to which the pattern belongs.

1.2.6 Pattern Variability

There are many situations when the features in a pattern undergo unspecified distortions each time the pattern is generated by the system. Human beings are able to recognize them due to some implicit interrelations among the features, which are not known precisely. Classification of such patterns falls into the category of pattern variability task.

1.2.7 Temporal Patterns

All the tasks discussed so far, refer to the features present in a given static pattern. It is sometimes required to capture the dynamic features present in a sequence of patterns as in a movie on a television. All such situations require handling multiple static patterns simultaneously, looking for changes in the features in the sub-patterns in adjacent pattern pairs.

1.2.8 Stability-plasticity Dilemma

In any pattern recognition task, if the input patterns keep changing, then it is difficult to freeze the categorization task based on a set of patterns used in the training set. If it is frozen, then the system cannot learn the category that a new pattern may suggest. In this case, the system is said to lose its plasticity. If the system is allowed to change its categorization continuously, based on new input patterns, it cannot be used for any application such as pattern classification or grouping, as it is not stable. This is called the stability-plasticity dilemma in pattern recognition.

1.3 Methods of Pattern Recognition Tasks

Generally speaking, the underlying and quantifiable statistical basis for generation of patterns and the underlying structure of the pattern, provide the fundamental information for pattern recognition tasks. Based on these, pattern recognition tasks can be classified into three groups that are described as follows.

1.3.1 Statistical Pattern Recognition

This is a statistical basis for classification of algorithms. A set of characteristic measurements is extracted from the input data and is used to assign each feature vector to one of the classes. Features are assumed generated by a state of nature, and therefore the underlying model is a state of nature or class-conditioned set of probabilities or probability function.

Various methods for finding decision functions in decision theoretic approach are matching, optimum statistical classifier and neural networks. There are two ways of matching- one is minimum distance classifier and the other is matching by correlation [4].

1.3.2 Syntactic Pattern Recognition

Many times the significant information in a pattern is not merely in the presence or absence of features. Rather, the interrelationships or interconnections of the features yield important structural information, which facilitates structural description or classification. This is the basis of syntactic pattern recognition. However, in syntactic pattern recognition approach, one must be able to quantify and extract structural information and to assess structural similarity of patterns.

Typically, syntactic pattern recognition approaches formulate hierarchical descriptions of complex patterns built up from simpler sub-patterns. At the lowest level, primitives, elements or characteristics of syntactic pattern recognition, involves the choice of primitives. Primitives must be sub-pattern or building blocks whereas features are any measurements.

1.3.3 Neural Pattern Recognition

Modern digital computers do not emulate the computational paradigm of biological systems. The alternative of neural computing emerged from attempts draw on knowledge of how biological neural systems store and manipulate information. This leads to a class of artificial neural systems termed neural networks. Neural networks are particularly suited for pattern association applications.

1.4 Implementation Techniques of Pattern Recognition

Pattern recognition techniques can be implemented either by a computer or by an optical system. Signal processing through computer for traditional pattern recognition has the temporal advantage, whereas optical system has an additional manipulation advantage. Digital pattern recognition techniques mainly suffer from the intensive

computation and implementation difficulties. On the other hand, optical pattern recognition techniques inherently provide parallelism, ultrahigh processing speed, non-interfering communication and massive interconnection capability. Therefore, optical pattern recognition technique offers itself as a suitable candidate for real time pattern recognition [5].

1.5 Optical Pattern Recognition

Optical pattern recognition is usually carried out on the basis of correlation, which can be performed easily, and efficiently with the help of a lens. Correlation is a sensor independent approach that generally does not require data specific operation. In correlation based pattern recognition approach, complexity of the system depends on the input image size but not on the number of objects in the scene. This avoids the computational bottlenecks faced by some other techniques due to scene complexity, clutter and no of objects present.

Optical correlators usually use the Fourier transform properties of lens. If an input image is placed at the front focal plane of a lens and illuminated by a laser source, then the Fourier transform of the input image will be got at the back focal plane of the lens [6-7]. This transforming of input transparency can be performed in real time and therefore making the optical correlator as a suitable candidate for real time target detection process.

To perform the correlation between reference and target image, a joint image containing reference and target images placed side-by-side is placed at front focal plane of a lens. At the back focal plane, the Fourier transform of reference and target images are multiplied in the frequency domain, which is equivalent to convolution in the spatial domain. A second lens is used to transform the frequency function back to spatial domain. Since an optical correlator provides parallel processing capability, it is a

suitable candidate to provide better speed advantages compared to digital counter parts. Other than lens, the optical devices that are used for optical pattern recognition technique are Spatial light modulator (SLM), charge coupled device (CCD), mirror, beam splitter, laser etc.

There are two widely used optical correlators namely VanderLugt Correlator (VLC) and Joint Transform correlator (JTC).

1.5.1 VanderLugt Correlator (VLC)

VanderLugt Correlator is also known as matched filter based correlator. It was first introduced in 1964 and since then it has spurred myriad of applications [8-13]. It requires a priori fabrication of a complex matched filter used in the Fourier and perfect alignment of this filter along the optical axis. Thus, it is not suitable for real time operation and slight mismatch results in output intensity degradation for optical correlation operation [14-16]. A schematic diagram of a VLC is shown in the figure 1.1.

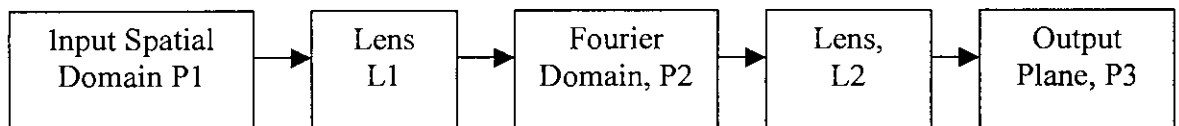


Fig. 1.1: VanderLugt Correlator

In Fig. 1.1, P1, P2 and P3 represent the input plane, Fourier plane and output plane respectively. If an input signal $f(x, y)$ is placed at the input plane P1 and illuminated by a point laser source, the complex light field distribution produced at the plane P2 is given by

$$F(u, v) = \mathfrak{F}[f(x, y)] \quad (1.1)$$

where \mathfrak{F} represents Fourier transform operation and u and v are frequency dependent variables scaled by a factor of $\frac{2\pi}{\lambda f}$, where λ is the wavelength of collimating light and f is the focal length of lens.

The phase fronts of the wave emanating from plane P1 and incident on P2 are both curved and distorted. The phase transmittance of the matched filter $H(u, v)$ at the plane P2 must be the conjugate match to the phase of the Fourier transform of $f(x, y)$. As a result, the phase variations of the beam incident on plane P2 is cancelled and the resulting distribution becomes a parallel beam. Lens L2 performs the inverse Fourier transform and produces the output at plane P3. If the input is different from $f(x, y)$, the phase-front will not exactly be cancelled by the filter and the light distribution in plane P3 will produce a spot of poor intensity or no spot at all. The complex light distribution at plane P3 can be expressed as

$$g(x, y) = \mathfrak{F}^{-1}[F(u, v)H(u, v)] \quad (1.2)$$

where $\mathfrak{F}^{-1}(\dots)$ represents inverse Fourier transformation.

The intensity of bright correlation spot is proportional to the degree to which the input and the filter functions are matched. Since the system is linear, superposition theorem holds. The aforementioned arguments are also valid for multiple objects present at different at different locations of the input plane. This correlation system provides a great deal of sensitivity since it is both phase matched and amplitude matched [5].

1.5.2 Joint Transform Correlator (JTC)

Joint transform correlator is a device consisting of two optical systems in which two signals are simultaneously transformed to produce their spectra, and these spectra are multiplied and inverse Fourier transformed to produce the correlation output. Fig 1.2 shows the schematic diagram of a classical JTC architecture.

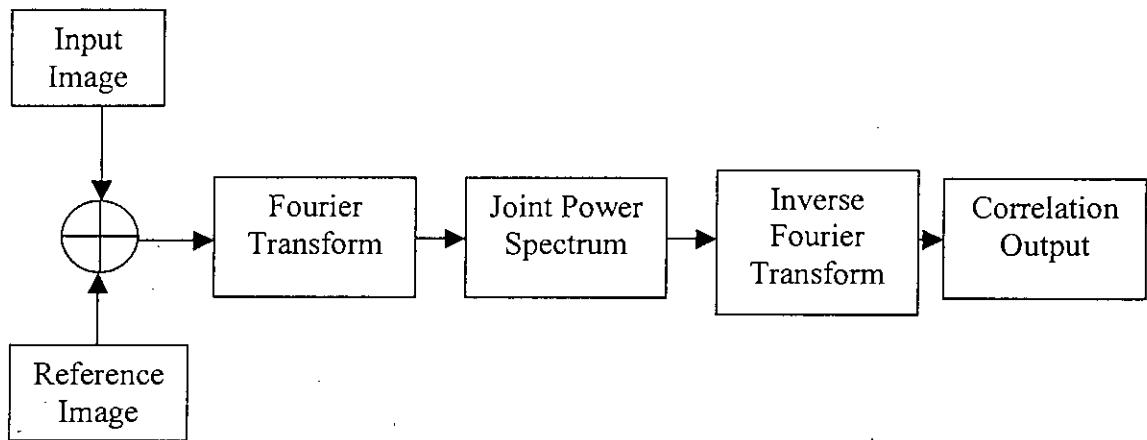


Figure 1.2: Joint Transform Correlator

Optical pattern recognition by using joint transform correlation (JTC) was first introduced by Goodman, in 1966 [17]. With the development of electronically addressable spatial light modulator (SLM), JTC has been introduced as a real-time programmable optical pattern recognition technique [18]. JTC does not require any complex matched filter fabrication and accurate alignment of the filter along the optical axis. It uses a spatial (impulse response) domain filter. In contrast to VLC, joint transform power spectrum in JTC is dependent on input signal. Some of the major advantages of JTC are that it allows real-time update of the reference image and permits parallel Fourier transform of the reference image and the unknown input scene. Although a classical JTC provides many attractive advantages, it is found to suffer from large correlation side-lobes, large correlation width, wide zero order peak and low optical efficiency [18-19].

To improve the performance of classical JTC, binary JTC has been proposed. A binary JTC is found to be superior to a classical JTC in terms of the correlation peak intensity, correlation width and discrimination sensitivity [20-22]. However, a binary JTC involves computationally intensive Fourier plane joint power spectrum (JPS)

binarization, which limits the system processing speed [23]. For multi-object input scene, in particular, the JPS binarization process introduces harmonic correlation peaks that may cause spurious correlation peaks and thus, results misses [24-25] and thereby complicating the target detection process. However, the binary JTC cannot completely eliminate the strong zero-order term at the output plane.

Recently, Tang et al. proposed a chirp-encoded JTC, which requires multiple spatial light modulators for the input planes (to display the reference and input scene) as well as multiple detectors for the output planes [26]. Hence, a large space-bandwidth product is required for chirp-encoded JTC for multiple target detection process.

To overcome or reduce the zero diffraction and enhance the discrimination ability of a classical JTC technique, several methods have been introduced. Different zero-order elimination techniques have been proposed in the literature [27-32]. DC blocking, using an opaque aperture to block the dc component can be easily carried out. But, when the input scenes are noised, zero order items become very complicated which results in the futility of the method. Fourier plane image subtraction [28], correlation plane image subtraction [29] and phase-shift power spectrum subtraction [30] all can yield a better correlation output. All of the methods need multiple processing steps to achieve it, which limits their applications to the highly required real-time recognition tasks.

However, the above-mentioned JTCs still produce relatively broad correlation peaks which complicates their application to multiple target detection process. To get a sharper correlation peaks, one of the recent apodization based approaches is the use of a fringe adjusted JTC (FJTC) based on Newton-Raphson algorithm [33]. Fringe adjusted filter (FAF)-based JTC avoids many problems otherwise associated with other JTC techniques. In FJTC technique, the JPS is multiplied by the transfer function of FAF before inverse Fourier transformation. This technique yields better correlation performance than alternate JTCs for the noise free single and multi-target binary input scenes under normal as well as poor illumination conditions [34]. However, for noise

corrupted input scenes, whenever the reference power spectrum contains very low values, the FJTC technique may suffer from low correlation output.

To overcome this problem, a fractional power fringe adjusted JTC has been proposed for single target detection. This technique employs a family of real valued filters, called generalized fringe adjusted filters (GFAF) [35]. By adjusting a parameter, one can obtain classical JTC, fringe adjusted JTC and phase-only JTC without actually fabricating these complex valued filters. It is found that the phase only JTC provide better noise robustness than the fringe adjusted JTC technique.

Again all of the above methods utilize at most 50% of the input and output plane SLM and therefore provides poor utilization of space bandwidth product. Again, multi-step processing techniques to eliminate the extraneous signals, limit their operation in such cases where processing speed is a crucial constant. To overcome the above-mentioned problem, a phase encoding principle has been adopted with FAF [36] to get a sharper and only a single correlation peak per target. This phase encoding operation is performed in such a way that it does not have any detrimental effect on the system processing speed. This technique uses separate reference and input planes and yields one correlation peak per target instead of a pair of peaks thus ensuring better utilization of space bandwidth product. This is significantly an improved target detection process for real time operations.

A number of various other non-linear joint transform correlation have been proposed for optical pattern recognition among which all-optical photo refractive crystal based JTC have been found to be particularly attractive [37-40]. There are also some other phase-encoding techniques that yield a single correlation peak per target and provide better utilization of space-bandwidth product [41-43].

1.6 Distortion Invariant Pattern Recognition

Target detection process may involve detection of a desired target object from a noisy or uncorrelated data. A target in input scene may be distorted or corrupted by a variety of ways. It may in-plane and out-of-plane rotation, scale variation or shifting of the position of the object. Clutter, noise, obscuration and low illumination are some of the other contribution factors of distortion. There are a number of different approaches for distortion invariant pattern recognition tasks. Among these, methods using transform coefficient features, algebraic features, visual features, moment-based methods and synthetic discriminant function based methods are well known.

Moment invariant methods are found to be effective in obtaining full geometrical invariance, such as rotation and translation [44-45]. Moment invariance method is often utilized for invariant character recognition [46-47]. The disadvantage of moment invariant method is the relatively time consuming computation and its inaminability of optical implementation. It also falls short of expectation in case of noise degradation. Circular harmonic (CH) method has been well utilized to deal with rotation invariant recognition [48-49]. Mellin radial harmonic method is employed for scale invariant recognition [50]. Fourier Mellin descriptors and Fang's self-transform employ the pattern itself in the kernel of a transform that is invariant to objects scale, rotation and translation variations [51-52]. Log-polar mapping is another popular method by which object scale and rotation can be transformed into translation [53].

Basic Synthetic discriminant function (SDF)-based technique uses a training set of distorted images, which are assumed to be representative of possible distortion [54]. The minimum variance synthetic discriminant filter (MVSDF) and minimum average correlation energy filter (MACE) are examples of two such composite filter fabrications [55-56]. Several researchers have used synthetic discriminant function for distortion invariant pattern recognition [57-58]. Javidi successfully applied the SDF formulation in a bipolar nonlinear JTC whereby SDF was used as the reference [59]. Recently, a

rotation-invariant FJTC and a distortion invariant fractional power FJTC technique have been proposed [60-61]. For discerning in-plane and out-of-plane 3D distortions, a modification of basic SDF has been proposed namely projection slice synthetic discriminant function (PSDF) [62]. Wavelet transformation and neural networks have also been efficiently utilized for distortion invariant pattern recognition [63-66].

1.7 Class Associative Pattern Recognition

Sometimes a class of similar or dissimilar objects needs to be detected from a group of different objects. For classification of objects from a given class, recently proposed filter-based techniques utilized phase information from the Fourier transform of each of the objects in that class. In this case consequent addition of objects requires a change in the phase of the filter. Casasent and Tefler proposed a pattern recognition technique that can perform hetero-association based pattern recognition [67]. Later Khoury *et. al.* proposed a new class-associative filter using cross-correlation enhancement. This technique has the advantage of adding additional objects by changing only the amplitude of the filter, i.e., it is phase restricted. In this technique the phase may result from an object of one class or even from an object that may be completely different from the classification set of objects. More, recently a class-associative target detection technique has been proposed by using a new multi-target detection algorithm [69]. Here an enhanced version of the fringe-adjusted filter called class associative fringe adjusted filter (CFAF) has been developed for class associative multiple target detection.

1.8 Objectives of the Thesis

Though distortion invariant recognition has been made for single or multiple target of same object, still no work has been reported on 3-D distortion invariant optical target detection with better utilization of space bandwidth product. Most of the methods of making distortion invariance are either too complex or not successful enough. Again

most of the existing multi-target detection algorithm generates a pair of correlation peak for each target, which may make the detection ambiguous in case of randomly appearing objects in input scene.

The main objective of this work is to develop a method for getting distortion-invariant optical pattern recognition for both binary and gray level images. The projection-slice theorem has been taken to create synthetic-discriminant-function based matched filters that are capable of discerning rotation and scale distortions. A modified multi-target recognition algorithm is aimed to get more prominent detection signal such as better utilization of space band width product of the input and output plane, even in the midst of very much similar non-targets and noise. By combining the above-mentioned PSDF and multi-target recognition algorithm a high performance optical correlator will be proposed that will give almost equal correlation peak for all the targets for both in-plane and out-of-plane distortions and whether there is noise in the input scene or not.

1.9 Thesis Outline

This thesis consists of five chapters. In the first chapter some introduction on optical pattern recognition has been given, especially on joint transform correlation. Also a brief description of the theories and problems associated with this work has been given. Chapter 2 deals with joint transform correlation in details. Classical JTC, fringe-adjusted JTC, fractional power fringe-adjusted JTC and phase-encoded fringe-adjusted JTC are analyzed there with theory and simulation. Chapter 3 deals with the theory and simulation of efficient multi-target detection algorithm that is a basis for both class-associative target detection and 3D distortion invariant target detection. Chapter 4 gives the theory and simulation of PSDF based composite image generation and its application with the class-associative target detection algorithm for detection of 3D distorted objects. Chapter 5 is the final chapter giving a conclusive remark on this work along with some suggestive future works in this area.

Chapter 2

JOINT TRANSFORM CORRELATION TECHNIQUE

2.1 Introduction

Joint transform correlation technique is one of the most widely used correlation technique for optical pattern recognition. In a JTC technique, the reference image and the unknown input image are placed side-by-side by using a spatial light modulator (SLM) and are illuminated by a coherent light source. A lens is used to produce the Fourier transformation of the input joint image. The Fourier transform patterns constructively interfere with each other to create an interference pattern called joint power spectrum (JPS). This joint power spectrum is recorded by a CCD camera and is further inverse Fourier transformed by using another lens to get the correlation output. For a match between the reference and input scene, a pair of correlation peaks or bright spots is produced and for a mismatch, negligible or no correlation peaks are produced. A classical JTC suffers from many disadvantages and therefore several modifications have been proposed to overcome these limitations. In this chapter the classical JTC and fringe-adjusted JTC are discussed along with effect of image subtraction technique.

2.2 Classical Joint Transform Correlation Technique

A classical JTC architecture is shown in Fig. 2.1, where the reference image and the input images are displayed side-by-side using an SLM illuminated by a coherent light source. The combined light distribution passes through the first Fourier lens L1 and the complex light distribution called JPS is recorded by a CCD array. In a classical JTC, the JPS is then introduced into the second lens to get the correlation output.

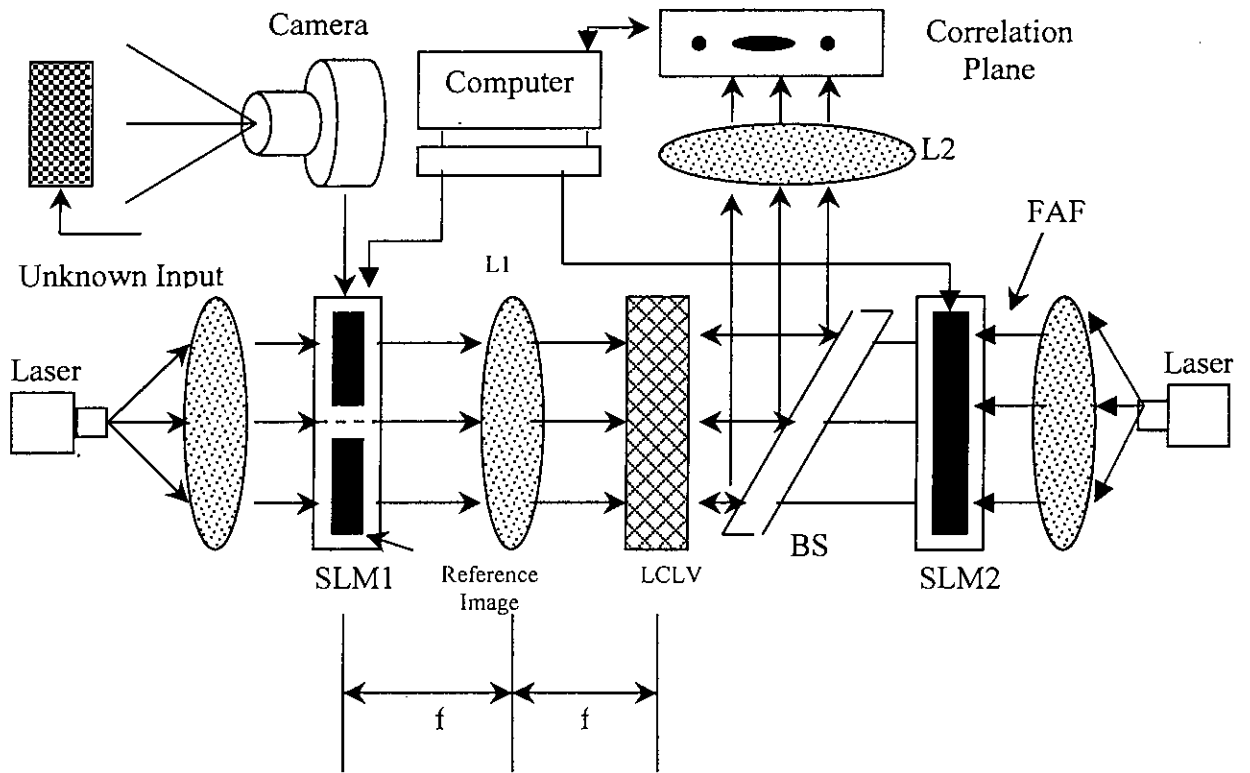


Fig 2.1: Classical JTC architecture

In case of fringe-adjusted JTC (FJTC), the JPS is multiplied by the real valued FAF before applying the inverse Fourier transformation to get sharp delta-like correlation peaks. In this section the theory and simulation of classical JTC will be discussed first for single target detection and then for multiple target detection. Then we will discuss the FJTC and the phase-encoded FJTC (PFJTC) for both single and multiple target detection.

2.2.1 Single Target Detection

2.2.1.1 Analysis

Let a reference image $r(x, y)$ and a target image $t(x, y)$ are displayed side-by-side in the input plane by using an SLM. With $r(x, y + y_0)$ representing the reference image and $t(x, y - y_0)$ representing the input scene in the input plane separated by a distance $2y_0$ along the y axis, the input joint image $f(x, y)$ can be expressed as

$$f(x, y) = r(x, y + y_0) + t(x, y - y_0) \quad (2.1)$$

Lens L1 performs the Fourier transform of $f(x, y)$ to yield

$$F(u, v) = |R(u, v)| \exp[j\Phi_r(u, v) + jvy_0] + |T(u, v)| \exp[j\Phi_t(u, v) - jvy_0] \quad (2.2)$$

where $R(u, v)$ and $T(u, v)$ are the amplitudes, $\Phi_r(u, v)$ and $\Phi_t(u, v)$ are the phases of the Fourier transforms of $r(x, y)$ and $t(x, y)$, respectively; u and v are mutually

independent frequency domain variables scaled by a factor of $\frac{2\pi}{\lambda}$, λ is the wavelength

of collimating light, f is the focal length of lenses L1 and L2. The complex light distribution produced in the back focal plane of L1, called the JPS, is then detected by a square-law device such as a CCD array. The JPS is given by

$$\begin{aligned} |F(u, v)|^2 &= F(u, v) \cdot F^*(u, v) \\ &= |R(u, v)|^2 + |T(u, v)|^2 + 2|R(u, v)||T(u, v)| \times \cos[\Phi_r(u, v) - \Phi_t(u, v) + 2vy_0] \end{aligned} \quad (2.3)$$

In a classical JTC, this JPS is introduced into second SLM which is illuminated by another laser of same frequency. Lens L2 performs the inverse Fourier transformation of the JPS to yield the correlation output. The final output can be expressed as

$$\begin{aligned} o(x', y') = & r(x', y') \otimes r^*(x', y') + t(x', y') \otimes t^*(x', y') \\ & + r(-x', -y' - 2y_0) \otimes t^*(-x', -y') + r^*(x', y' - 2y_0) \otimes t(x', y') \end{aligned} \quad (2.4)$$

From Eq. (2.4), it is evident that the correlation output contains autocorrelation of the reference image and the input scene objects and cross-correlation between the reference image and the input scene objects. The first and second terms of Eq. (2.4) produce the zero-order peak and only the third term generates the desired cross correlation peaks.

2.2.1.2 Simulation Results

To analyze the performance of a classical JTC, a binary character 'E' of English Alphabet has been taken as the reference image. The size of the character is 32×32 pixels and it is placed in a joint image of size 256×256 pixels. The joint image also contains a target image as shown in Fig. 2.2 and 2.3. The simulations are performed using FFT2 routine of MATLAB software and the outputs are plotted using the 3-D plotting routine.

In Fig. 2.2 (a), the target image is the same (i.e. 'E' in this case) as the reference image and the corresponding classical JTC output is shown in Fig 2.2 (b). From the figure, it is evident that there is the presence of a strong zero-order term at the middle of the output plane and a pair of correlation peaks is produced for the target object.

The same analysis has been made with a non-target image 'M' for the same reference image 'E' as shown in Fig. 2.3. In this case a pair of correlation peaks is obtained with negligible intensity.

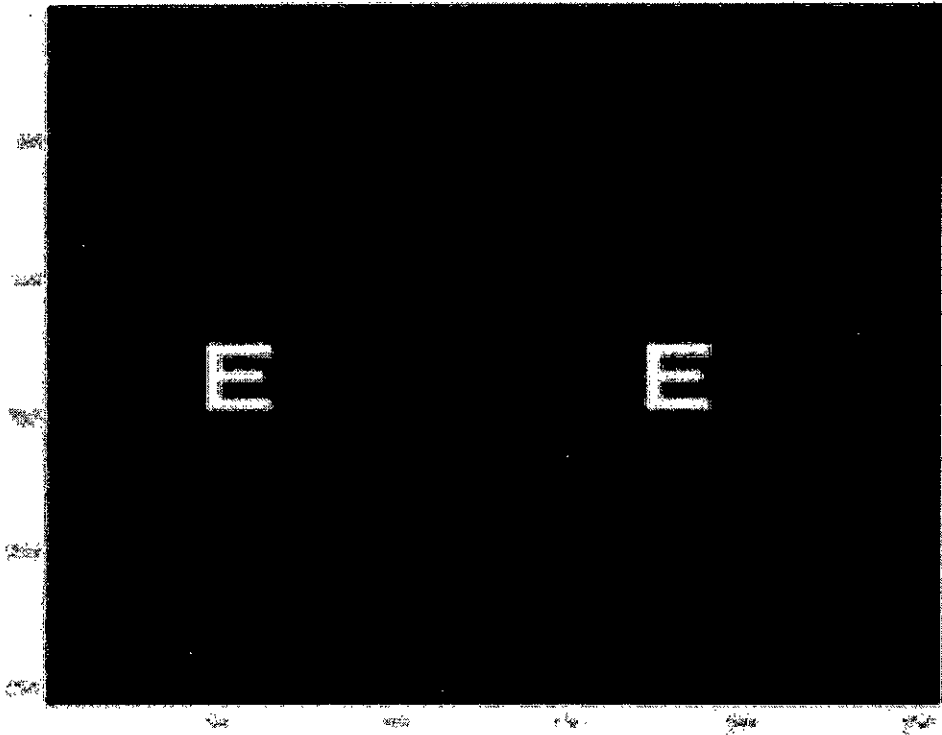


Fig. 2.2 (a): Input joint image of identical target object

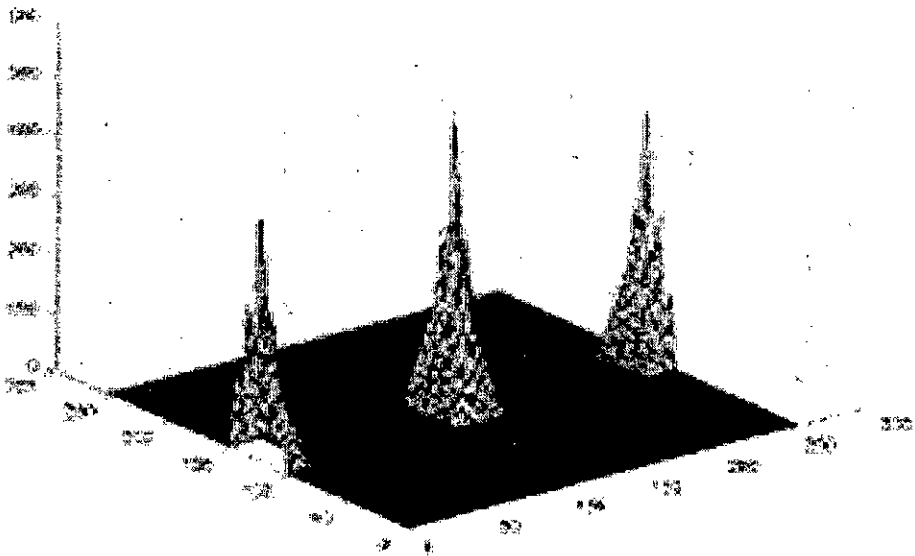


Fig. 2.2 (b): Classical JTC output of Fig. 2. 2 (a)

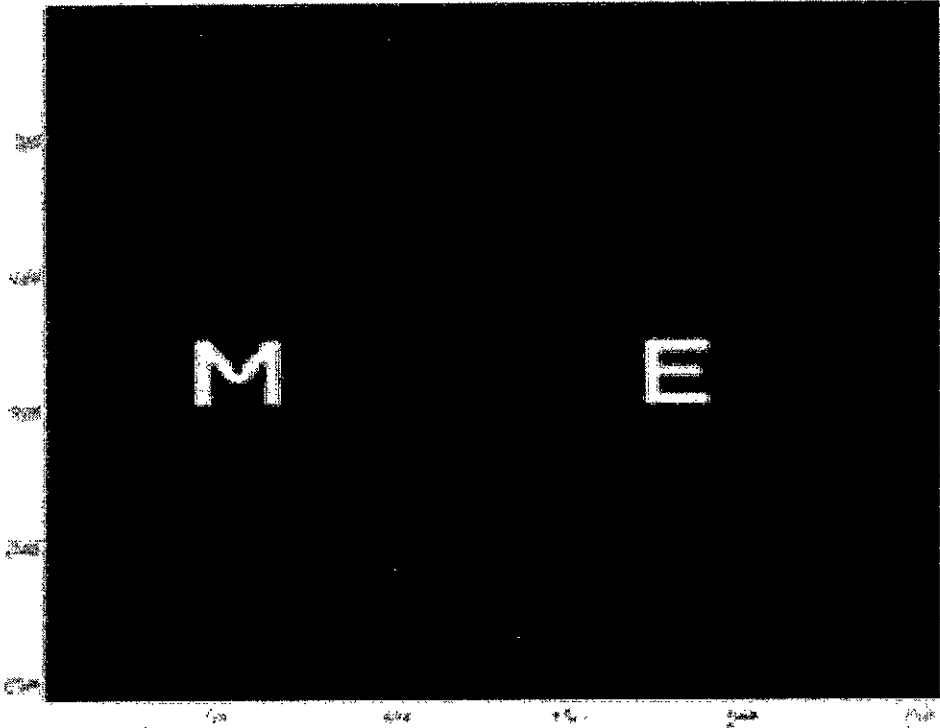


Fig. 2.3 (a): Input joint image of non-target object

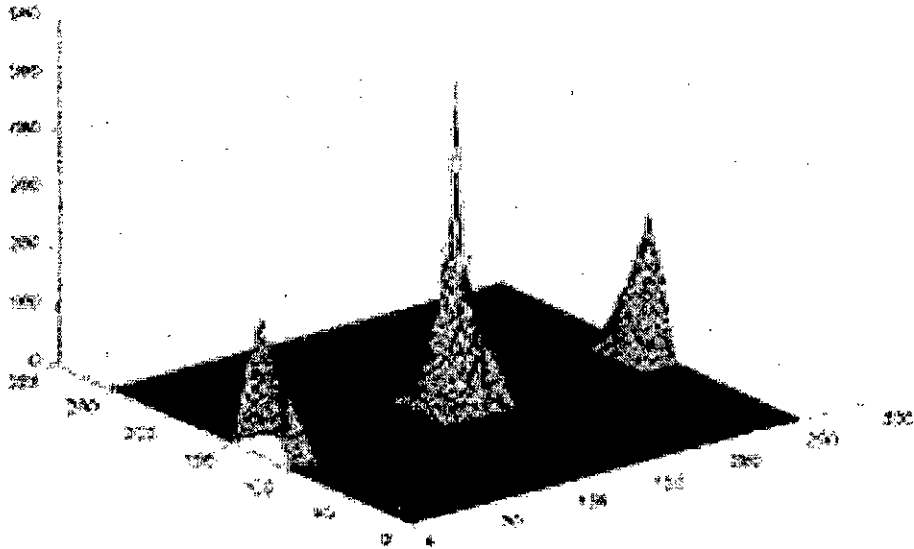


Fig. 2.3 (b): Classical JTC output of Fig. 2.3 (a)

From Fig. 2.2 and Fig. 2.3, it is evident that a DC term or zero-order diffraction is obtained at the center of the output plane with a very high correlation peak intensity compared to the target peaks which deteriorates the detection process and lowers the optical efficiency. The position of the zero-order term is always fixed i.e. at the center of the correlation plane while the positions of cross-correlation terms mainly depend on the relative position of reference image in the joint image. If the reference image is placed at the center of the joint image, then one of the correlation peaks will be generated at the exact position of the target image. But the major disadvantage of this method is that in this case the reference image may be superimposed to any input scene image and therefore performance of the system may be deteriorated. The existence of any cross-correlation peak provides the information about the presence of any object in the input scene while the height of the cross-correlation peaks indicates the degree of similarity between the reference image and the target images present in the input scene.

2.2.2 Multiple Target Detection

2.2.2.1 Analysis

If $r(x, y + y_0)$ represents the reference image and $t(x, y - y_0)$ represents the input scene containing n objects $t_1(x - x_1, y - y_1)$, $t_2(x - x_2, y - y_2)$, $t_n(x - x_n, y - y_n)$, the input joint image may be expressed as

$$f(x, y) = r(x, y + y_0) + \sum_{i=1}^n t_i(x - x_i, y - y_i) \quad (2.5)$$

Applying Fourier transform to the input joint image by the lens L1 yields

$$F(u, v) = |R(u, v)| \exp[j\Phi_r(u, v) + jvy_0] + \sum_{i=1}^n |T_i(u, v)| \exp[j\Phi_{t_i}(u, v) - jux_i - jvy_i] \quad (2.6)$$

where $R(u, v)$ and $T(u, v)$ are the amplitudes, $\Phi_r(u, v)$ and $\Phi_i(u, v)$ are the phases of the Fourier transforms of $r(x, y)$ and $t_i(x, y)$, respectively; u and v are mutually independent frequency domain variables scaled by a factor of $\frac{2\pi}{\lambda}$, λ is the wavelength of collimating light, f is the focal length of lenses L1 and L2. The complex light distribution produced in the back focal plane of L1, called the JPS, is then detected by a square-law detector such as a CCD array. The JPS is given by

$$\begin{aligned}
 |F(u, v)|^2 = & |R(u, v)|^2 + \sum_{i=1}^n |T_i(u, v)|^2 + 2 \sum_{i=1}^n |T_i(u, v)| |R(u, v)| \cos[\Phi_i(u, v) - \Phi_r(u, v) \\
 & - ux_i - vy_i - 2vy_0] + 2 \sum_{i=1}^n \sum_{k=1, k \neq i}^n |T_i(u, v)| |T_k(u, v)| \cos[\Phi_i(u, v) - \Phi_{ik}(u, v) \\
 & - ux_i + ux_k - vy_i + vy_k]
 \end{aligned} \quad (2.7)$$

From Eq. (2.7), it is evident that the correlation output contains autocorrelation of the reference image and the input scene objects, cross-correlation between the reference image and the input scene objects, and cross-correlations between different input scene objects. The first term produces the auto-correlation of the reference image only while the second term produces the same for different input scene image only. The fourth term produces the cross-correlation between different input scene images without considering the reference image. Therefore, the first and second term of Eq. (2.7) produce the strong zero-order peak and the fourth term produces the false alarms. Only the third term generates the desired correlation peaks between the reference image and different input scene images.

2.2.2.2 Simulation Results

To simulate the performance of a classical joint transform correlator in case of multiple objects (target and non-targets) present simultaneously in the input scene, the same 256×256 zero-padded image matrix is taken for the input scene. Here again the alphabet 'E' is used as the reference image and is placed at the right side of each joint

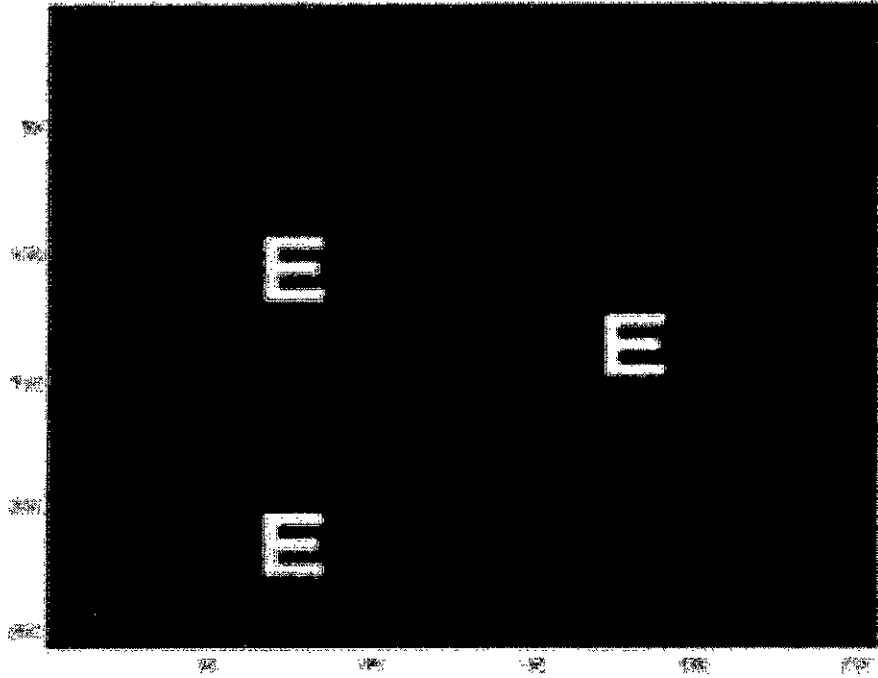


Fig 2.4 (a): Input joint image with identical targets

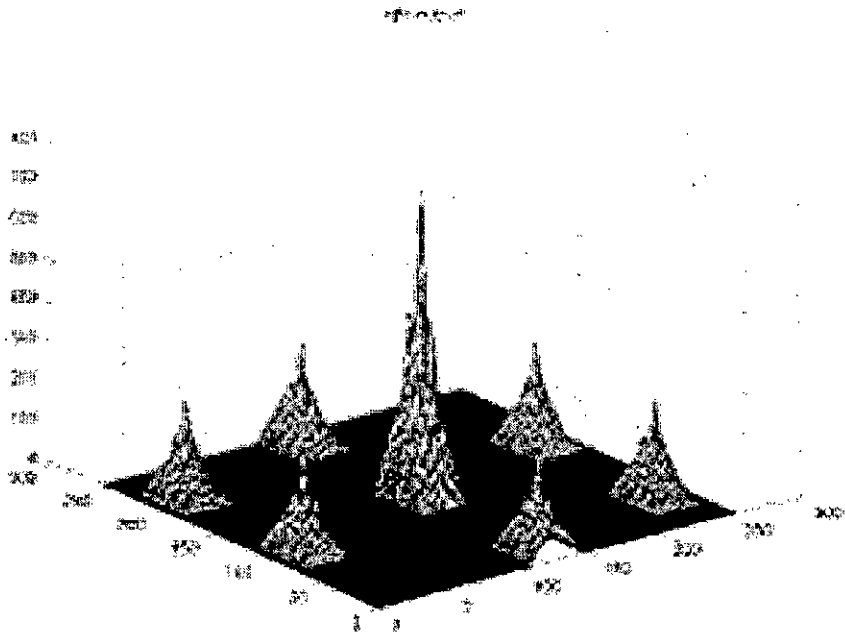


Fig. 2.4 (b): Classical JTC output of Fig. 2.4 (a)

image as shown in Fig 2.4 (a), 2.5 (a) and 2.6 (a). Three cases have been shown for three combinations of input objects.

Case 1: Two target images 'E' are placed at the left side of the joint image as shown in Fig 2.4 (a). The classical JTC output of Fig 2.4 (a) is shown in Fig 2.4 (b). The correlation output shows the strong zero-order peak at the center. Two pairs of cross-correlation peaks indicate two target images. But another pair of auto-correlation peaks between the targets produces the false alarm. If the intensity of the target images is higher than the reference image then the intensity of the false alarm may be higher than the desired cross-correlation peaks and therefore simply obscures the detection.

Case 2: One target image 'E' and one non-target image 'M' are the input objects placed at the left side of the joint image as shown in Fig 2.5 (a). The joint image is processed to find the correlation output of Fig 2.5 (b). The correlation output shows a strong zero-order peak, a pair of cross-correlation peaks for the target, a pair of cross-correlation peak for the non-target (which is smaller than the target peak) and a pair of cross-correlation peaks between the target and non-target.

Case 3: Multiple target images and multiple non-target images are the input objects placed at the left side of the input plane as shown in Fig 2.6 (a). The joint image of Fig 2.6 (a) has been processed to find the correlation output of Fig 2.6 (b). The correlation output shows the strong zero-order peak at the center of the output plane. It also shows several cross-correlation peaks for the targets and non-targets and several auto-correlation peaks among the target and non-target objects. Due to the presence of strong false alarms, it is quite difficult to locate the actual position of the target images.

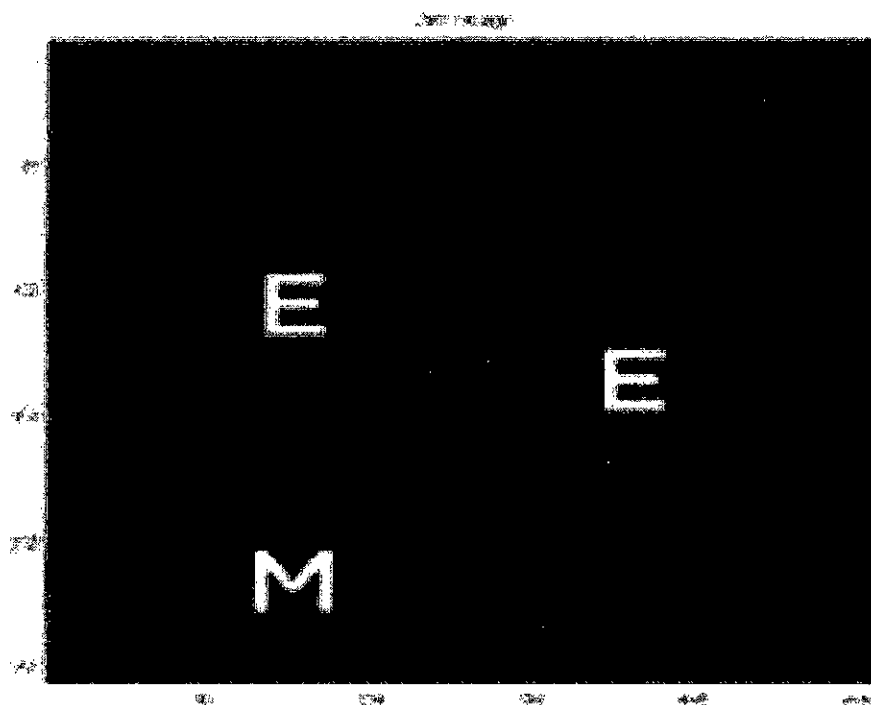


Fig 2.5 (a): Input joint image with one target and one non-target

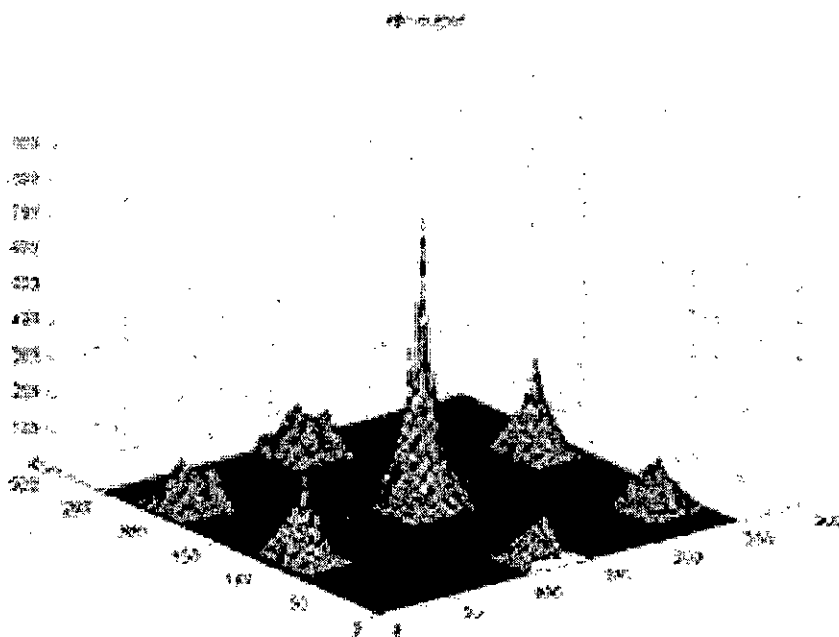


Fig. 2.5 (b): Classical JTC output of Fig. 2.5 (a)

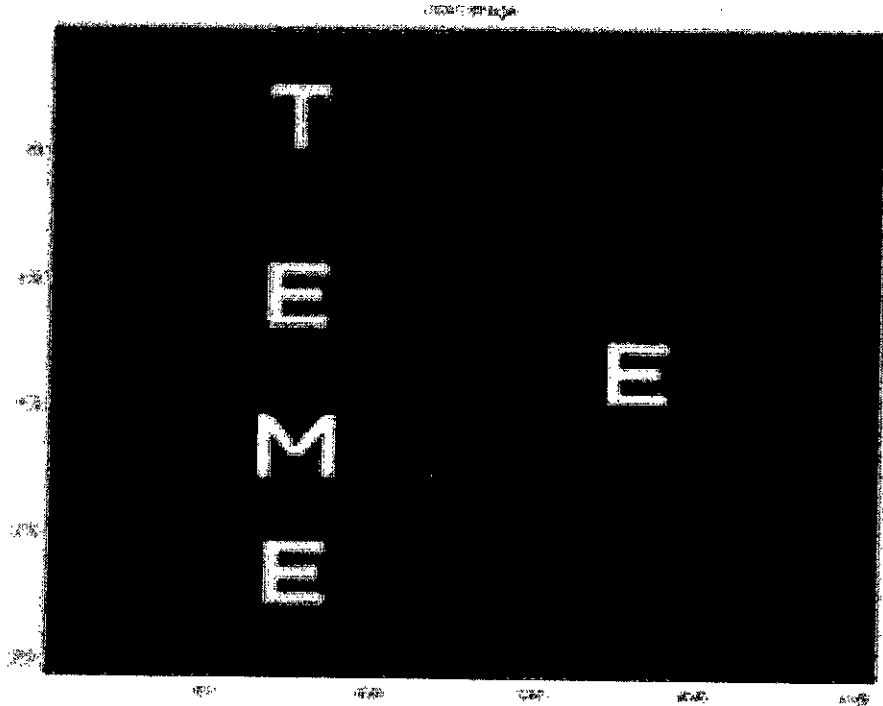


Fig 2.6 (a): Input joint image with multiple target and non-target objects

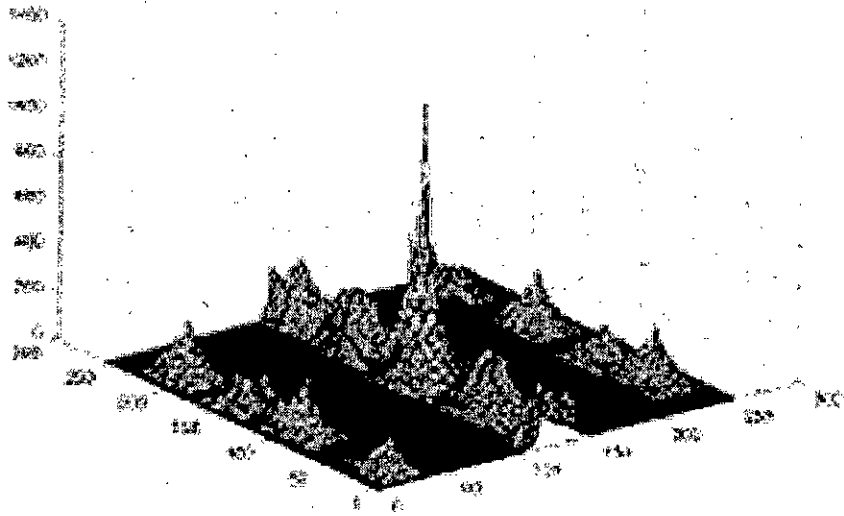


Fig. 2.6 (b): Classical JTC output of Fig. 2.6 (a)

From this study, it is obvious that presence of strong zero-order peak and false alarms make the multi-target detection process quite difficult in a classical JTC scheme.

2.3 Classical Joint Transform Correlation with Image Subtraction

2.3.1 Analysis

If the zero-order term of the joint power spectrum is eliminated then the on-axis correlation distribution can be avoided. Since the zero-order correlation are derived from the individual power spectrum of $r(x, y)$ and $t(x, y)$, these power spectrum can be eliminated from the joint power spectrum by the use of computer. After subtracting these two terms from Eq. (2.3), the JPS for single object input scene becomes

$$P(u, v) = |F(u, v)|^2 - |R(u, v)|^2 - |T(u, v)|^2 \quad (2.8)$$

Now inverse Fourier transform of Eq. (2.8) produces the final correlation output as

$$g(x', y') = r(-x' - 2x_0, -y') \otimes t^*(-x', -y') + r^*(x' - 2x_0, y') \otimes t(x', y') \quad (2.9)$$

In Eq. (2.9) there are only the cross correlation terms between the reference and the target image of the input scene.

For the multi-object input scene, the zero-order power spectra are derived from the Fourier transformation of the reference image $r(x, y + y_0)$ and the input scene image containing targets of $t_1(x - x_1, y - y_1), t_2(x - x_2, y - y_2), \dots \dots \dots t_n(x - x_n, y - y_n)$ respectively. By subtracting the reference only power spectra and the input scene only power spectra, we get the modified JPS for multi-object input scene as given by

$$\begin{aligned}
P(u, v) &= |F(u, v)|^2 - |R(u, v)|^2 - \sum_{i=1}^n |T_i(u, v)|^2 \\
&= \sum_{i=1}^n |R(u, v)| |T_i(u, v)| \exp[j(\Phi_r(u, v) - \Phi_{i_i}(u, v) + ux_i + v(y_0 + y_i))] \\
&\quad + \sum_{i=1}^n |R(u, v)| |T_i(u, v)| \exp[-j(\Phi_r(u, v) - \Phi_{i_i}(u, v) + ux_i + v(y_0 + y_i))]
\end{aligned} \tag{2.10}$$

Inverse Fourier transform of Eq. (2.10) gives the correlation output for the multi-object input scene as expressed by

$$\begin{aligned}
g(x', y') &= \sum_{i=1}^n r(-x' - x_i, -y' - y_i - 2y_0) \otimes t^*(-x', -y') \\
&\quad + \sum_{i=1}^n r^*(x' - x_i, y' - y_i + 2y_0) \otimes t(x', y')
\end{aligned} \tag{2.11}$$

In Eq. (2.11), again there are only cross-correlation terms between the reference and target images only.

The above process of subtracting the reference only power spectra and the input scene only power spectra from the joint power spectra at the Fourier plane is known as Fourier plane image subtraction technique. The Fourier plane image subtraction technique eliminates the zero-order term and other false alarms generated by the cross-correlation among different objects in the input scene.

2.3.2 Simulation Results

To eliminate the zero-order term and other false alarms, the Fourier plane image subtraction technique is employed and the simulation result is shown for the joint image as shown in Fig. 2.7 (a). The character 'E' on the right side represents the reference image and all the characters on the left side represent the input scene images. Thus the input scene contains both target and non-target objects. Fig. 2.7 (b) shows the classical JTC output of Fig. 2.7 (a) with Fourier plane image subtraction. From the figure, it is evident that the zero-order diffraction term and other false alarms are completely eliminated from the output plane and there exists only a pair of cross-correlation peaks

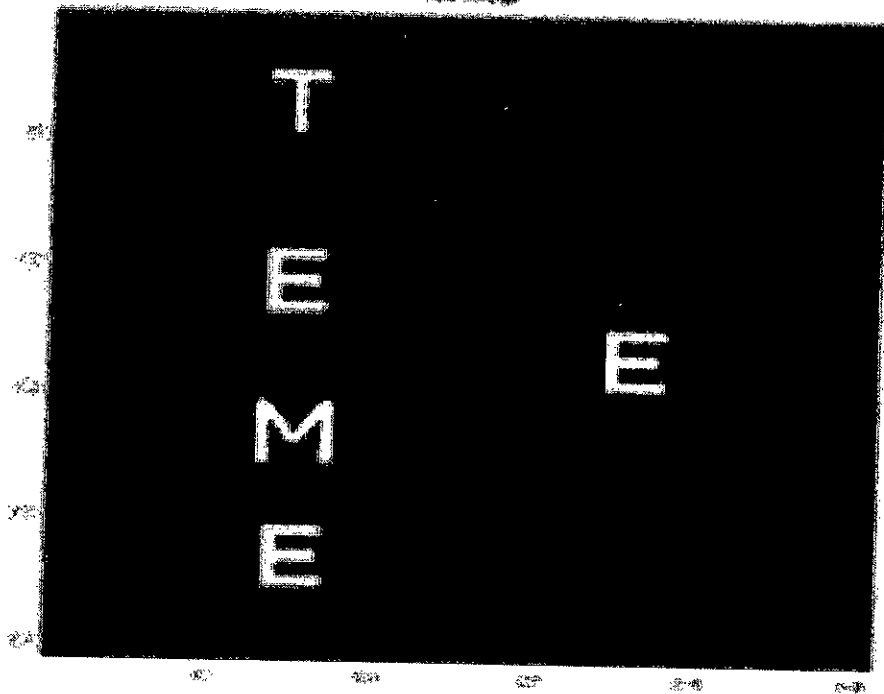


Fig. 2.7 (a): Input joint image with multiple identical targets and non-target objects

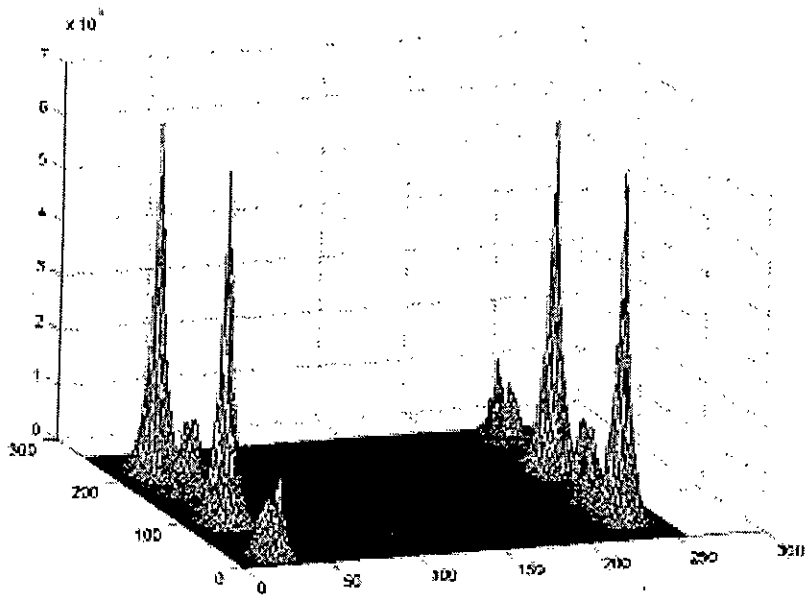


Fig. 2.7 (b): Classical JTC output after image subtraction of Fig. 2.7 (a)

for each input scene object. At the same time, the correlation peak intensity for the target object is comparatively higher than that of the non-target object.

Thus the Fourier plane image subtraction technique provides better detection of target objects while eliminating the false alarms and other extraneous signals. It also provides comparatively better utilization of space bandwidth product and therefore, input scene objects can be located as closely as we wish. But still there exists some problems with the Fourier plane image subtraction technique. From the correlation output, it is obvious that the target peak intensity is not much higher than the non-target peaks. Again, large correlation width and large side-lobes lower the optical efficiency and therefore, deteriorate the detection scheme.

2.4 Fringe-adjusted Joint Transform Correlation with Image Subtraction

2.4.1 Analysis

The classical JTC output with Fourier plane image subtraction suffers from lower optical efficiency and wide correlation peaks. To yield better correlation performance compared to alternate JTCs under various illumination conditions of the input scene, the fringe-adjusted JTC (FJTC) technique has been proposed. In FJTC technique, the JPS is multiplied by the real valued fringe-adjusted filter (FAF) transfer function before applying the inverse Fourier transformation to get the correlation output. The transfer function of the FAF is given by

$$H_{faf}(u, v) = \frac{C(u, v)}{D(u, v) + |R(u, v)|^2} \quad (2.12)$$

where $C(u, v)$ and $D(u, v)$ are either constants or functions of u and v . When $C(u, v)$ is properly selected, one can avoid an optical gain greater than unity. With very small values of $D(u, v)$, the pole problem otherwise associated with a inverse filter is

eliminated, and it is still possible to obtain a large auto-correlation peak. When $D(u, v) \ll |R(u, v)|^2$, $C(u, v)=1$, the transfer function of the FAF may be approximated as $H_{\text{far}}(u, v) \approx |R(u, v)|^{-2}$. The function $C(u, v)$ can also be used to suppress noise or band limit the signal or both. Thus, in FJTC, the amplitude matching can be implemented more effectively to yield sharper and larger correlation peak intensity. When $C(u, v)=1$ and $|R(u, v)|^2 \geq D(u, v)$, the FAF becomes a perfect real-valued inverse filter function unlike the VanderLugt filter, which consists of both magnitude and phase terms. Therefore, the FAF is more suitable for practical implementation and computation associated with the FAF can be performed before the practical implementation. Therefore, it has no detrimental effect on system processing speed.

For single target input scene, the modified JPS after image subtraction is multiplied by the transfer function of the FAF, $H_{\text{far}}(u, v)$ and is thus given by

$$G_r(u, v) = H_{\text{far}}(u, v) \times P_s(u, v) \approx |R(u, v)|^{-2} \times P_s(u, v) \quad (2.13)$$

For multiple target input scene, the modified JPS after image subtraction is multiplied by the transfer function of the FAF, $H_{\text{far}}(u, v)$ and is thus given by

$$G_m(u, v) = H_{\text{far}}(u, v) \times P_m(u, v) \approx |R(u, v)|^{-2} \times P_m(u, v) \quad (2.14)$$

Inverse Fourier transformation of Eq. (2.13) and Eq. (2.14) produce a pair of sharp delta-like correlation peak for each target object and almost negligible correlation peak for non-target object. Simulation results prove the effectiveness of this method.

2.4.2 Simulation Results

The simulation for the fringe-adjusted joint transform correlation has been performed with Fourier plane image subtraction so as to avoid the zero-order terms and unwanted false alarms. The input joint image used for the simulation purpose is shown in Fig. 2.8 (a). The simulation is done for various values of parameter $D(u, v)$ of the fringe-adjusted filter. The corresponding correlation output is shown in Fig. 2.8 (b) to Fig. 2.8(e).

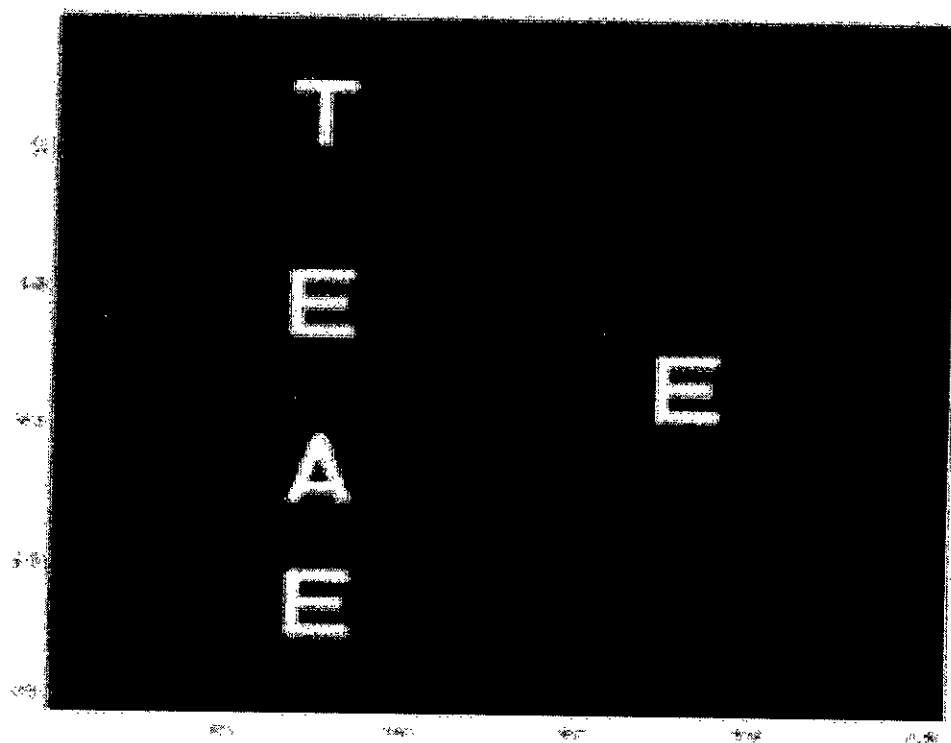


Fig. 2.8 (a): Input joint image with multiple identical targets and multiple non-target objects

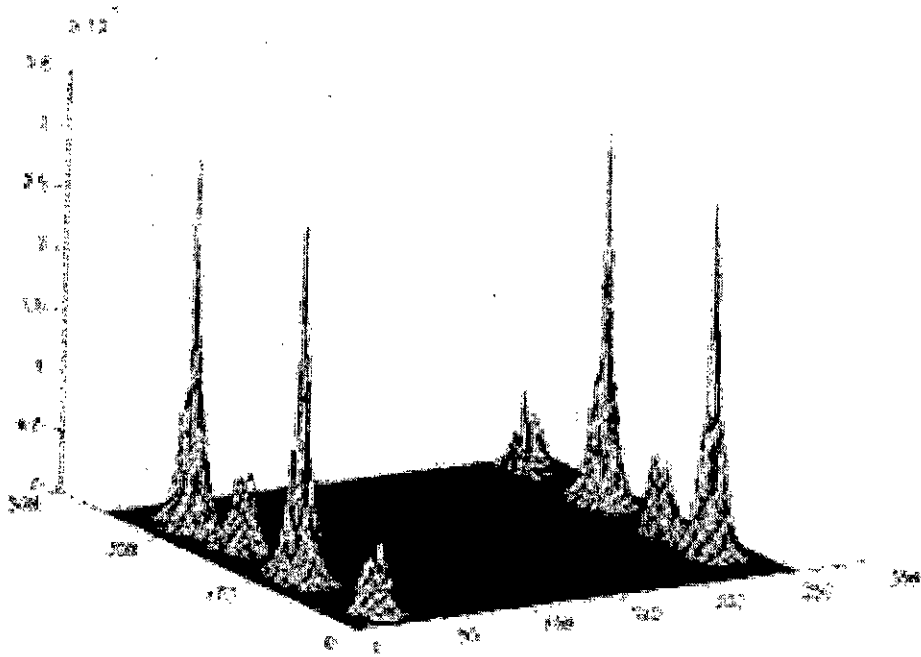


Fig. 2.8 (b): FJTC output of fig. 2.8 (a) with $C=1$ and $D=1e-1$

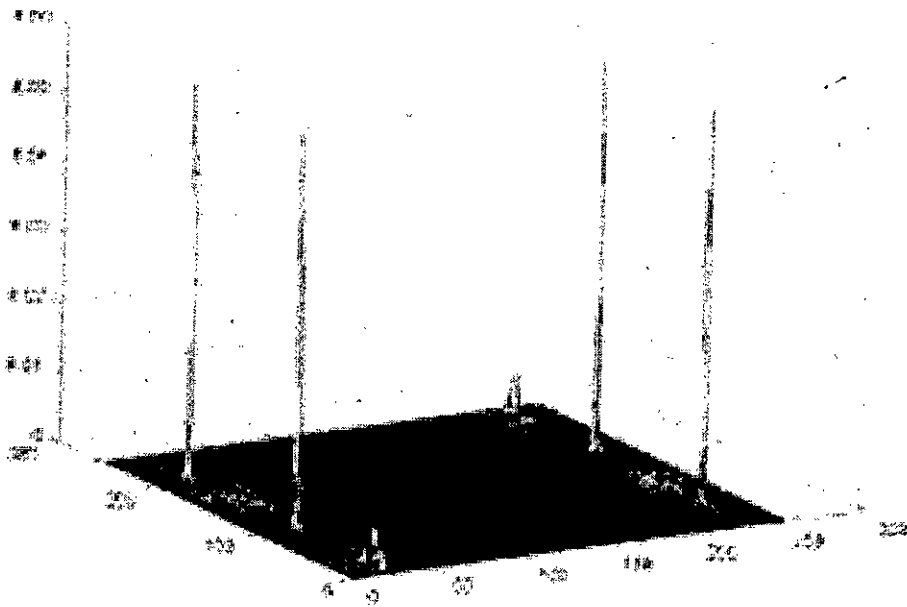


Fig. 2.8 (c): FJTC output fig. 2.8 (a) with $C=1$ and $D=1e-4$

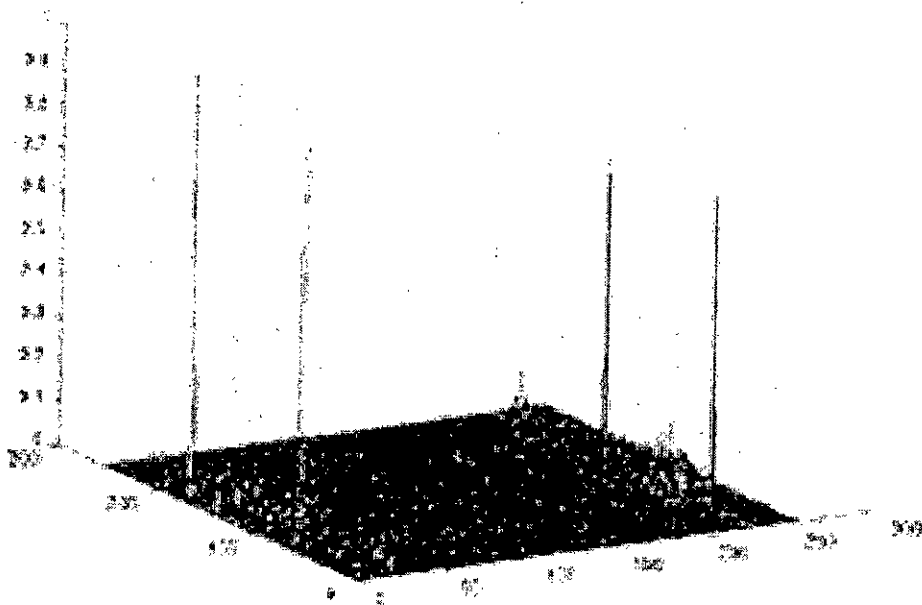


Fig. 2.8 (d): FJTC output fig. 2.8 (a) with $C=1$ and $D=1e-9$

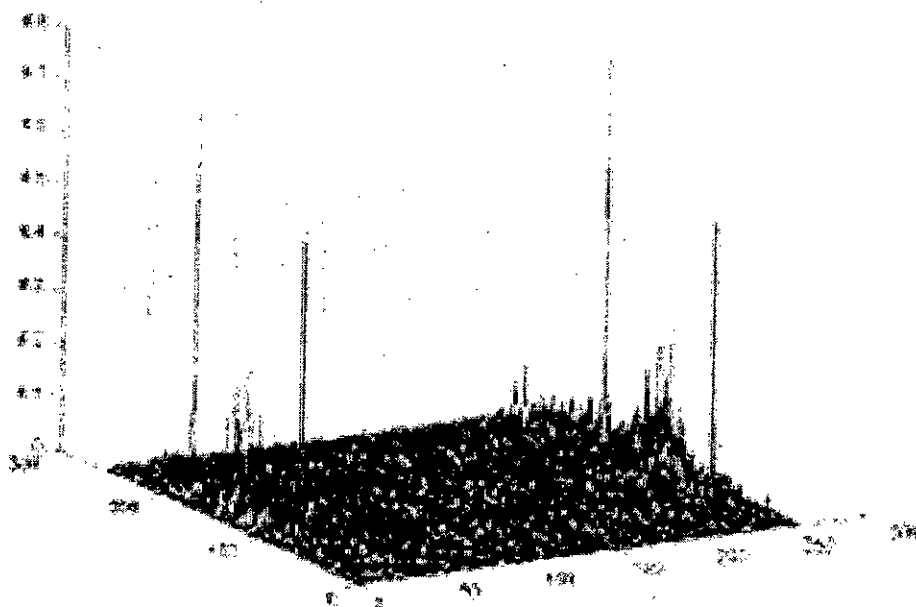


Fig. 2.8 (e): FJTC output fig. 2.8 (a) with $C=1$ and $D=1e-12$

From the correlation output as shown in Fig. 2.8 (b) to Fig. 2.8(e), it is obvious that the FJTC technique produces a pair of sharp delta-like correlation peak for each target. At the same time, it produces almost negligible correlation peaks for non-target objects. The target peaks are almost of same intensity or height. For noise free input scene, the value of $D(u,v)$ parameter of FAF, plays an important role for correlation output. As in this case, higher the value of $D(u,v)$ parameter from an optimal value, the target peaks become more and more sharper. But at the same time, the correlation plane generates spurious peaks and deteriorates the correlation performance. Now, if the value of $D(u,v)$ is set higher than the optimal value, then the width correlation peaks becomes more wider and again hampers the system performance. Therefore, selection of an optimal value of $D(u,v)$ parameter is a must which is done here by trial and error method. An intensive study is required to choose the best value of FAF parameters.

2.5.1 Fractional Power Fringe-adjusted Joint Transform Correlation

For noise-free input scenes, the fringe-adjusted JTC has been found to yield significantly better correlation output compared to alternate JTCs. However, for noisy input scenes, whenever the reference power spectrum contains very low values of $\left[|R(u,v)|^2\right]$, it may accentuate the noise component of the input scene, which may degrade the system performance. To overcome the aforementioned problems, and at the same time to utilize the fringe adjusted filter as a versatile tool for various applications, a fractional power fringe-adjusted JTC (FPFJTC) has been proposed, also termed as generalized fringe-adjusted JTC (GFJTC) [35]. In GFJTC the generalized fringe-adjusted filter transfer function is given by

$$H_{gfaf}(u,v) = \frac{C(u,v)}{D(u,v) + |R(u,v)|^m} \quad (2.15)$$

where m is a constant and $C(u,v)$ and $D(u,v)$ are either constants or functions of u and v . When $C(u,v)$ is properly selected, one can avoid an optical gain greater than unity. With

very small values of $D(u,v)$, the pole problem otherwise associated with an inverse filter can be avoided. By adjusting the values of m , a family of different types of real-valued filters can be achieved.

When $m=0$, $D(u,v)=0$ and $C(u,v)=1$, the GFAF-based JTC corresponds to classical JTC. In this case the transfer function of the filter becomes

$$H_{gfaf}^c = 1 \quad (2.16)$$

For $m=1$, $D(u,v) \ll |R(u,v)|^2$ and $C(u,v)=1$, the GFJTC technique corresponds to phase-only JTC and in this case the transfer function of the filter becomes

$$H_{gfaf}^p(u,v) = \frac{C(u,v)}{D(u,v) + |R(u,v)|} \quad (2.17)$$

When $m=2$, $D(u,v) \ll |R(u,v)|^2$ and $C(u,v)=1$, the GFAF-based JTC corresponds to fringe-adjusted JTC. In this case the transfer function of the filter becomes

$$H_{gfaf}^f(u,v) = \frac{C(u,v)}{D(u,v) + |R(u,v)|^2} \quad (2.18)$$

Thus, all important types of matched filter based correlators can be implemented in real time using the fractional power fringe-adjusted JTC while avoiding the limitations of matched filter based correlators. For noise free input scene, the fringe-adjusted JTC produces the highest correlation output. In case of noise corrupted input scene, the fractional power fringe-adjusted JTC produces better correlation output than fringe-adjusted JTC.

2.5.2 Simulation Results

To investigate the performance of the GFAF, we have chosen the same joint image as shown in fig. 2.9 (a). The character 'E' on the right half portion represents the reference image and all other characters on the left half portion represent the input scene image.

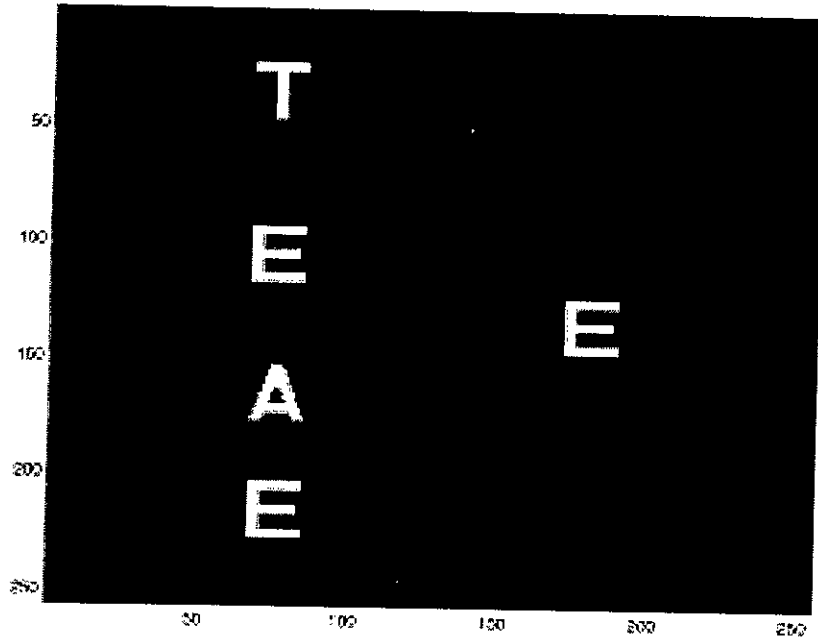


Fig. 2.9 (a): Binary joint image

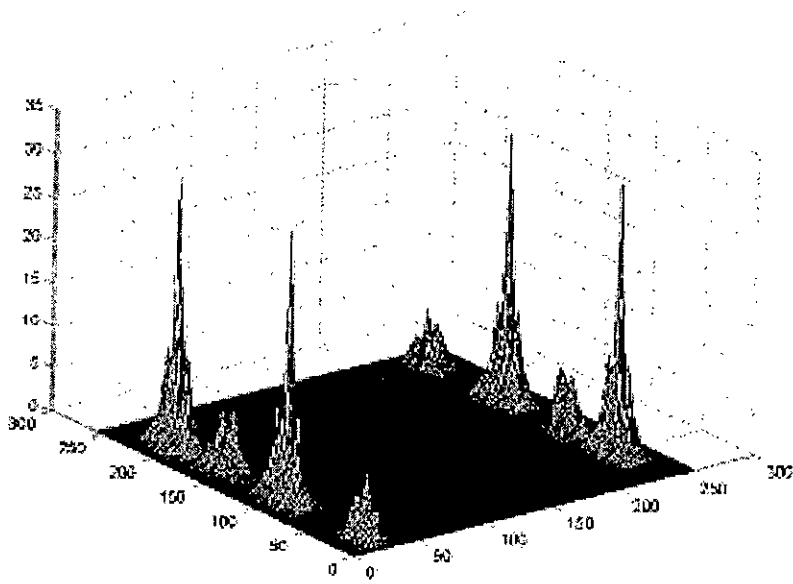


Fig. 2.9 (b): Performance of GFAF with $m=0$

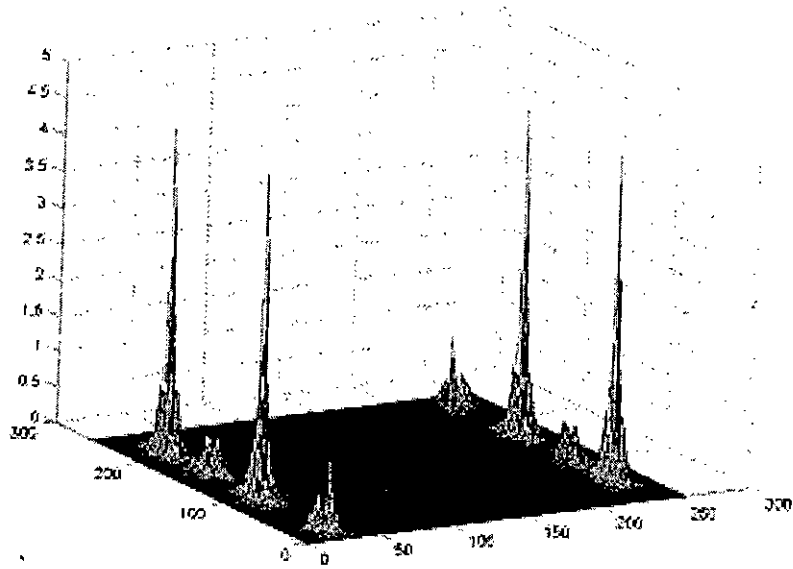


Fig. 2.9 (c): Performance of GFAF with $m=1$

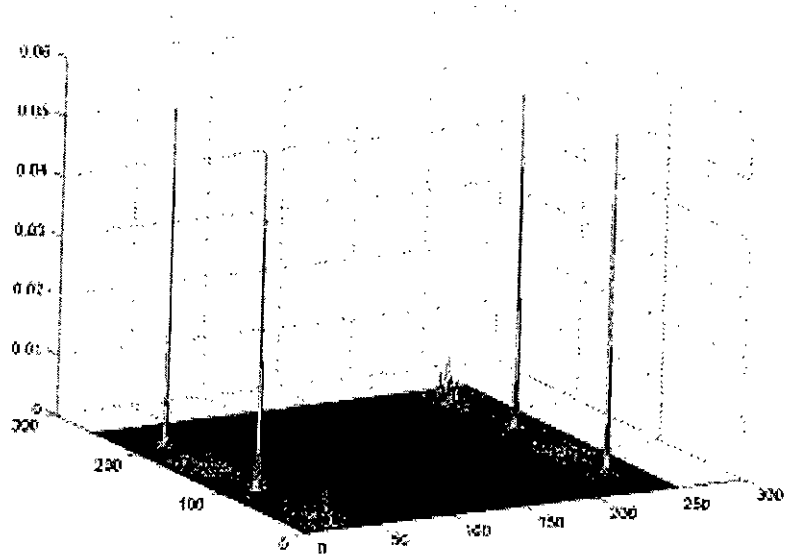


Fig. 2.9 (d): Performance of GFAF with $m=2$

Fig. 2.9 (b) to fig. 2.9 (d) show the effect of GFAF with $C(u,v)=1$ and an optimal value of $D(u,v)$ parameter i.e. $D(u,v)=1e-4$. For different values of m , GFAF may act as a classical JTC filter, phase only filter or fringe-adjusted filter. For $m=0$, it will act as a classical JTC filter; for $m=1$, it will act as a phase only filter and for $m=2$, it will act as a fringe-adjusted filter.

Fig. 2.9 (b) shows the effect of GFAF with $m=0$. Here the correlation peaks are wide and there is presence of side-lobes. There is also presence of wide correlation peaks for non-target objects. All these characteristics match with that of the classical JTC technique. Therefore, GFAF with $m=0$, acts as a classical JTC filter.

Fig. 2.9 (c) shows the effect of GFAF with $m=1$. Here the correlation peaks are much sharper than that of the classical JTC technique but still there is presence of side-lobes in the output plane. Though, there is presence of non-target peaks, the discrimination between target and non-target peaks are much higher. Thus, GFAF with $m=1$, acts a phase-only filter.

Fig. 2.9 (d) depicts the performance of GFAF with $m=2$. In this case sharp delta-like correlation peaks are generated for each target object and there is almost no or negligible correlation peaks for non-target objects. The performance of the correlation output is much higher than that of classical or phase-only JTC technique. Therefore, GFAF with $m=2$, acts a fringe-adjusted filter.

2.6 Phase-encoded Fringe adjusted Joint Transform Correlation

2.6.1 Analysis

For multiple target detection, either Fourier plane image subtraction [28], or correlation plane image subtraction [29], or phase-encoded FJTC [36] may be used depending on the system constraint. The Fourier plane and correlation plane image subtraction

techniques involve multiple processing steps and generate a pair of correlation peaks for each target. The phase-encoded fringe adjusted JTC (PFJTC) technique overcomes these problems by yielding one peak per target in one processing step. Fig. 2.10 shows the block diagram of a PFJTC technique.

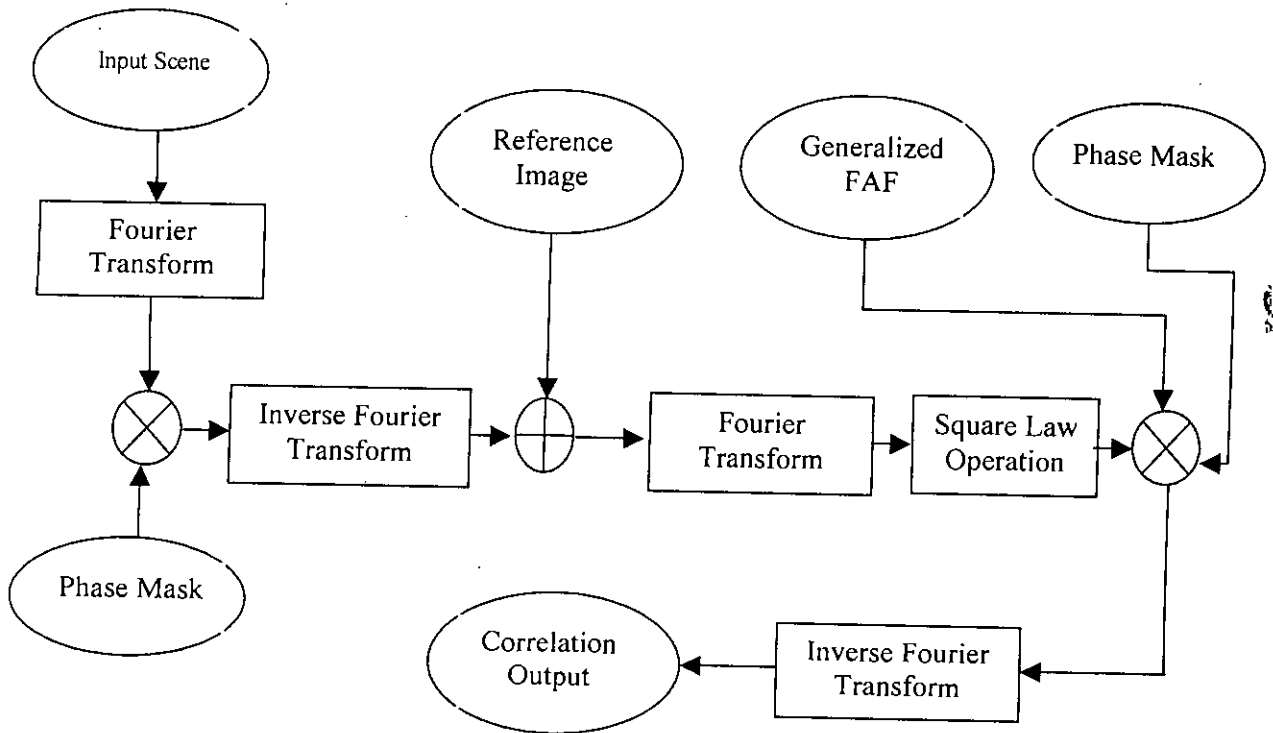


Fig 2.10: Block diagram of PFJTC Technique

In the PFJTC technique, a random phase mask, defined by

$$\Phi(u, v) = \exp[j\Psi(u, v)] \quad (2.19)$$

is multiplied by the input scene power spectrum, yielding

$$|T_{ph}(u, v)| = \sum_{i=1}^n |T_i(u, v)| \exp[j\Phi_{ii} - j(ux_i - vy_i)] \times \Phi(u, v) \quad (2.20)$$

An inverse Fourier transform of Eq. (2.20) yields the phase encoded input scene, given by

$$t_{ph}(x, y) = \sum_{i=1}^n t_i(x - x_i, y - y_i) \otimes \phi(x, y) \quad (2.21)$$

where \otimes represents convolution operation. The phase-encoded input scene of Eq. (2.21) is combined with the reference image to form the input joint image, which is then Fourier transformed to yield the phase-encoded JPS, given by

$$\begin{aligned} |F(u, v)|^2 &= |R(u, v)|^2 + \sum_{i=1}^n |T_i(u, v)|^2 \\ &+ \sum_{i=1}^n |R(u, v)| |T_i(u, v)| \exp[j(\Phi_r(u, v) - \Phi_{ii}(u, v) + ux_i + v(y_0 + y_i))] \times \Phi^*(u, v) \\ &+ \sum_{i=1}^n |R(u, v)| |T_i(u, v)| \exp[-j(\Phi_r(u, v) - \Phi_{ii}(u, v) + ux_i + v(y_0 + y_i))] \times \Phi(u, v) \\ &+ 2 \sum_{i=1}^n \sum_{k=1, k \neq i}^n |T_i(u, v)| |T_k(u, v)| \cos[\Phi_{ii}(u, v) - \Phi_{kk}(u, v) - u(x_i - x_k) - v(y_i - y_k)] \end{aligned} \quad (2.22)$$

where the superscript * represents the complex conjugate. The phase-encoded JPS is again multiplied by the same phase mask $\Phi(u, v)$ to yield the modified JPS, given by

$$\begin{aligned} O(u, v) &= |R(u, v)|^2 \Phi(u, v) + \sum_{i=1}^n |T_i(u, v)|^2 \Phi(u, v) \\ &+ \sum_{i=1}^n |R(u, v)| |T_i(u, v)| \exp[j(\Phi_r(u, v) - \Phi_{ii}(u, v) + ux_i + v(y_0 + y_i))] \\ &+ \sum_{i=1}^n |R(u, v)| |T_i(u, v)| \exp[-j(\Phi_r(u, v) - \Phi_{ii}(u, v) + ux_i + v(y_0 + y_i))] \times \Phi(u, v) \Phi(u, v) \\ &+ 2 \sum_{i=1}^n \sum_{k=1, k \neq i}^n |T_i(u, v)| |T_k(u, v)| \cos[\Phi_{ii}(u, v) - \Phi_{kk}(u, v) - u(x_i - x_k) - v(y_i - y_k)] \Phi(u, v) \end{aligned} \quad (2.23)$$

It is obvious from Eq. (2.23) that the third term, which yields the desired correlation peak, is not affected by the phase mask. Rather the unwanted or extraneous correlation signals are intentionally disturbed by the phase mask. If the phase mask is chosen randomly, the extraneous correlation signals can be scattered in various directions so that the zero-order terms and false alarms can automatically be eliminated from the

desired portion of the correlation plane. Also, only a single correlation peak is generated for each target [36].

To achieve high discrimination capability, phase-encoded JPS, is multiplied by FAF given by

$$O_{faf}(u, v) = \frac{C(u, v) \times O(u, v)}{D(u, v) + |R(u, v)|^2} \quad (2.24)$$

However, it is observed that the PFJTC technique is sensitive to noise and the number of target and non-target objects present in the input scene. The aforementioned problem can be tackled by adjusting the FAF parameter, $D(u, v)$, which is not always a convenient option for practical applications. With the increase of non-target objects in the input scene, the amplitude of noisy surface in the correlation plane starts to rise which may complicate the target detection process. Assume that the input scene contain two identical targets, $t_1(x-x_1, y-y_1)$ and $t_2(x-x_2, y-y_2)$, which are identical to the reference image $r(x, y)$. Then the modified JPS of Eq. (2.24) can be expressed as

$$O_{faf}(u, v) = \frac{C(u, v)}{D(u, v) + |R(u, v)|^2} \times \left[\begin{aligned} &|R(u, v)|^2 \Phi(u, v) + 2 |R(u, v)|^2 \Phi(u, v) \\ &+ \sum_{i=1}^2 |R(u, v)|^2 \exp[j(ux_i + v(y_0 + y_i))] \\ &+ \sum_{i=1}^2 |R(u, v)|^2 \exp[-j(ux_i + v(y_0 + y_i))] \times \Phi(u, v) \\ &+ 2 \sum_{i=1}^2 \sum_{k=1, k \neq i}^2 |R(u, v)|^2 \cos[-ux_i - ux_k - vy_i - vy_k] \times \Phi(u, v) \end{aligned} \right] \quad (2.25)$$

From Eq. (2.25), it is evident that all terms involve constant magnitude spectra, and only the third term produces a sharp correlation peak while the remaining correlation peaks are steared outside the correlation due to the application of the phase mask.

Now consider a case involving one target $t_1(x-x_1, y-y_1)$ and a non-target object $t_2(x-x_2, y-y_2)$. Then, Eq. (2.24) can be written as

$$\begin{aligned}
O_{\text{ref}}(u, v) = & \frac{C(u, v)}{D(u, v) + |R(u, v)|^2} \times [2 |R(u, v)|^2 \times \Phi(u, v) \\
& + |T_2(u, v)|^2 \times \Phi(u, v) + |R(u, v)|^2 \exp[j(ux_1 + v(y_0 + y_1))] \\
& + |R(u, v)|^2 \exp[-j(ux_1 + v(y_0 + y_1))] \times \Phi(u, v) \\
& + |T_2(u, v)| \times |R(u, v)| \exp[j(ux_2 + v(y_0 + y_2))] \\
& + |T_2(u, v)| \times |R(u, v)| \exp[-j(ux_2 + v(y_0 + y_2))] \times \Phi(u, v) \\
& + 2 \sum_{i=1}^2 \sum_{k=1, k \neq i}^2 T_i(u, v) \| T_k(u, v) | \cos[\Phi_r(u, v) - \Phi_{ik}(u, v) - ux_i - ux_k - vy_i - vy_k] \times \Phi(u, v)
\end{aligned}
\tag{2.26}$$

In Eq. (2.26), only the first, third and fourth terms contain constant magnitude spectra, while the remaining terms contain magnitude spectra amplitude-modulated by the corresponding non-target objects. Therefore, $D(u, v)$ must be selected to ensure that the desired correlation peak is significantly higher than false alarms.

2.6.2 Simulation Results

The performance of PFJTC technique is investigated for binary images of character as shown in Fig. 2.11 (a). In the Figure, the character 'E' on the right side represents the reference image and the characters 'E', 'A', 'T' and 'E' on the left side represent the input scene images. Thus, as in this case, the input scene contains both target and non-target images.

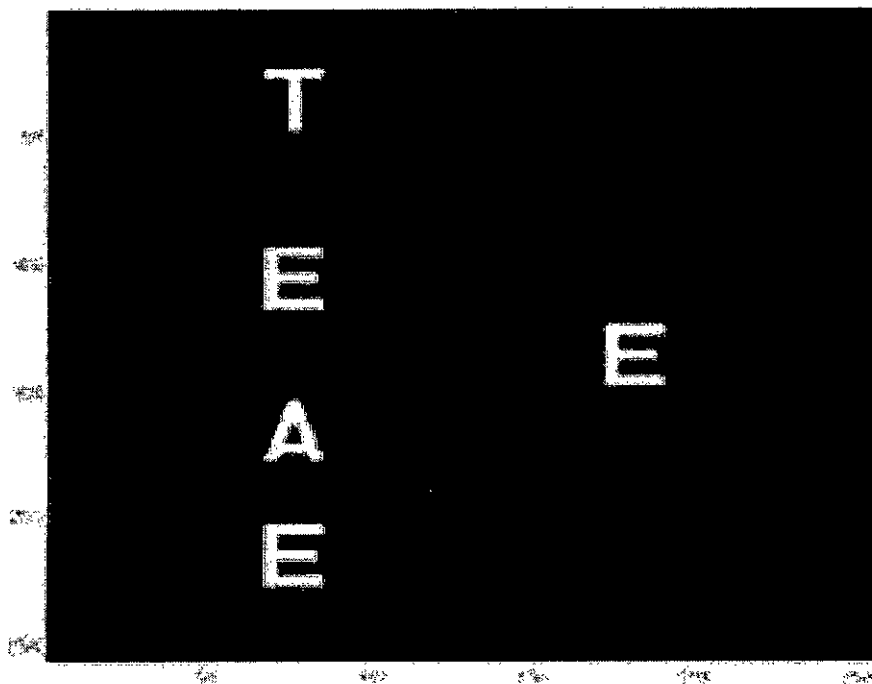


Fig. 2.11 (a): Input joint image with multiple identical targets and multiple non-target objects

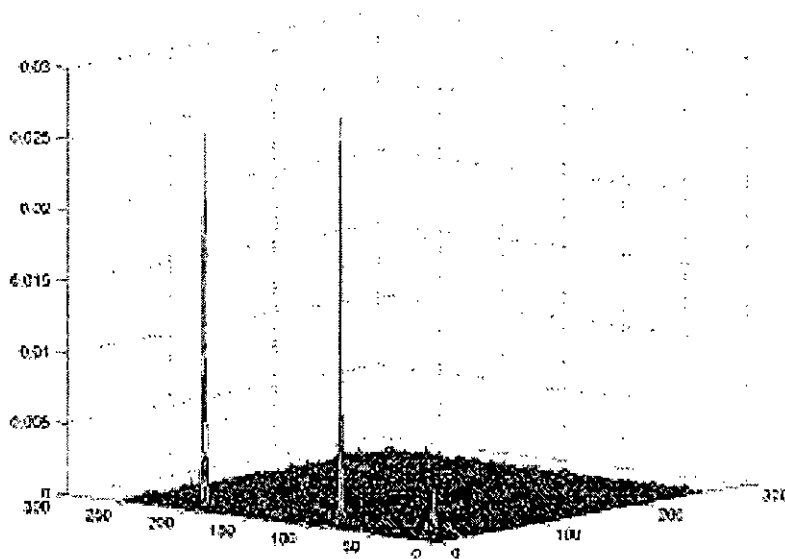


Fig. 2.11 (b): PFJTC output of fig. 2.11 (a) with $C=1$ and $D=1e-4$

The simulation is performed for an optimal value of $D(u,v)$ to get the best performance of fringe-adjusted filter. Here the values used for FAF parameters are $C(u,v)=1$ and

$D(u,v)=10^{-4}$. Fig. 2.11 (b), shows the PFJTC output for the joint image of Fig. 2.11 (a). From Fig. 2.11 (b), it can be easily deduced that the PFJTC technique efficiently eliminates the zero-order terms and other false alarms. This technique also produces a single delta-like sharp peak per each target instead of a pair of peaks and thus ensures better utilization of space bandwidth product.

For a pair of peaks per target object, it is always a complicated task to locate the exact position of the object. In FJTC technique, with the increase of target objects, the correlation plane is over crowded with pair of peaks and makes the target detection job more and more tedious. On the other hand, in PFJTC technique only a single peak is produced for each target and makes the target detection job easier. Here all the processing tasks can be performed in a single step and therefore, this technique is always a preferred one where processing time is a severe constraint. However, the performance of the system is very much input scene dependent and therefore, a modification is proposed in the next chapter.

2.7 Conclusion

This chapter describes the various JTC techniques for optical correlation purposes. A classical JTC technique is the simplest one but it suffers from large correlation width, strong zero-order term and low optical efficiency. Image subtraction technique with classical JTC technique, eliminates the false alarms and zero-order terms present in the classical JTC technique. But here still the correlation peaks are wide and there is presence of side-lobes in the output plane. To overcome this problem, the modified JPS after image subtraction, is multiplied by the transfer function of a real valued fringe-adjusted filter. Incorporation of FAF, produces sharp delta-like correlation peaks for each target object and there is almost none or negligible correlation peaks for non-target objects. Instead of producing a single fringe-adjusted filter, a family of filters is generated by using a generalized FAF transfer function. Here by varying a single

parameter (i.e. m), the GFAF can be performed as a classical JTC filter, phase-only filter or fringe-adjusted filter. All of the above techniques are multi-step processing techniques and therefore, these cannot be used where processing time is a severe constraint. To overcome this problem, a PFJTC technique is employed where a random phase mask is used to steer the unnecessary signal outside of the correlation plane. In PFJTC technique, all the processing is performed in a single step and furthermore, it produces single correlation peak for each target object instead of a pair of peaks. Therefore, PFJTC technique provides ultra-high-speed processing with better utilization of space bandwidth product of output plane. Simulation results validate the performance of all the techniques.

Chapter 3

EFFICIENT MULTIPLE TARGET DETECTION ALGORITHM

3.1 Introduction

If the input scene contains multiple identical target objects, then in addition to desired cross-correlation peaks between the reference and target object, additional peaks will be generated due to dc terms and other false alarms produced by the correlation of the target objects themselves. Presence of these extraneous signals complicates the target detection process to a great extent. The exclusion of correlation peaks produced by similar target or non-target objects eliminates the possibility of false alarms in the output plane. Fourier plane image subtraction, correlation plane image subtraction or power-spectrum subtraction technique may be employed to remove the zero-order diffraction term and other false alarms. All these techniques require a multi-step processing technique and therefore these cannot be applied where processing time is crucial constraint. Moreover, existing JTCs utilize at most 50% of the input plane spatial light modulator (SLM) to display the input scene and only 50% of the output plane to detect the correlation output. Thus these techniques cause a poor utilization of space bandwidth product (SBP). The phase-encoded FJTC technique can overcome this problem by using separate input and reference planes and yields a single correlation peak per target instead of a pair of peak. Thus PFJTC ensures the better utilization of space bandwidth product. However, for multi-object target in noisy input scenes, the PFJTC technique may require the adjustment of a parameter used in the FAF formulation for successfully detecting a target. To overcome the aforementioned problem and, at the same time, to enhance the correlation peak intensity, a phase-shifted PFJTC (PSPFJTC) technique is proposed here, which involves the phase-shifted power spectrum subtraction technique and the PFJTC technique. This technique requires only

one extra processing step compared to PFJTC technique but the performance of the system is much than that of the PFJTC technique.

3.2 Multi-target Detection

3.2.1 Theoretical Analysis

Let $r(x, y + y_0)$ represents the reference image and $t(x, y - y_0)$ represents the input scene containing n objects $t_1(x - x_1, y - y_1)$, $t_2(x - x_2, y - y_2)$, $t_n(x - x_n, y - y_n)$. Fig. 3.1 shows the block diagram of the proposed scheme. A random phase mask is used in this technique as expressed in Eq. 2.19. Here, the reference image is Fourier transformed and fed to two channels, as shown in Fig. 3.1, where one channel introduces 180° phase shift. Then random phase mask is applied to both the channels and the resultant signals are given by

$$S_1(u, v) = |R(u, v)| \exp[j\Phi_r(u, v) + jvy_0] \times \Phi(u, v) \quad (3.1)$$

$$S_2(u, v) = |R(u, v)| \exp[j\pi] \exp[j\Phi_r(u, v) + jvy_0] \times \Phi(u, v) \quad (3.2)$$

Applying inverse Fourier transform to Eqs. (3.1) and (3.2), we get

$$s_1(x, y) = r(x, y + y_0) \otimes \phi(x, y) \quad (3.3)$$

$$s_2(x, y) = -r(x, y + y_0) \otimes \phi(x, y) \quad (3.4)$$

The input scene is then added to both channels to form two input joint images, given by

$$f_1(x, y) = r(x, y + y_0) \otimes \phi(x, y) + \sum_{i=1}^n t_i(x, y - y_i) \quad (3.5)$$

$$f_2(x, y) = -r(x, y + y_0) \otimes \phi(x, y) + \sum_{i=1}^n t_i(x, y - y_i) \quad (3.6)$$

The 180° phase-shifted version of the reference image, $s_2(x, y)$, can be displayed in a SLM by adding a dc value as suggested by Yu *et. al.* [30]. Applying Fourier transform to Eqs. (3.5) and (3.6), we get

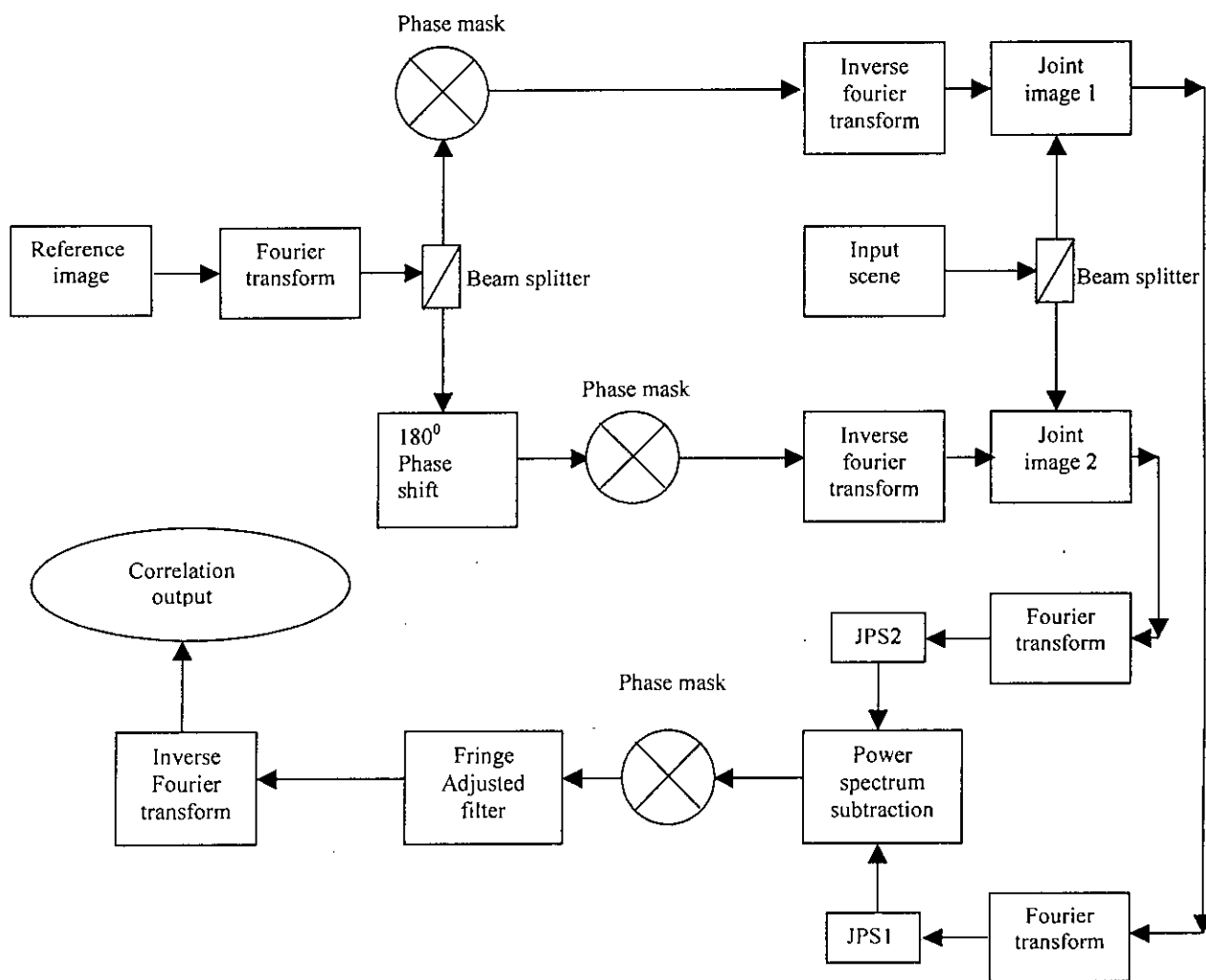


Figure 3.1: Architecture of the proposed PSPFJTC scheme

$$\begin{aligned}
|F_1(u, v)|^2 &= |R(u, v)|^2 + \sum_{i=1}^n |T(u, v)|^2 + \sum_{i=1}^n R(u, v)T_i^*(u, v)\exp[j(vy_0 - vy_i)] \times \Phi(u, v) \\
&+ \sum_{i=1}^n R^*(u, v)T_i(u, v)\exp[-j(vy_0 - vy_i)] \times \Phi^*(u, v) + \sum_{i=1}^n \sum_{k=1, k \neq i}^n T_i(u, v)T_k^*(u, v)\exp[-j(vy_i - vy_k)] \\
&+ \sum_{i=1}^n \sum_{k=1, k \neq i}^n T_i^*(u, v)T_k(u, v)\exp[j(vy_i - vy_k)]
\end{aligned} \tag{3.7}$$

and

$$\begin{aligned}
|F_2(u, v)|^2 &= |R(u, v)|^2 + \sum_{i=1}^n |T(u, v)|^2 - \sum_{i=1}^n R(u, v)T_i^*(u, v)\exp[j(vy_0 - vy_i)] \times \Phi(u, v) \\
&- \sum_{i=1}^n R^*(u, v)T_i(u, v)\exp[-j(vy_0 - vy_i)] \times \Phi^*(u, v) + \sum_{i=1}^n \sum_{k=1, k \neq i}^n T_i(u, v)T_k^*(u, v)\exp[-j(vy_i - vy_k)] \\
&+ \sum_{i=1}^n \sum_{k=1, k \neq i}^n T_i^*(u, v)T_k(u, v)\exp[j(vy_i - vy_k)]
\end{aligned} \tag{3.8}$$

The JPS corresponding to Eqs. (3.7) and (3.8) can be calculated in parallel and does not have detrimental effect on the processing speed. The phase-encoded JPS of Eq. (3.8) is then subtracted from that of Eq. (3.7) and the corresponding output is multiplied by the phase mask yielding

$$\begin{aligned}
P(u, v) &= [|F_1(u, v)|^2 - |F_2(u, v)|^2] \times \Phi(u, v) \\
&= 2 \left[\begin{aligned} &\sum_{i=1}^n R(u, v)T_i^*(u, v)\exp[j(vy_0 - vy_i)] \times \Phi(u, v) \\ &+ \sum_{i=1}^n R^*(u, v)T_i(u, v)\exp[-j(vy_0 - vy_i)] \times \Phi^*(u, v) \end{aligned} \right] \times \Phi(u, v) \\
&= 2 \sum_{i=1}^n R(u, v)T_i^*(u, v)\exp[j(vy_0 - vy_i)] \times \Phi(u, v) \times \Phi(u, v) \\
&+ 2 \sum_{i=1}^n R^*(u, v)T_i(u, v)\exp[-j(vy_0 - vy_i)] \times \Phi^*(u, v) \times \Phi(u, v)
\end{aligned} \tag{3.9}$$

It is obvious from Eq. (3.9) that the zero-order term and the false alarms are eliminated due to the application of the random phase mask. The modified JPS obtained in Eq. (3.9) is then multiplied by the FAF transfer function, yielding

$$\begin{aligned}
P_{fof}(u, v) &= P(u, v) \times H(u, v) \\
&= \frac{2C(u, v)}{D(u, v) + |R(u, v)|^2} \times \left[\sum_{i=1}^n R(u, v) T_i^*(u, v) \exp[j(vy_0 - vy_i)] \times \Phi^2(u, v) \right. \\
&\quad \left. + 2 \sum_{i=1}^n R^*(u, v) T_i(u, v) \exp[-j(vy_0 - vy_i)] \right]
\end{aligned} \tag{3.10}$$

Finally, the inverse Fourier transformation of Eq. (3.10) produces the desired correlation output. Assume that the input scene contains two objects, where $t_1(x, y-y_1)$ is the target object and $t_2(x, y-y_2)$ is the non-target object. Assuming $C(u, v)=1$ for unity gain and $|R(u, v)|^2 \gg D(u, v)$, Eq. (3.10) becomes

$$\begin{aligned}
P_{fof}(u, v) &\approx 2P(u, v) \times |R(u, v)|^{-2} \\
&\approx 2 \exp[-jvy_1] + 2 \frac{T_2^*(u, v)}{R(u, v)} \exp[j(\Phi_r(u, v) - \Phi_{t2}(u, v) - vy_2)] + \dots \dots
\end{aligned} \tag{3.11}$$

Here, the first term produces a delta like correlation output for the target object, but the second term contributes negligible correlation output due to the mismatch between the reference and non-target objects.

3.2.2 Edge Extraction of Images

Edge extracted images produce higher correlation peak intensity and therefore enhance target detection performance. Here the two joint images as expressed by Eq. (3.5) and (3.6), are edge extracted by using Roberts operator. The kernel of the extractor is given by

$$R^+ = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad R^- = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{3.12}$$

Convolution of these two kernels with the two joint images and summation of the absolute values of the outputs give the edges of the images. This is an all-digital process. Since this technique requires the edge extraction of input scene images that are unknown before the real time operation, this technique cannot be used where processing time is a severe constraint.

3.2.3 Simulation Results

Performance of the proposed PSPFJTC technique is investigated by considering a binary input scene involving characters as shown in Fig. 3.2 (a), and a gray-level input scene containing tanks and truck, as shown in Fig. 3.3 (a). In Fig. 3.2 (a), the letter E on the right side is used as the reference image and a set of four binary characters E, A, T, E, is used as the input scene. In Fig. 3.3 (a), the tank in the upper half of the input joint image represents the reference image and the set of two tanks and a truck represents the input scene. The simulation tests are performed using the MATLAB software package.

The correlation output obtained using classical JTC technique is shown in Figs. 3.2 (b) and 3.3 (b) for binary and gray-level input scenes, respectively, from where it is evident that the classical JTC yields a poor correlation output; strong zero order term which is truncated, as shown in figures 3.2 (b) and 3.3 (b), to make visible the desired cross-correlation peaks; and a pair of false correlation peaks due to autocorrelation between two similar input scene targets. To eliminate the zero-order term and false alarms, and to suppress the effects of input scene clutter, Fourier plane image subtraction is applied [15] and the corresponding correlation output for Figs. 3.2 (a) and 3.3 (a) are shown in Figs. 3.2 (c) and 3.3 (c), respectively. Although the Fourier plane image subtraction improves the correlation output of a classical JTC, the width of the correlation peak is still very broad.

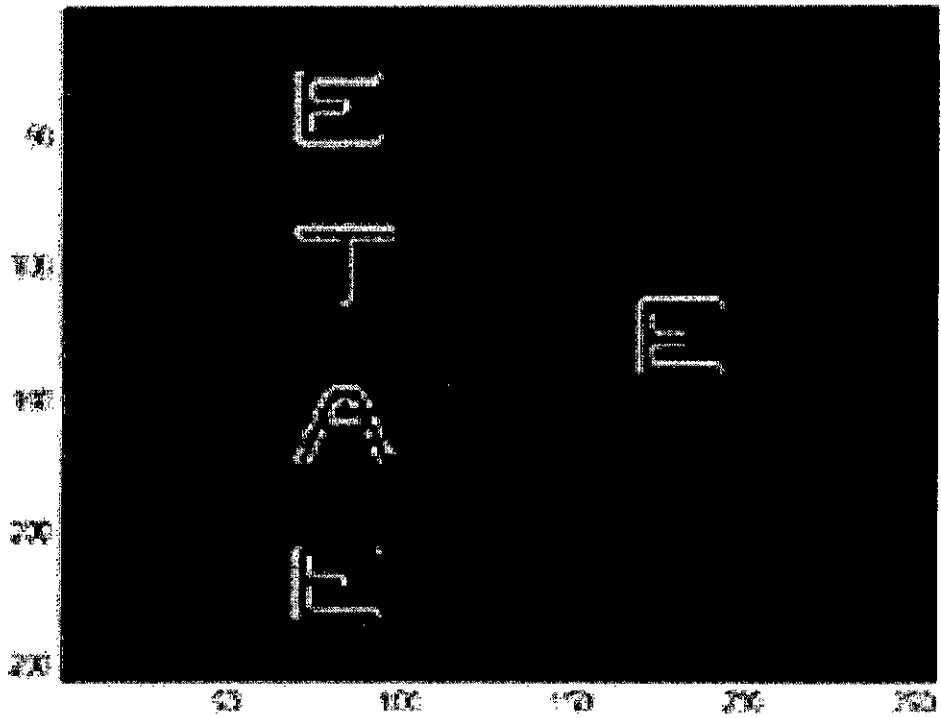


Fig. 3.2 (a): Binary joint input image

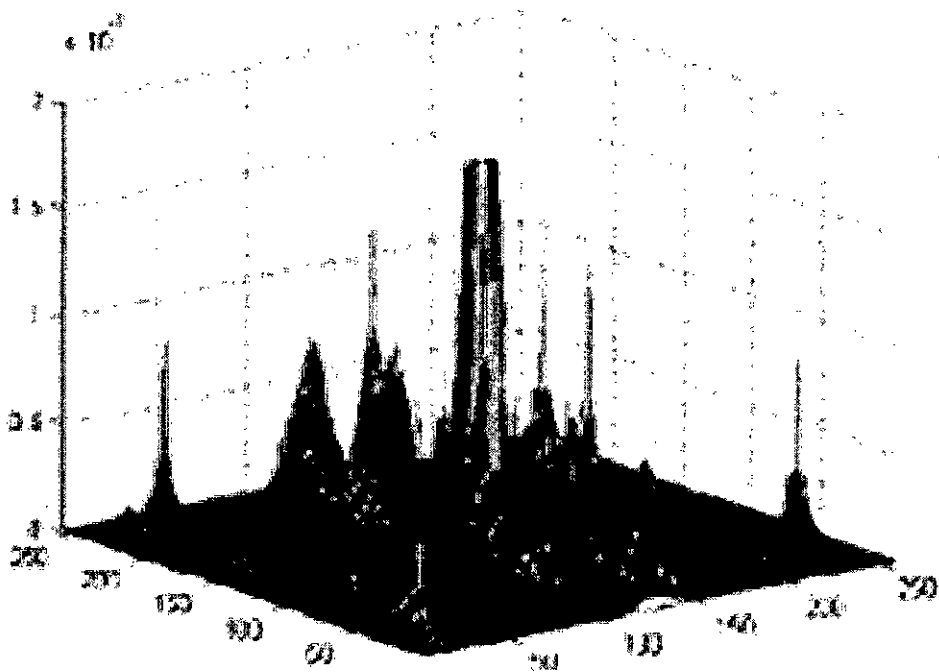


Fig. 3.2 (b): Classical JTC output of fig. 3.2 (a)

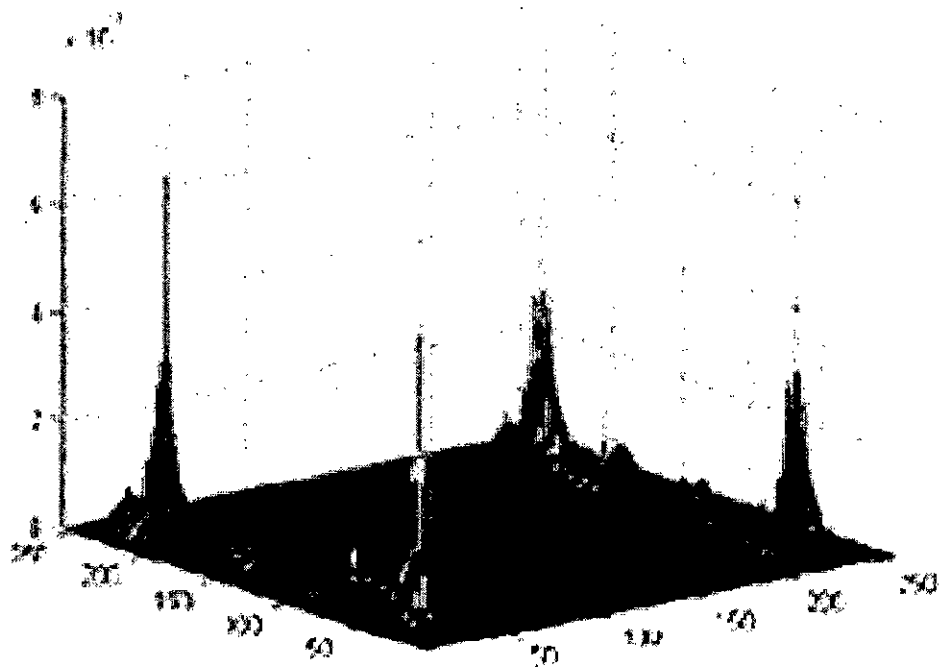


Fig. 3.2 (c): JTC output of fig. 3.2 (a) after image subtraction

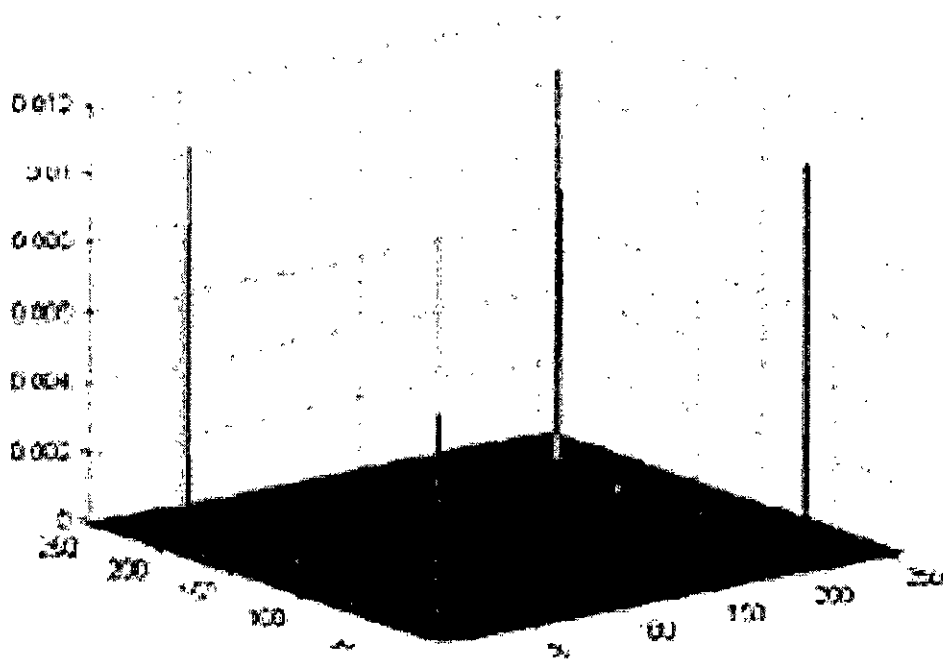


Fig. 3.2 (d): FJTC output of fig. 3.2 (a) with image subtraction

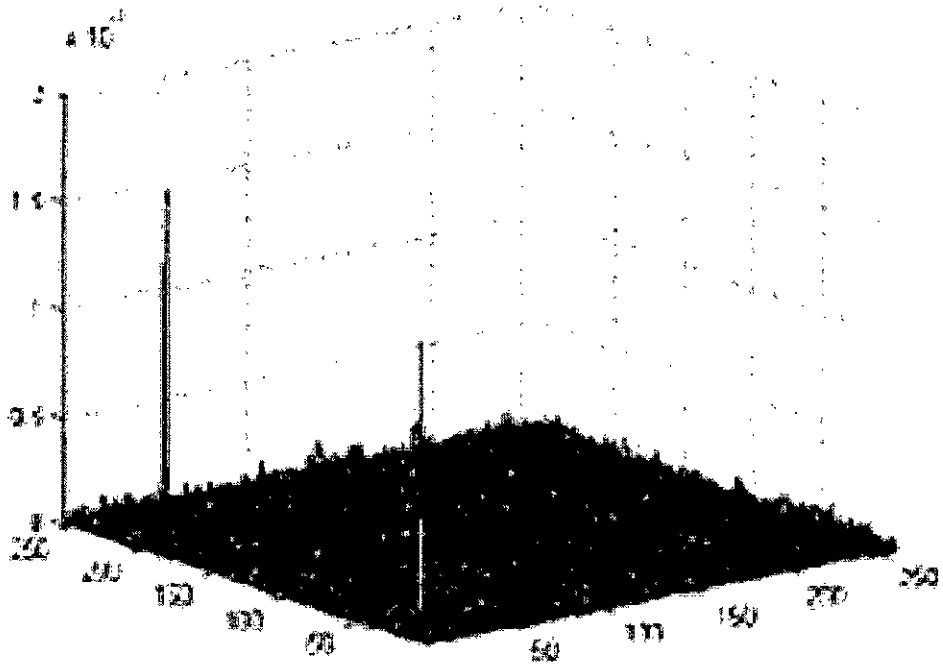


Fig. 3.2 (e): PFJTC output of fig. 3.2 (a)

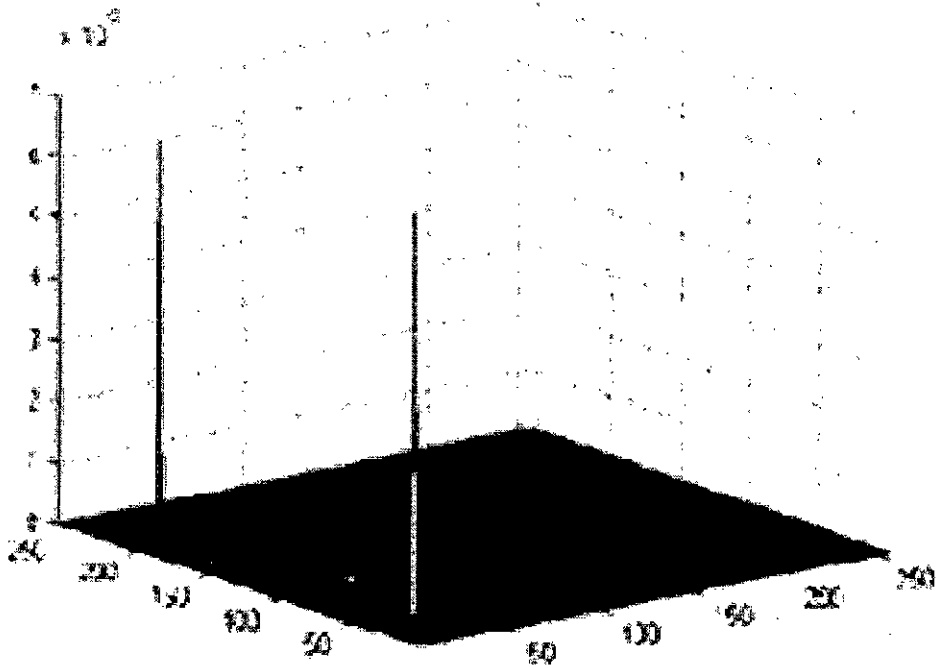


Fig. 3.2 (f): PS PFJTC output of fig. 3.2 (a)

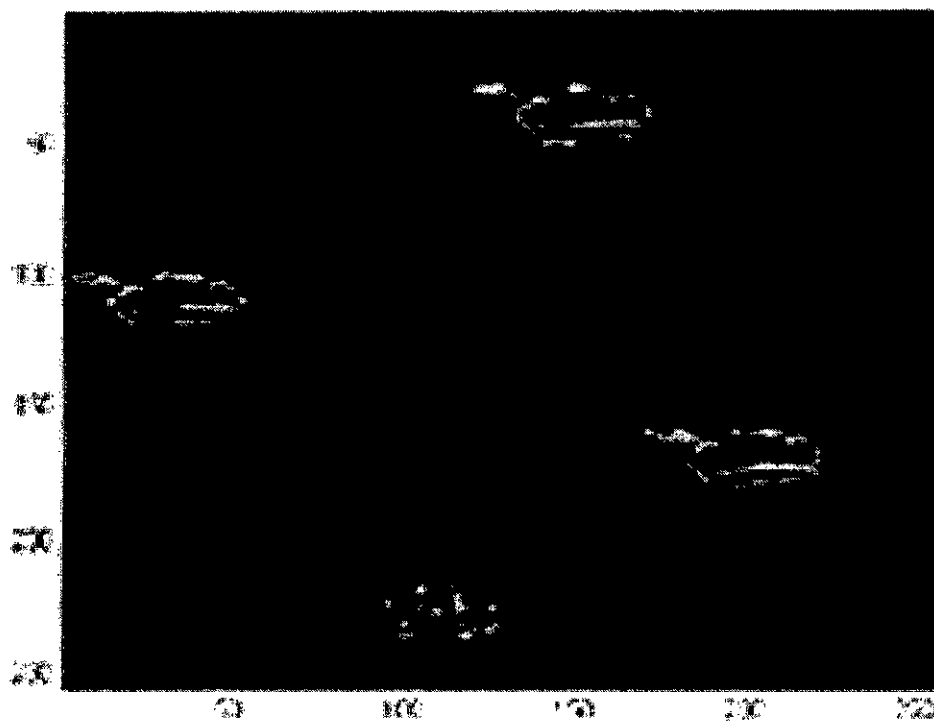


Fig. 3.3 (a): Gray level joint input image

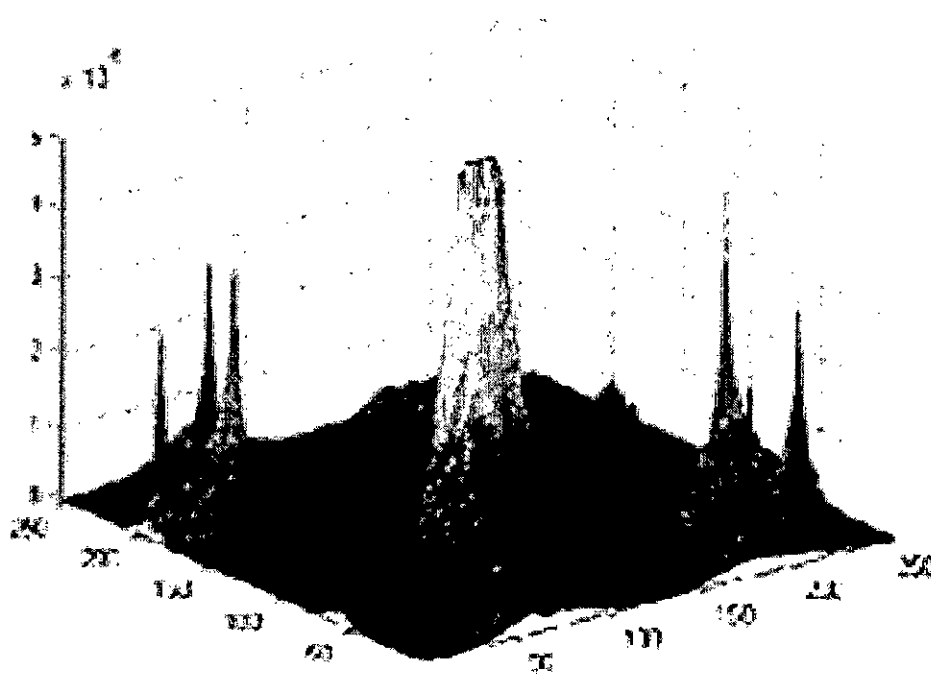


Fig. 3.3 (b): Classical JTC output of fig. 3.3 (a)

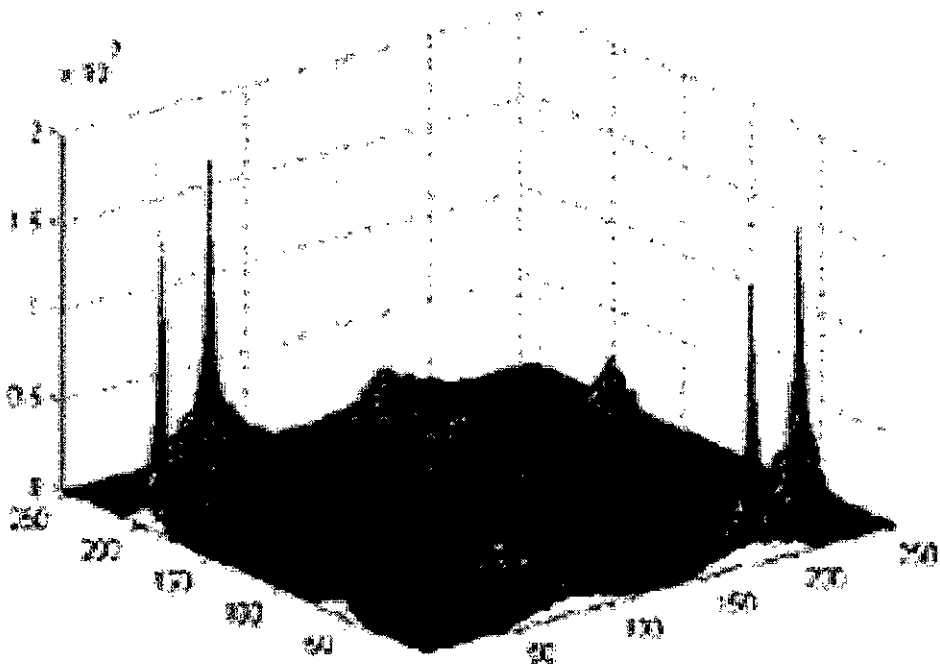


Fig. 3.3(c): JTC output after image subtraction of fig. 3.3 (a)

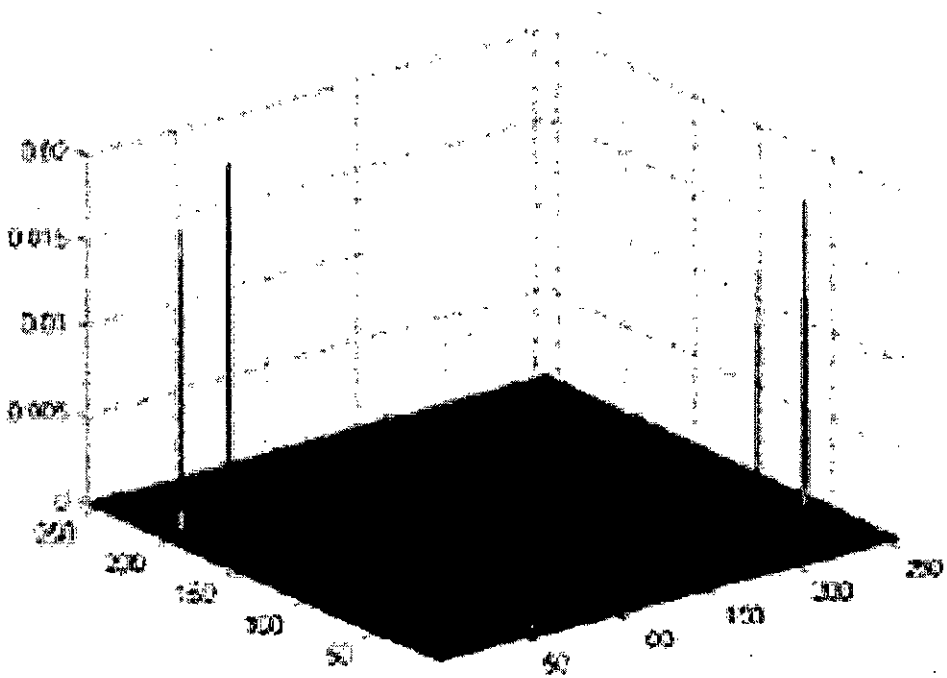


Fig. 3.3 (d): FJTC output with image subtraction of fig. 3.3 (a)

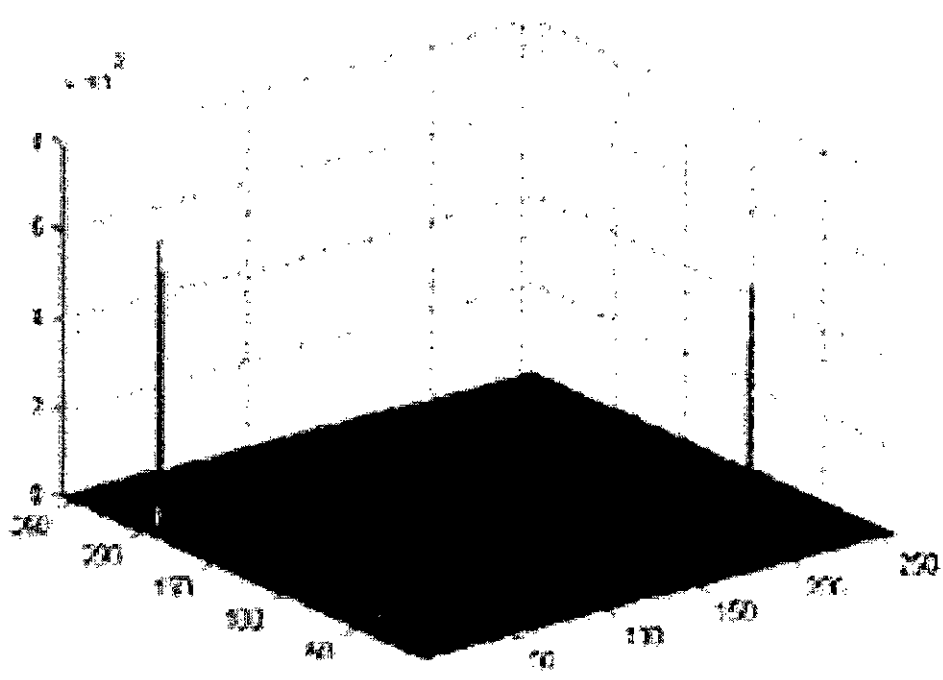


Fig. 3.3 (e): PFJTC output of fig. 3.3 (a)

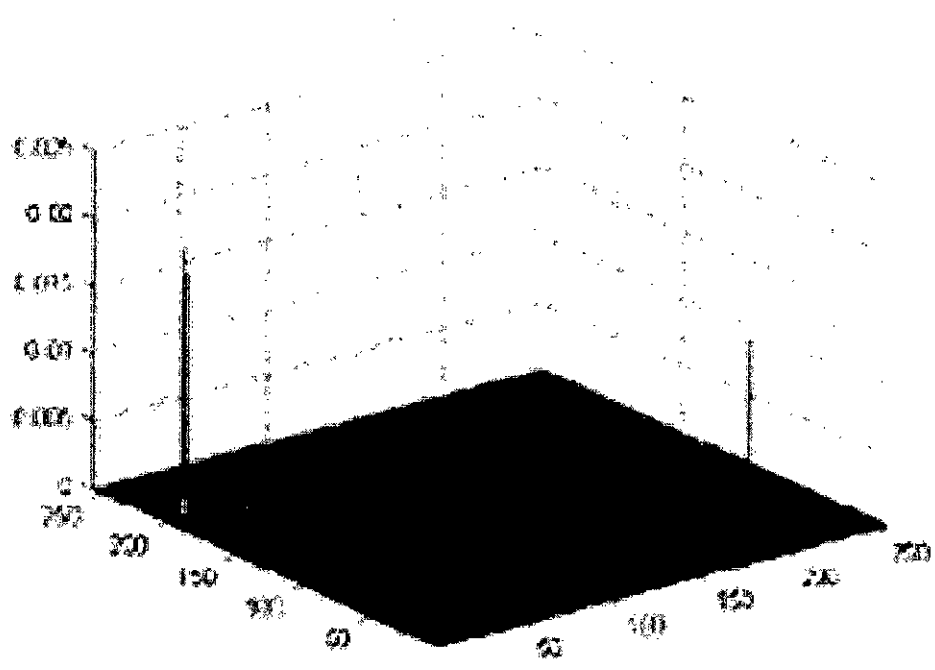


Fig. 3.3 (f): Phase shifted PFJTC output of fig. 3.3 (a)

When the FJTC technique is applied to the input scenes of Figs. 3.2 (a) and 3.3 (a), the corresponding correlation output using Fourier plane image subtraction are shown in Figs. 3.2 (d) and 3.3 (d), respectively. The simulation test was performed by setting the parameters $C(u,v) = 1$ and $D(u,v) = 10^{-4}$ and the figures depict that the FJTC technique yields a pair of delta-function-like correlation peaks for each input scene target.

Figures 3.2 (e) and 3.3 (e) show the correlation output generated by the PFJTC whereas figures 3.2 (f) and 3.3 (f) by the proposed PSPFJTC technique for the binary and gray-level images of Figs. 3.2 (a) and 3.3 (a), respectively. In both cases a single peak is generated for each target object and thus ensuring better SBP utilization.

To compare the performance of the PFJTC technique with the proposed PSPFJTC technique, an input scene with identical targets (five tank images) is considered as shown in Fig. 3.4 (a). The correlation outputs for the PFJTC and PSPFJTC techniques are shown in Figs. 3.4 (b) and 3.4 (c), respectively. Here the same filter parameters, $C(u,v) = 1$ and $D(u,v) = 10^{-4}$, are used for both the techniques. Figures 3.4 (b) and 3.4 (c) show that both PFJTC and PSPFJTC techniques produce one correlation peak per target, however, the correlation peak intensity (CPI), is higher in PSPFJTC than that in PFJTC and PSPFJTC satisfies the ECP (equal correlation peak) intensity criteria for all identical targets.

Next simulation is carried out for different cases of input scene having 5 objects where the number of targets and number of non-targets are varied gradually replacing each desired target tank by each non-target truck till the input image contains only one target tank and four non-target trucks. Both the PFJTC and PSPFJTC technique can produce distinct correlation peaks with the same FAF parameters as in previous simulation. The results are summarized in Table 3.1 for different combinations of input scene and as an example, the result is shown for one condition having two tanks and 3 trucks in Fig. 3.5 (a).

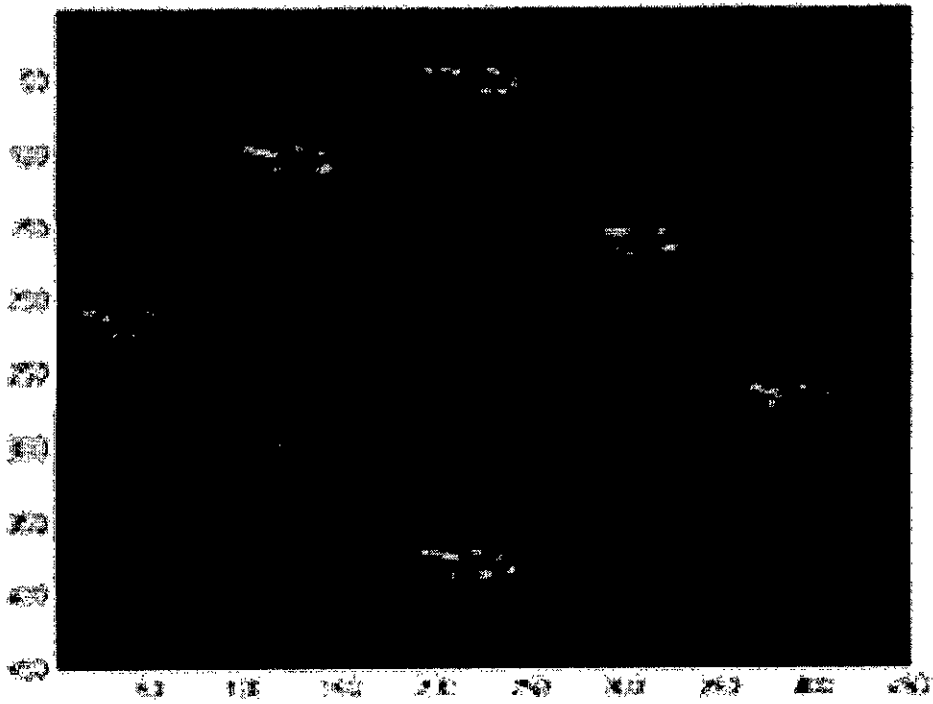


Fig. 3.4 (a): Gray level joint input scene of 5 tanks as target, without noise

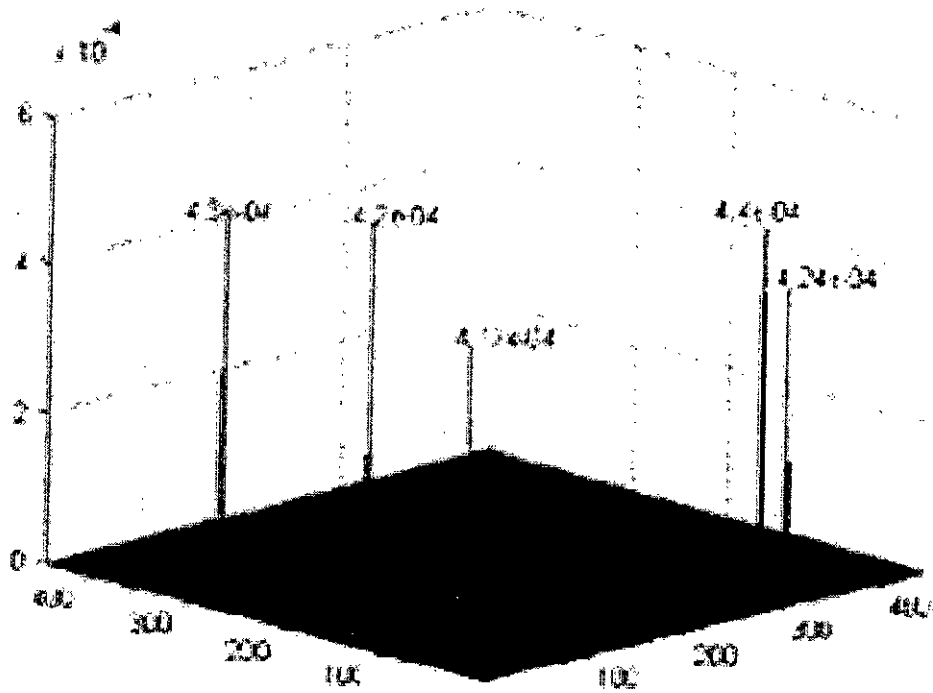


Fig. 3.4 (b): PFJTC output of fig. 3.4 (a) with $D=10^{-4}$

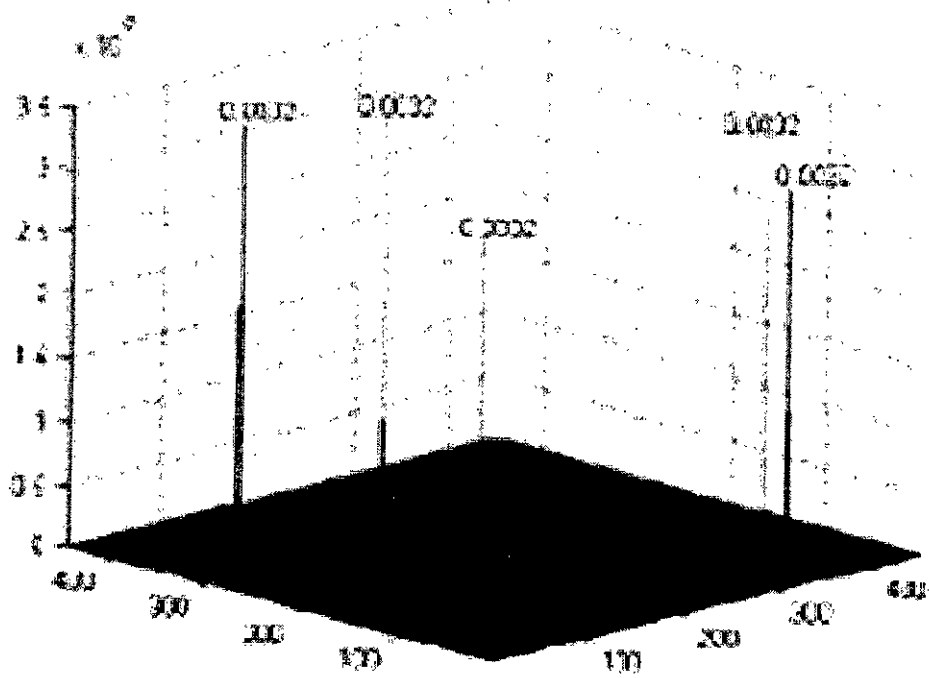


Fig. 3.4 (c): PSPFJTC output of fig. 3.4 (a) with $D=10^{-4}$

99112

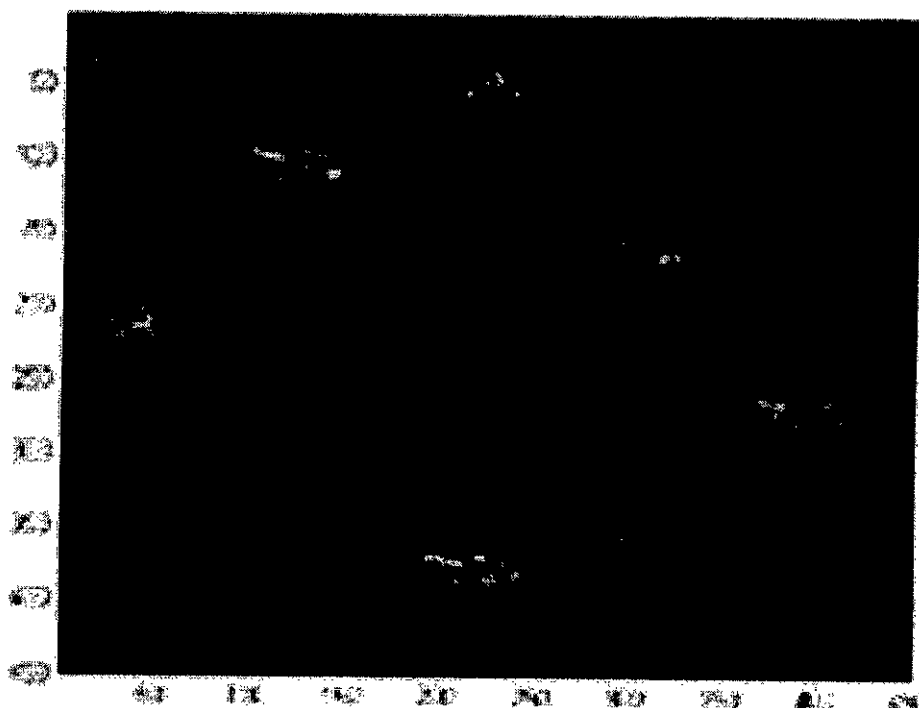


Fig. 3.5 (a): Gray level joint input scene without noise of 2 tanks as target and 3 trucks as non-target

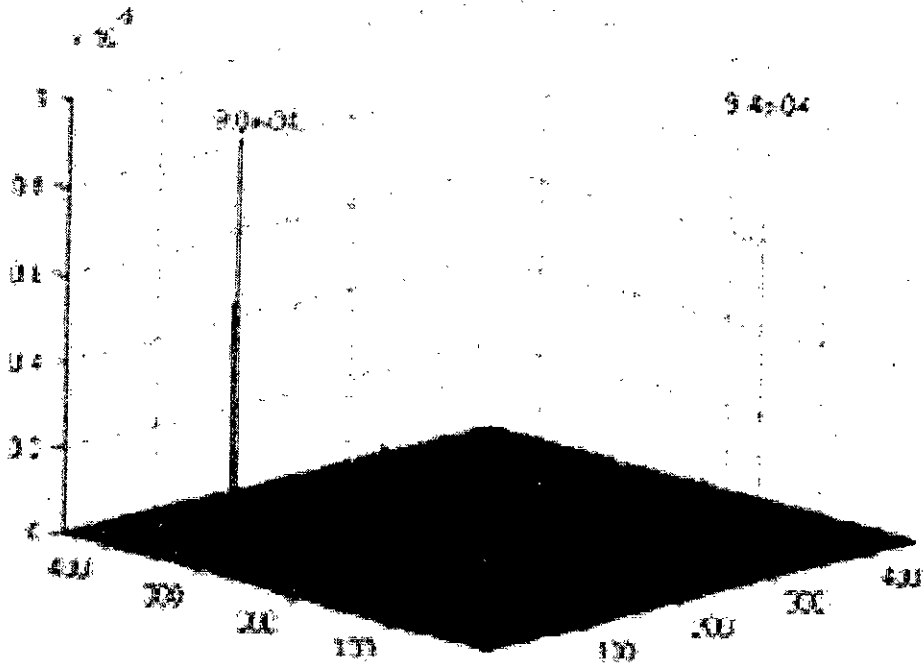


Fig. 3.5 (b): PFJTC output of fig. 3.5 (a) with $D=10^{-4}$

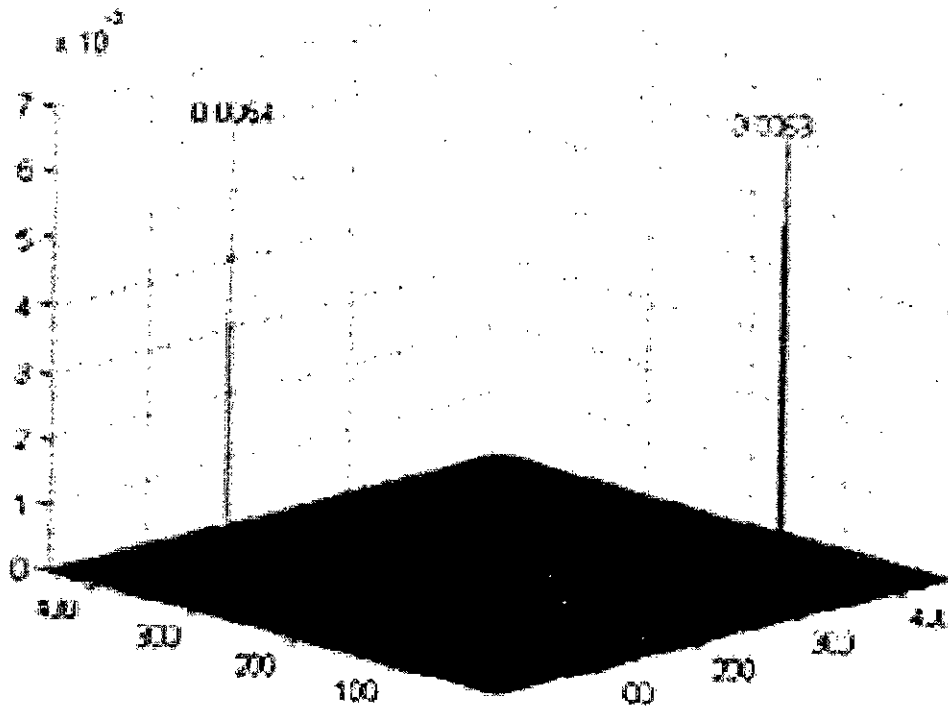


Fig. 3.5 (c): PS PFJTC output of fig. 3.5 (a) with $D=10^{-4}$

To investigate the noise robustness of the PSPFJTC technique, additive noise with signal-to-noise ratio (SNR) of 0 dB is added to the input scene as shown in Fig. 3.6 (a) and 3.7 (a). It is obvious from the correlation outputs in Fig. 3.6 (b), 3.6 (c), 3.7 (b) and 3.7 (c) that PSPFJTC technique can produce distinct correlation peaks with the same FAF parameters used without noise whereas PFJTC failed to detect desired targets for input scene including 3 or more identical targets. Though PFJTC can detect single or two target(s) in the noisy input scene, the correlation output is corrupted by strong noise as shown in figure 8(b) for the input scene with two tanks and three trucks in Fig. 3.7 (a). To detect 3 or more identical targets in presence of noise by PFJTC the value of the parameter D should be greater than that is used for noise-free scene and still with the changed value, it produces correlation peaks of non-uniform heights for similar targets. Moreover, the correlation output is overflowed by such noise that the difference between the minimum target peak and maximum noise peak is negligible as shown in figure 3.6 (d) as an example with five similar tanks. As far we have examined, PFJTC provides best result for noisy input scene with $D=10^{-3}$ irrespective of number of targets as shown in figures 3.6 (d) and 3.7 (d) while PSPFJTC offers best result for either $D=10^{-4}$ or $D=10^{-3}$ with noise free or noisy input scene having any number of similar targets as shown in figures 3.6 (c), 3.6 (e), 3.7 (c) and 3.7 (e). So the PSPFJTC technique has better noise robustness compared to the PFJTC technique and it produces almost uniform CPI for all the targets and most importantly in case of PFJTC technique, the FAF parameters need to be readjusted for getting correlation peaks having 3 or more targets in the noisy input scene which is not a feasible condition for real time target detection process.

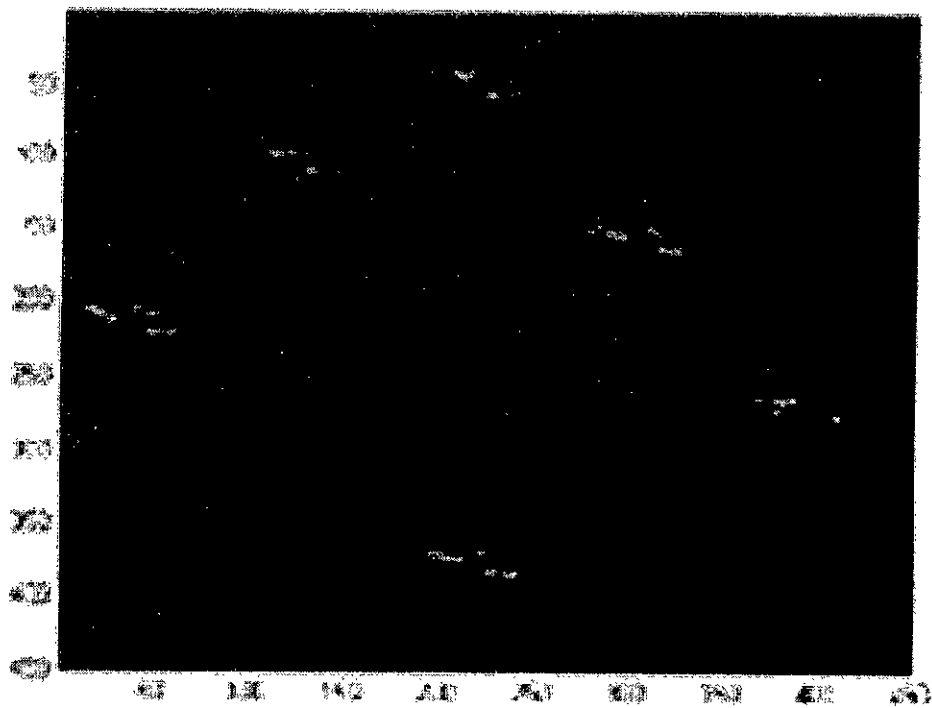


Fig. 3.6 (a): Gray level joint input scene with noise of 5 tanks as target

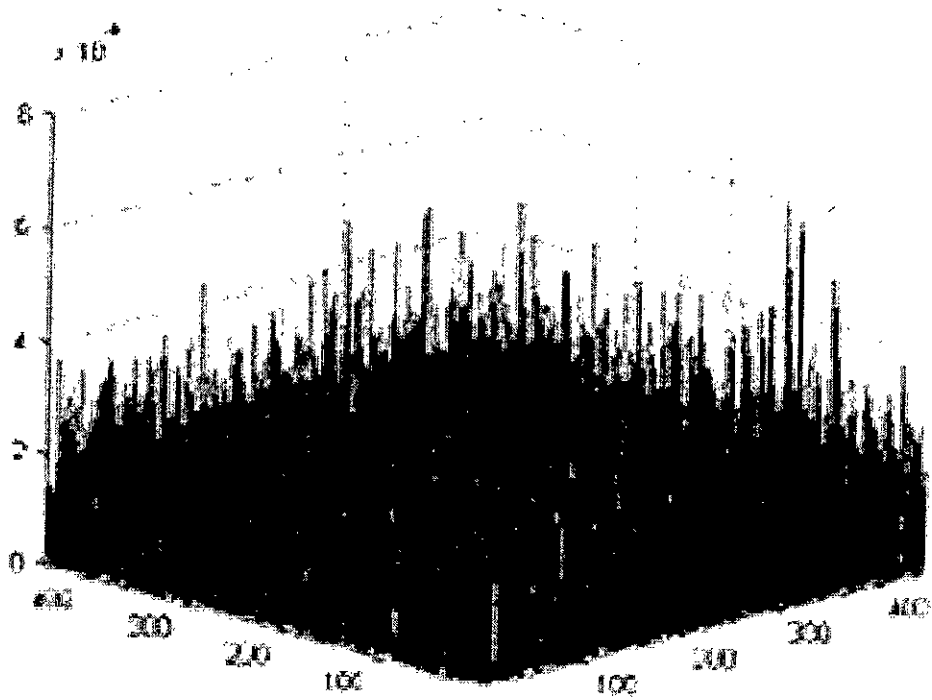


Fig. 3.6 (b): PFJTC output of fig. 3.6 (a) with $D=10^{-4}$

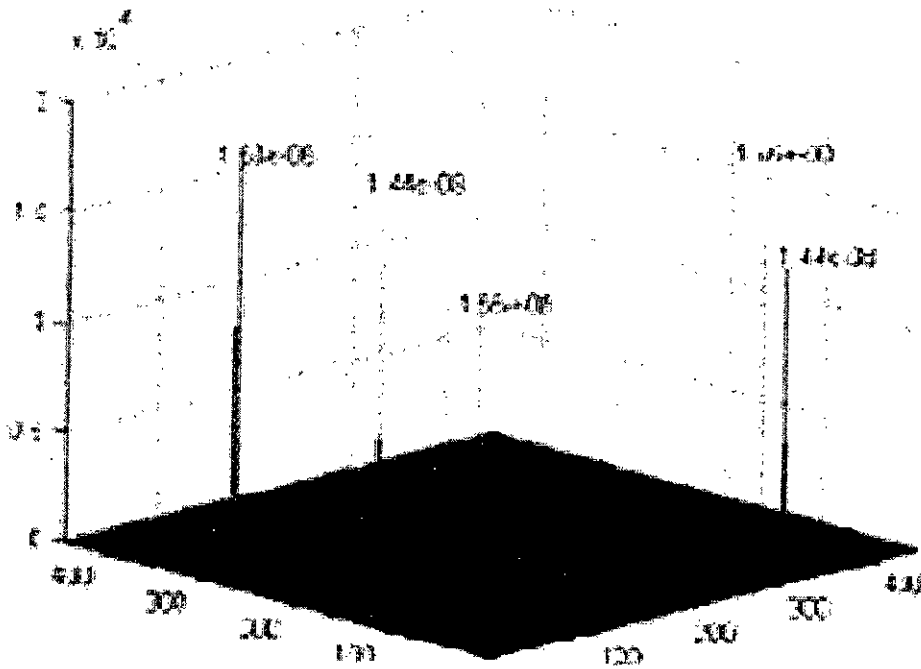


Fig. 3.6 (c): PS PFJTC output of fig. 3.6 (a) with $D=10^{-4}$

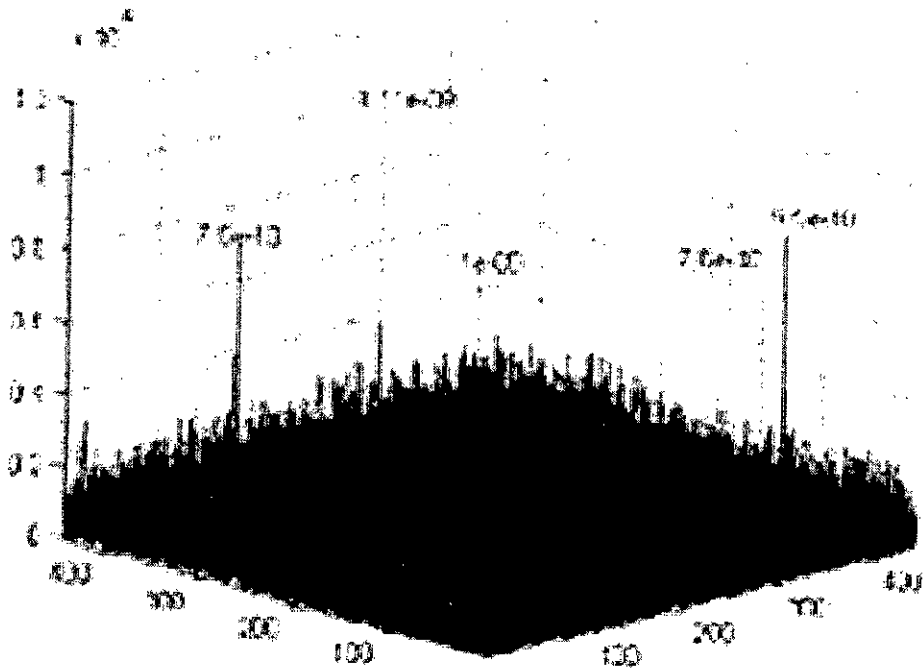


Fig. 3.6 (d): PFJTC output of fig. 3.6 (a) with $D=10^{-3}$

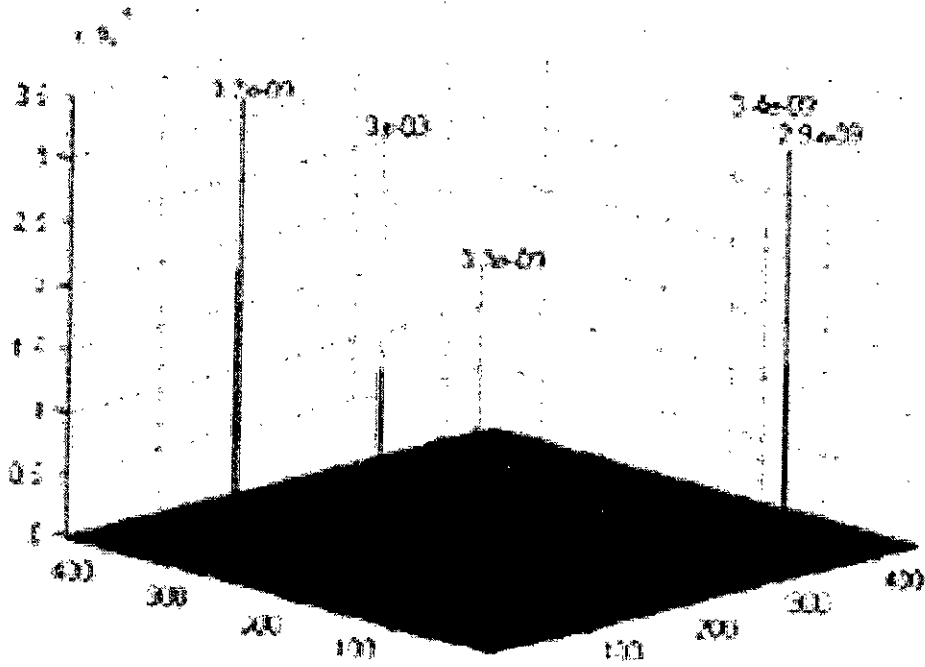


Fig. 3.6 (e): PSPFJTC output of fig. 3.6 (a) with $D=10^{-3}$

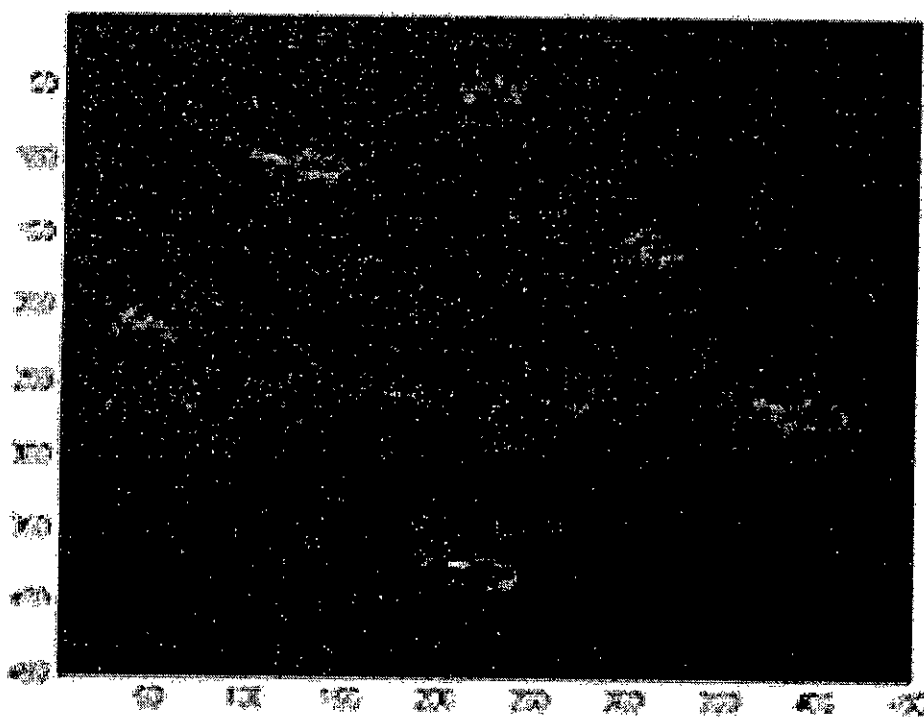


Fig. 3.7 (a): Gray level joint input scene with noise of 2 tanks as target and 3 trucks as non-target

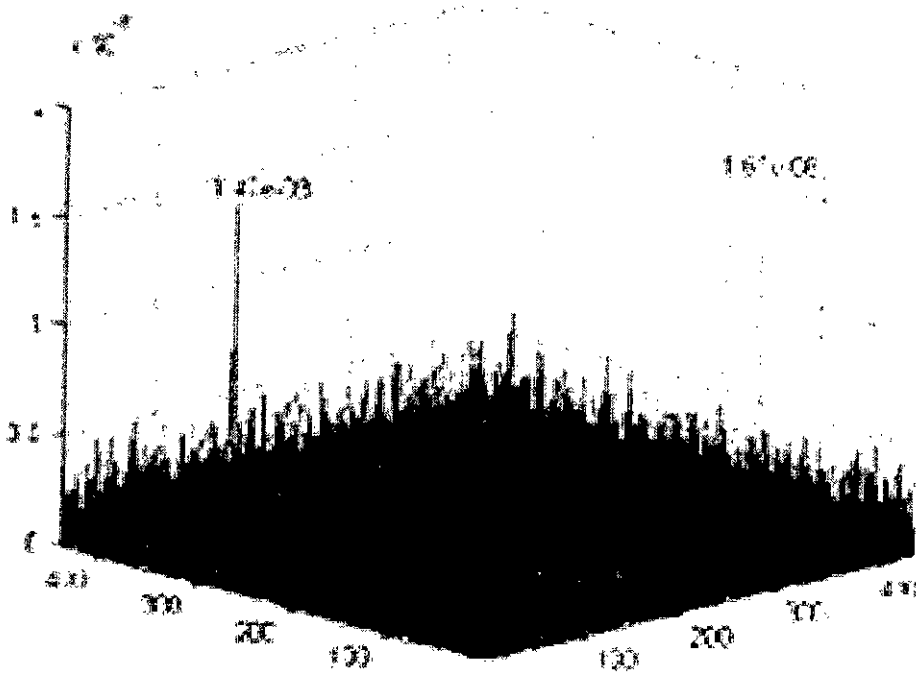


Fig. 3.7 (b): PFJTC output of fig. 3.7 (a) with $D=10^{-4}$

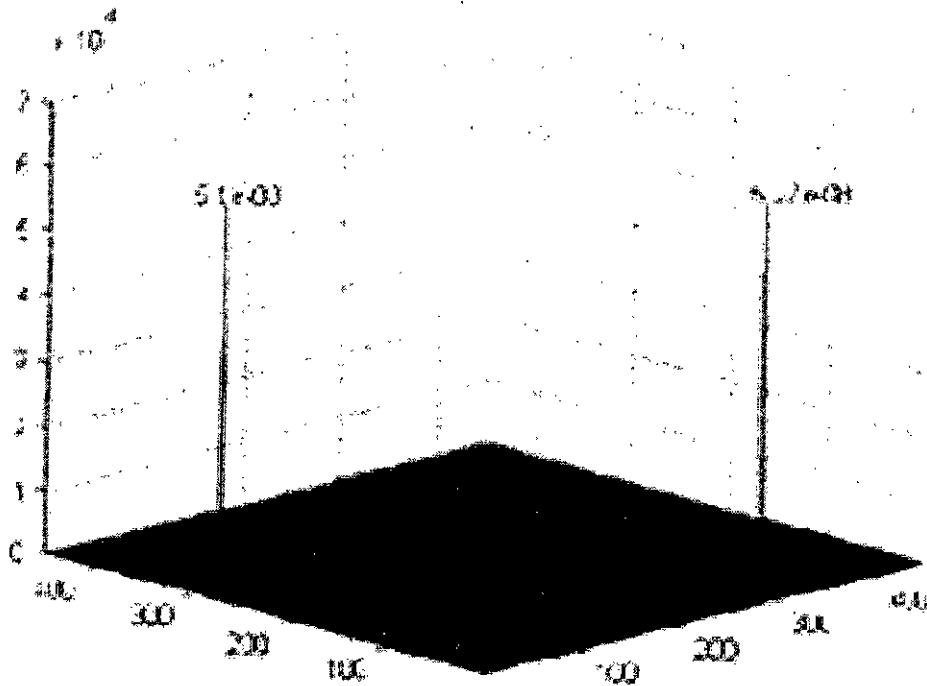


Fig. 3.7 (c): PS PFJTC output with $D=10^{-4}$

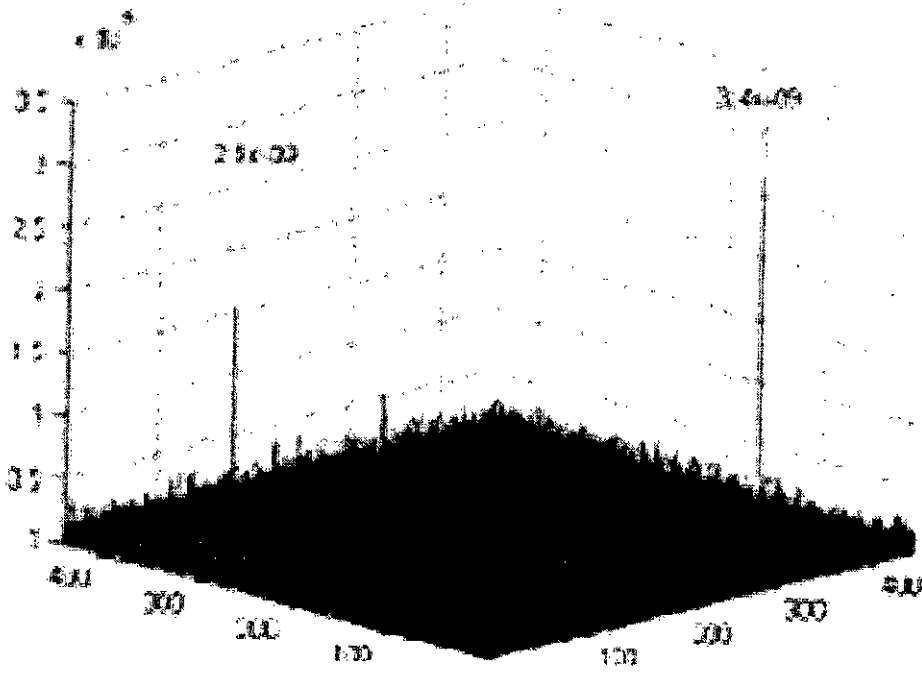


Fig. 3.7 (d): PFJTC output of fig. 3.7 (a) with $D=10^{-3}$

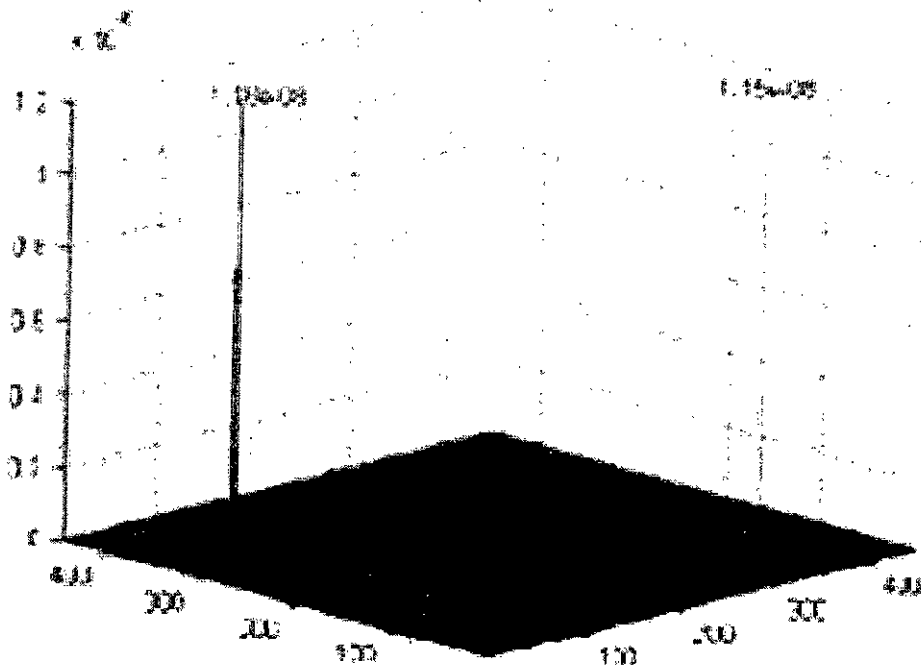


Fig. 3.7 (e): PSPFJTC output of fig. 3.7 (a) with $D=10^{-3}$

The detection performance by PFJTC and PSPFJTC are summarized in Table 3.1 and 3.2 for noise free and noisy input scene respectively. From the comparative study it is obvious that the PSPFJTC technique produces the desired uniform correlation peak irrespective of the number of objects and presence of noise in the input scene, which makes it an efficient method for multiple target detection.

Table 3.1: Performance comparison for noise free input scene

Input Scene Objects without Noise	PFJTC				PSPFJTC			
	D	Peak _{max}	Peak _{min}	P _{min} /P _{max}	D	Peak _{max}	Peak _{min}	P _{min} /P _{max}
5 tanks	10^{-4}	4.4×10^{-4}	4.17×10^{-4}	0.95	10^{-4}	3.2×10^{-3}	3.2×10^{-3}	1
4 tanks, 1 truck	10^{-4}	5.45×10^{-4}	5.36×10^{-4}	0.98	10^{-4}	4×10^{-3}	4×10^{-3}	1
3 tanks, 2 trucks	10^{-4}	7×10^{-4}	6.82×10^{-4}	0.97	10^{-4}	5.2×10^{-3}	5.2×10^{-3}	1
2 tanks, 3 trucks	10^{-4}	9.4×10^{-4}	9×10^{-4}	0.96	10^{-4}	6.4×10^{-3}	6.3×10^{-3}	0.98
1 tank, 4 trucks	10^{-4}	1.3×10^{-3}			10^{-4}	8.3×10^{-3}		

Table 3.2: Performance comparison for noisy input scene

Input Scene Objects with Noise	PFJTC				PSPFJTC			
	D	Peak _{max}	Peak _{min}	P _{min} /P _{max}	D	Peak _{max}	Peak _{min}	P _{min} /P _{max}
5 tanks	10^{-4}	Can not detect			10^{-4}	1.66×10^{-8}	1.44×10^{-8}	0.87
4 tanks, 1 truck	10^{-4}	Can not detect			10^{-4}	2.26×10^{-8}	1.83×10^{-8}	0.81
3 tanks, 2 trucks	10^{-4}	Can not detect			10^{-4}	3.43×10^{-8}	3.24×10^{-8}	0.94
2 tanks, 3 trucks	10^{-4}	1.67×10^{-8}	1.47×10^{-8}	0.88	10^{-4}	5.27×10^{-8}	5.1×10^{-8}	0.97
1 tank, 4 trucks	10^{-4}	3×10^{-8}			10^{-4}	1.04×10^{-7}		

Input Scene Objects with Noise	PFJTC				PSPFJTC			
	D	Peak _{max}	Peak _{min}	P _{min} /P _{max}	D	Peak _{max}	Peak _{min}	P _{min} /P _{max}
5 tanks	10^{-3}	1.11×10^{-9}	7.5×10^{-10}	0.68	10^{-3}	3.4×10^{-9}	2.9×10^{-9}	0.85
4 tanks, 1 truck	10^{-3}	1.84×10^{-9}	1.31×10^{-9}	0.71	10^{-3}	4.43×10^{-9}	4.24×10^{-9}	0.96
3 tanks, 2 trucks	10^{-3}	2.6×10^{-9}	1.6×10^{-9}	0.61	10^{-3}	7.26×10^{-9}	7×10^{-9}	0.96
2 tanks, 3 trucks	10^{-3}	3.4×10^{-9}	2.9×10^{-9}	0.85	10^{-3}	1.18×10^{-9}	1.15×10^{-9}	0.97
1 tank, 4 trucks	10^{-3}	5×10^{-9}			10^{-3}	2.17×10^{-8}		

3.3 Class-associative Target Detection

3.3.1 Theoretical Analysis

Class-associative target detection means detecting a group of objects, which may be similar, slightly similar or totally dissimilar. The class-associative target detection technique is formulated by combining the algorithm discussed in the previous section and an enhanced version of the generalized fringe-adjusted filter. The main advantage of the proposed scheme is that here the no of processing steps is always fixed irrespective of the number of included objects in a class.

3.3.1.1 Two Objects in a Class

Let, $r_1(x, y)$ and $r_2(x, y)$ be the two reference images in a class and $t(x, y)$ be the input scene image containing various unknown objects. First, each of the reference image in a class, is placed side-by-side using a SLM to form a composite joint reference image as expressed by

$$r(x, y) = r_1(x, y - y_1) + r_2(x, y - y_2) \quad (3.13)$$

This composite joint reference image is then phase encoded and phase shifted as expressed by Eq. (3.1) to Eq. (3.4) in the previous section and then two joint images are formed which are also expressed by Eq. (3.5) and Eq. (3.6).

After Fourier transformation of these two joint images, we get JPS as

$$\begin{aligned}
|F_1(u, v)|^2 &= \sum_{i=1}^2 |R_i(u, v) \times \Phi(u, v)|^2 + R_1(u, v)R_2^*(u, v) + R_1^*(u, v)R_2(u, v) \\
&+ \sum_{i=1}^n |T_i(u, v)|^2 + \sum_{i=1}^n R_1(u, v)T_i^*(u, v) \exp[j(vy_1 - vy_i)] \times \Phi(u, v) \\
&+ \sum_{i=1}^n R_2(u, v)T_i^*(u, v) \exp[j(vy_2 - vy_i)] \times \Phi(u, v) \\
&+ \sum_{i=1}^n R_1^*(u, v)T_i(u, v) \exp[-j(vy_1 - vy_i)] \times \Phi^*(u, v) \\
&+ \sum_{i=1}^n R_2^*(u, v)T_i(u, v) \exp[-j(vy_2 - vy_i)] \times \Phi^*(u, v) \\
&+ \sum_{i=1}^n \sum_{k=1, k \neq i}^n T_i(u, v)T_k^*(u, v) \exp[-j(vy_i - vy_k)] \\
&+ \sum_{i=1}^n \sum_{k=1, k \neq i}^n T_i^*(u, v)T_k(u, v) \exp[j(vy_i - vy_k)]
\end{aligned} \tag{3.14}$$

$$\begin{aligned}
|F_2(u, v)|^2 &= \sum_{i=1}^2 |R_i(u, v) \times \Phi(u, v)|^2 + R_1(u, v)R_2^*(u, v) + R_1^*(u, v)R_2(u, v) \\
&+ \sum_{i=1}^n |T_i(u, v)|^2 - \sum_{i=1}^n R_1(u, v)T_i^*(u, v) \exp[j(vy_1 - vy_i)] \times \Phi(u, v) \\
&- \sum_{i=1}^n R_2(u, v)T_i^*(u, v) \exp[j(vy_2 - vy_i)] \times \Phi(u, v) \\
&- \sum_{i=1}^n R_1^*(u, v)T_i(u, v) \exp[-j(vy_1 - vy_i)] \times \Phi^*(u, v) \\
&- \sum_{i=1}^n R_2^*(u, v)T_i(u, v) \exp[-j(vy_2 - vy_i)] \times \Phi^*(u, v) \\
&+ \sum_{i=1}^n \sum_{k=1, k \neq i}^n T_i(u, v)T_k^*(u, v) \exp[-j(vy_i - vy_k)] \\
&+ \sum_{i=1}^n \sum_{k=1, k \neq i}^n T_i^*(u, v)T_k(u, v) \exp[j(vy_i - vy_k)]
\end{aligned} \tag{3.15}$$

Again, calculation of these two JPS has no detrimental effect on processing speed, as in both the channels processing can be performed in parallel. Again, the steps are always fixed irrespective of the number of objects in the class. Then the phase encoded JPS in

Eq. (3.15) is subtracted from that in Eq. (3.14) and the resultant signal is again multiplied by the phase mask. Thus,

$$\begin{aligned}
P(u, v) &= \left[|F_1(u, v)|^2 - |F_2(u, v)|^2 \right] \times \Phi(u, v) \\
&= 2 \left[\sum_{i=1}^n R_1(u, v) T_i^*(u, v) \exp[j(vy_1 - vy_i)] \times \Phi(u, v) \right. \\
&\quad + \sum_{i=1}^n R_2(u, v) T_i^*(u, v) \exp[j(vy_2 - vy_i)] \times \Phi(u, v) \\
&\quad + \sum_{i=1}^n R_1^*(u, v) T_i(u, v) \exp[-j(vy_1 - vy_i)] \times \Phi^*(u, v) \\
&\quad \left. + \sum_{i=1}^n R_2^*(u, v) T_i(u, v) \exp[-j(vy_2 - vy_i)] \times \Phi^*(u, v) \right] \times \Phi(u, v) \\
&= 2 \sum_{i=1}^n R_1(u, v) T_i^*(u, v) \exp[j(vy_1 - vy_i)] \times \Phi(u, v) \times \Phi(u, v) \\
&\quad + 2 \sum_{i=1}^n R_2(u, v) T_i^*(u, v) \exp[j(vy_2 - vy_i)] \times \Phi(u, v) \times \Phi(u, v) \\
&\quad + 2 \sum_{i=1}^n R_1^*(u, v) T_i(u, v) \exp[-j(vy_1 - vy_i)] \times \Phi^*(u, v) \times \Phi(u, v) \\
&\quad + 2 \sum_{i=1}^n R_2^*(u, v) T_i(u, v) \exp[-j(vy_2 - vy_i)] \times \Phi^*(u, v) \times \Phi(u, v)
\end{aligned} \tag{3.16}$$

Due to random nature of phase mask, the first and second terms of Eq. (3.16) are scattered in various directions, and only one cross-correlation term exists between the reference images and each target object, which will contribute single correlation peak per target in the output plane.

Next we have developed an enhanced version of GFAF, which can be expressed as

$$H(u, v) = \frac{C(u, v)}{D(u, v) + \alpha |R_1(u, v)|^m + \beta |R_2(u, v)|^m} \tag{3.17}$$

where $\alpha + \beta = 1$. The ratio of α and β can be varied depending on the energy content of the power spectra of the reference images in the class to get equal correlation peak. Since the power spectra of reference images can be pre-calculated and stored, implementation of this filter does not deteriorate the processing speed in real time.

Now, the JPS as found in Eq. (3.16) is multiplied by the enhanced version of the GFAF as given by

$$\begin{aligned} P_{faf}(u, v) &= H(u, v) \times P(u, v) \\ &= \frac{C(u, v) \times P(u, v)}{D(u, v) + \alpha |R_1(u, v)|^m + \beta |R_2(u, v)|^m} \end{aligned} \quad (3.18)$$

Inverse Fourier transformation of Eq. (3.16), gives a sharp delta like correlation peak for each target object in the class.

3.3.1.2 Three Objects in a Class

If there are three objects in a class, then there will be three references. Accordingly the composite reference image can be expressed as

$$r(x, y) = r_1(x, y - y_1) + r_2(x, y - y_2) + r_3(x, y - y_3) \quad (3.19)$$

Now all the processing steps are almost same as expressed in the previous section. The only difference is that we have to formulate the GFAF as follows

$$H(u, v) = \frac{C(u, v)}{D(u, v) + \alpha |R_1(u, v)|^m + \beta |R_2(u, v)|^m + \gamma |R_3(u, v)|^m} \quad (3.20)$$

where $\alpha + \beta + \gamma = 1$ and $R_1(u, v)$, $R_2(u, v)$ and $R_3(u, v)$ are the Fourier transforms of $r_1(x, y)$, $r_2(x, y)$ and $r_3(x, y)$ respectively.

Next the phase encoded JPS, calculated using the reference image as expressed in Eq. (3.19), is then multiplied by the modified GFAF as given by

$$\begin{aligned} P_{faf}(u, v) &= H(u, v) \times P(u, v) \\ &= \frac{C(u, v) \times P(u, v)}{D(u, v) + \alpha |R_1(u, v)|^m + \beta |R_2(u, v)|^m + \gamma |R_3(u, v)|^m} \end{aligned} \quad (3.21)$$

where $P(u, v)$ is the phase encoded JPS that has been calculated using the reference image as expressed in Eq. (3.19). Inverse Fourier transformation of Eq. (3.21) produces a single correlation peak for each target in the class.

3.3.1.3 Multiple Objects in a Class

For multiple objects in a class, we have just to modify the GFAP. All other processing steps are always fixed. Again, since the GFAP fabrication can be performed and stored before the real time operation, this modification has no detrimental effect on system processing speed. In summary, the no of processing steps will be always fixed for two or more objects in a class and this is the major advantage of this proposed scheme.

3.3.2 Simulation Results

Let us consider that English alphabets 'H' and 'F' are two objects forming a group or class to be detected in a binary input scene. Then these are the two reference objects forming the group or class. First let us assume that the input scene contains only these two objects. Fig. 3.8 (a) shows the joint image where the left half portion indicates the reference images and the right half portion indicates the input scene images. Now class-associative target detection algorithm is applied as described in section 3.3.1 and the corresponding PSPFJTC output is obtained as shown in Fig. 3.8 (b). In PSPFJTC technique, the FAF parameters are set as $C=1$, $D=1e-9$ and $m=1$. Thus, here we get the phase encoded phase only correlation output of the joint image. Initially, the value of power spectra parameters are $\alpha=0.5$ and $\beta=0.5$. Fig. 3.8 (c) shows the side view of the correlation output. Here the values of unequal correlation peaks are 159.6665 and 148.5638. This difference is due to the difference in energy contents of the two objects. So by adjusting the power spectra parameters as $\alpha=0.45$ and $\beta=0.55$ the correlation output for both the targets of the class can be made equal as shown in Fig. 3.8 (d) and (e). In this case the peak intensities are 153.9597 and 153.3183 respectively.

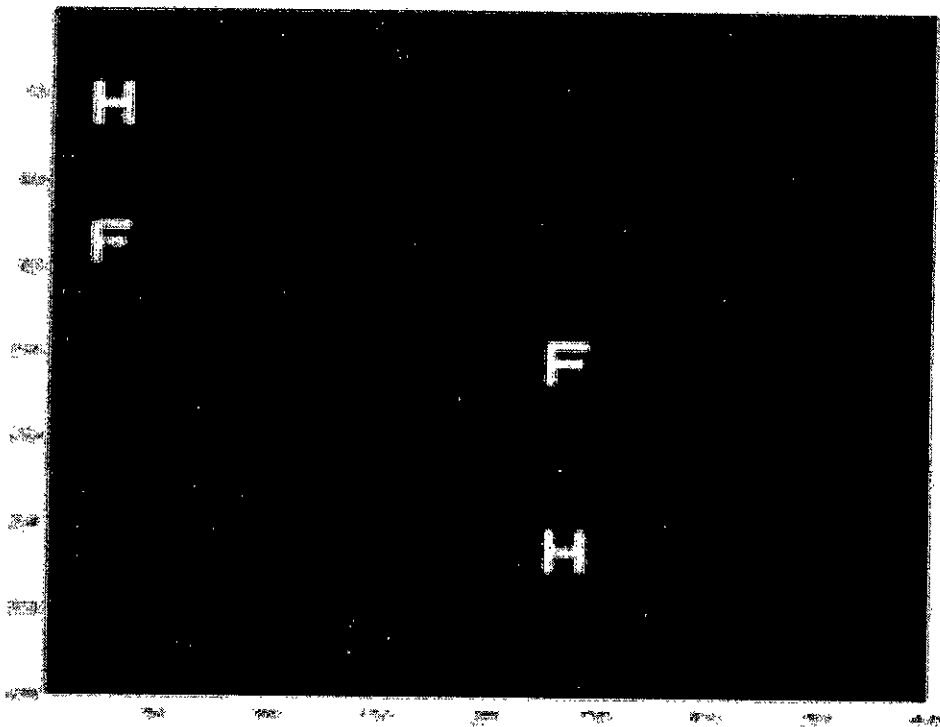


Fig. 3.8 (a): Binary joint image with two objects in the class and in the input scene

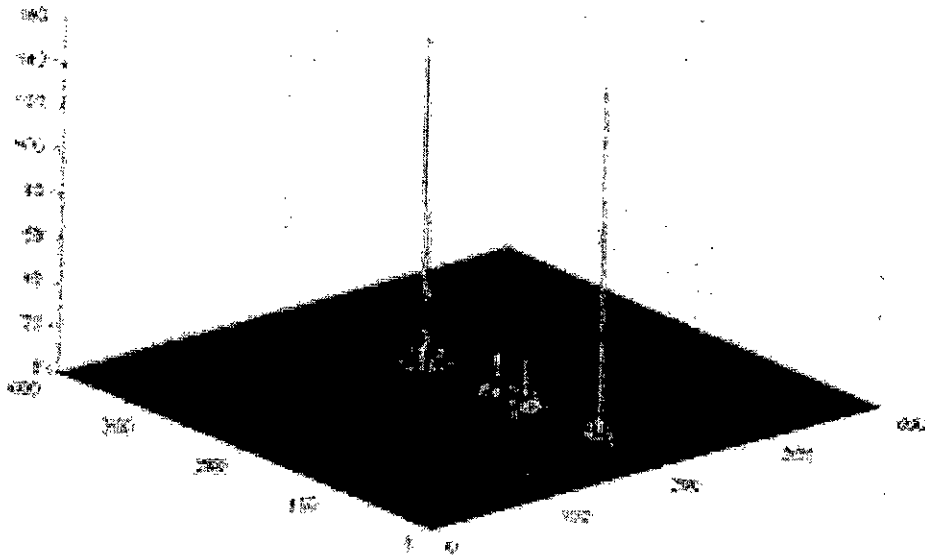


Fig. 3.8 (b): Correlation output of fig. 3.8 (a) with equal values of power spectra parameter

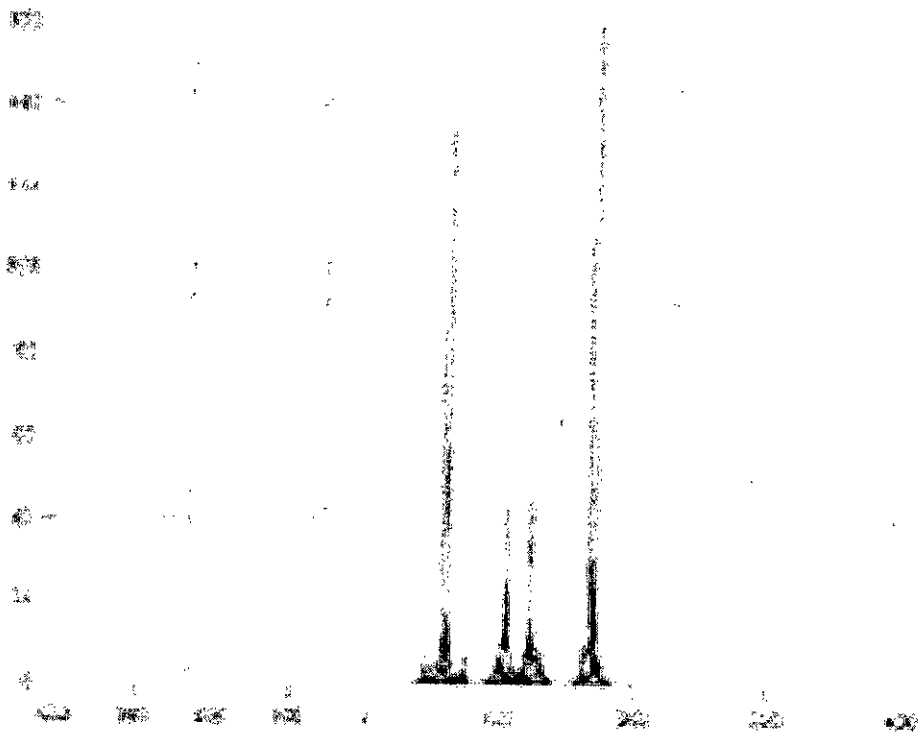


Fig. 3.8 (c): Side view of Fig. 3.8 (b)

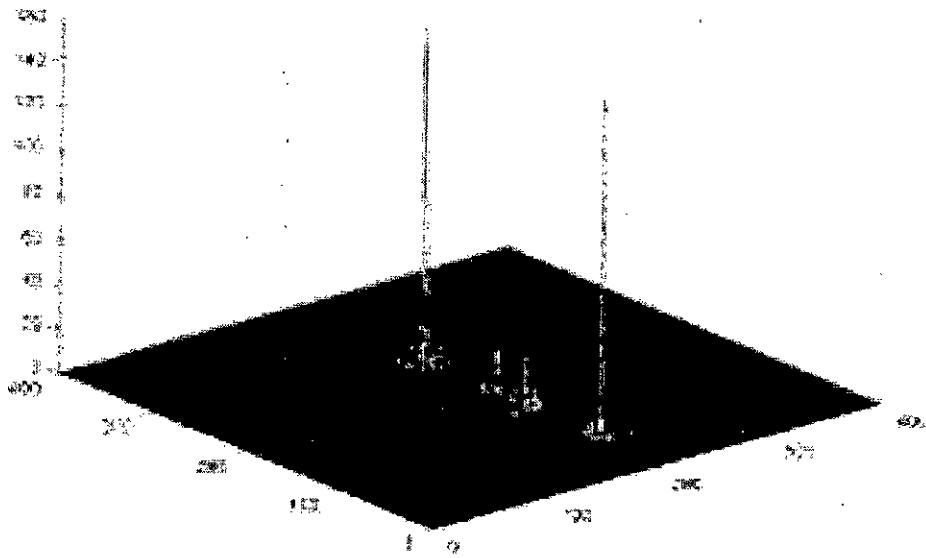


Fig. 3.8 (d): Correlation output of Fig. 3.8 (a) with unequal values of power spectra parameter

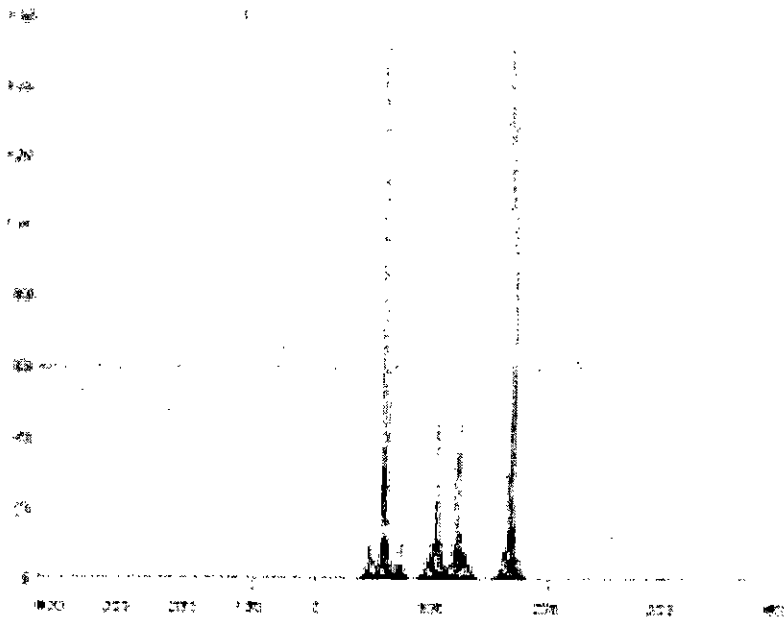


Fig. 3.8 (e): Side view of Fig. 3.8 (d)

Next we have considered the input scene with multiple objects as shown in Fig. 3.9 (a). In this case the input scene contains two target objects representing the member of the class and four non-target objects. The correlation output with phase-only filter and power spectra parameters adjusted is given in Fig. 3.9 (b). In this case the values of power spectra parameters (α and β) are chosen as has been taken in the previous case. Fig. 3.9 (c) shows the side view of Fig. 3.9 (b).

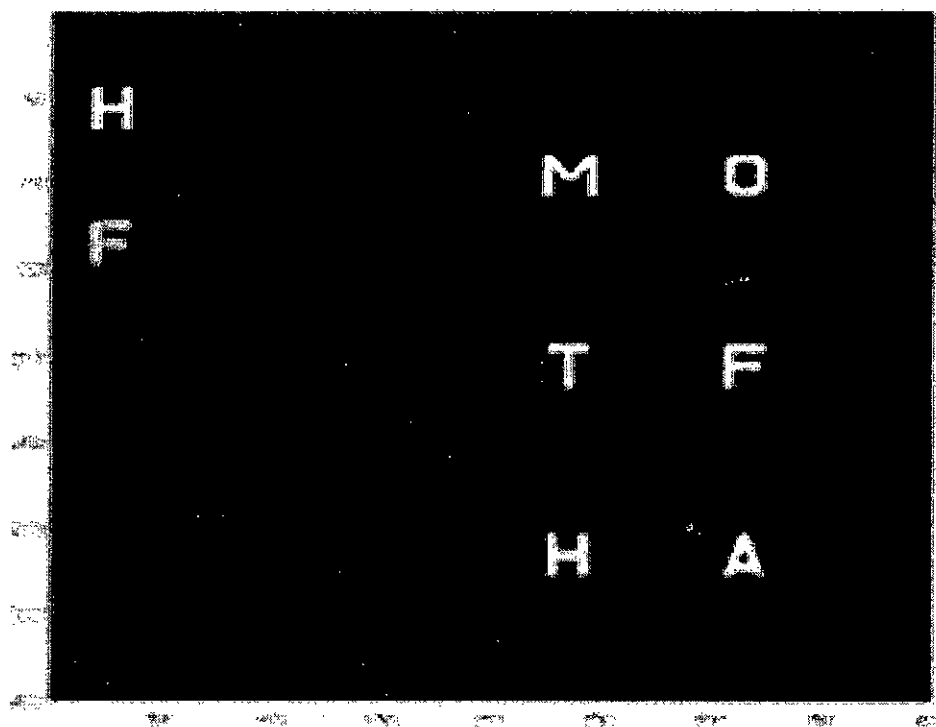


Fig. 3.9 (a): Binary joint image with 2 objects in the class and multiple objects in the input scene

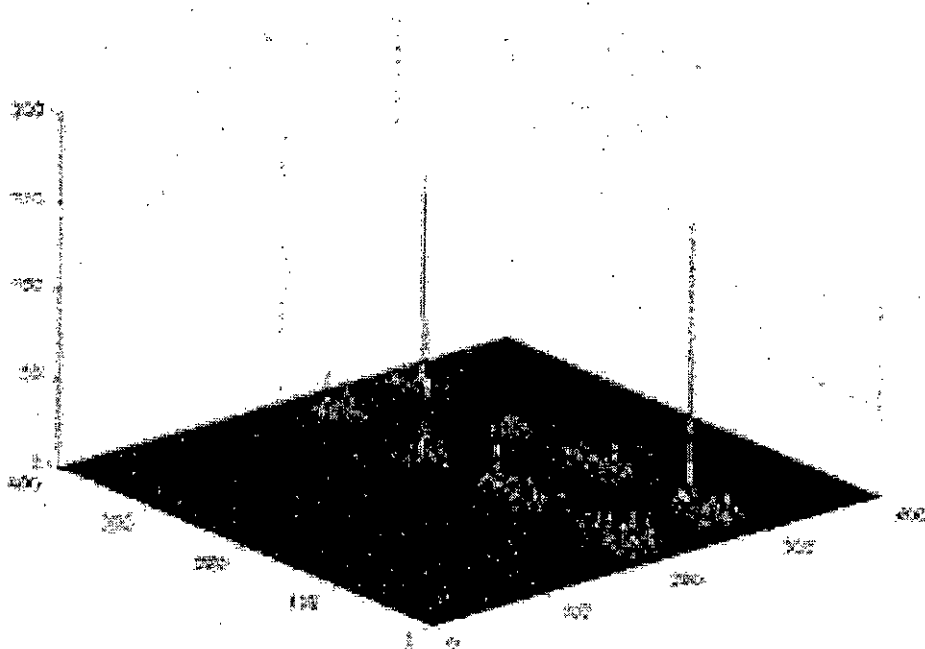


Fig. 3.9 (b): Correlation output of Fig. 3.9 (a)

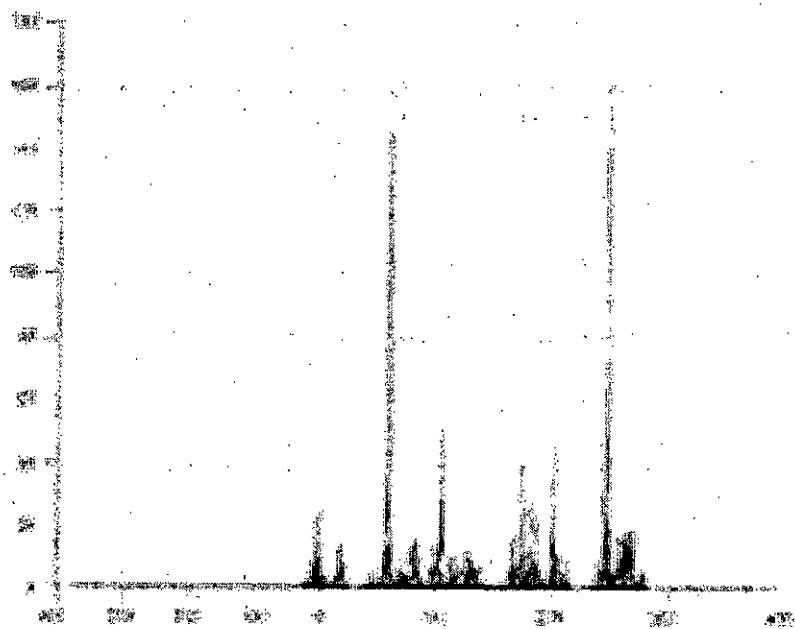


Fig. 3.9 (c): Side view of Fig. 3.9 (b)

The values of correlation peaks are 160.8953, 158.9873, 52.2917, 46.1420, 40.5454 and 29.5725 where the first two values represent the peaks for two target objects. Here it is observed that the peak intensities for the targets are almost equal and non-target peaks are negligible compared to target peaks. Attempts could be made by adjusting the values of α and β to make the peak intensities much nearer. But it is insignificant and not necessary as with the present values of α and β the ratio of maximum target peak to minimum target peak is only 1.0120.

So, once the targets of a class are defined or chosen, the required values of power spectra parameters can be selected by making test runs with only the desired objects of the class in the input scene. For these however several trial and error attempts are needed. The determined values can then be used for a real input scene with multiple targets and non-targets with no or negligible degradation of performance.

Now let us take three objects as the member of the target class. So there will be three reference objects in the class or group. The joint image is shown in Fig. 3.10 (a) where the input scene contains only the three target images. The class-associative PSPFJTC output is shown in Fig. 3.10 (b) with equal values of power spectra parameters (i.e. $\alpha=\beta=\gamma=0.33$).

Fig. 3.10 (c) shows the side view of Fig. 3.10 (b) and here the target peak intensities are 173.0164, 162.7981 and 145.6842. By adjusting the power spectra parameters i.e. $\alpha=0.328$ $\beta=0.467$ and $\gamma=0.205$, almost equal correlation peaks are obtained as shown in Fig. 3.10 (d) and its side view in Fig. 3.10 (e). In this case the correlation peak intensities become 158.7356, 158.0923 and 158.0336.

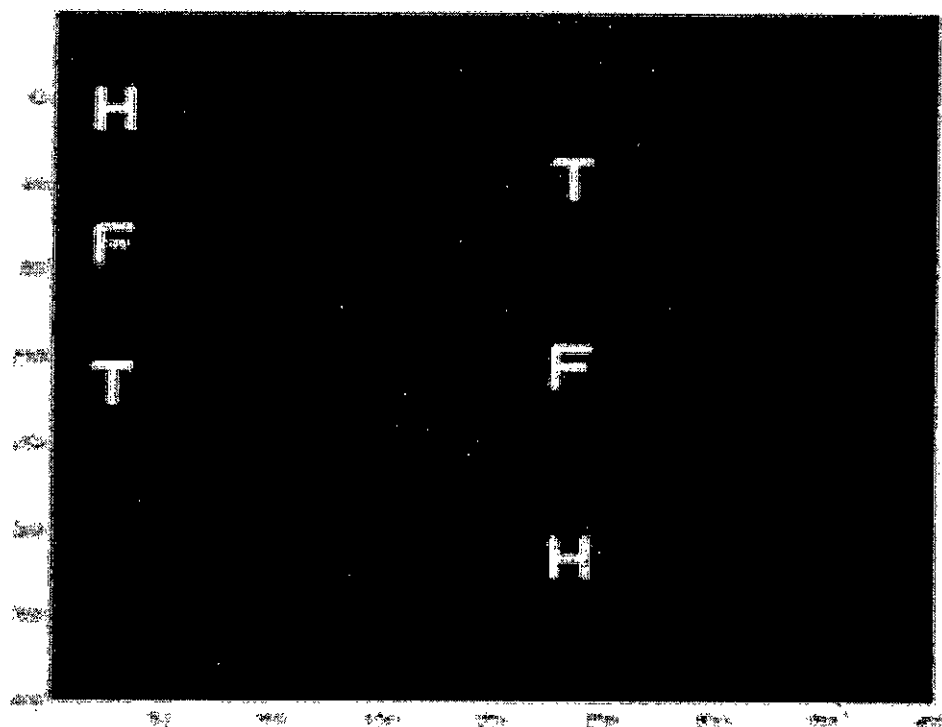


Fig. 3.10 (a): Binary joint image with three objects in the class and in the input scene

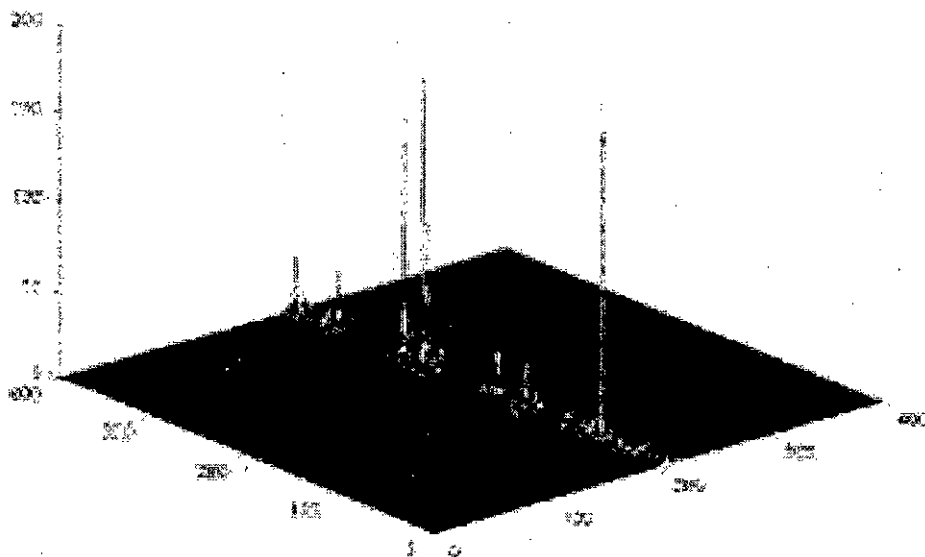


Fig. 3.10 (b): Correlation output of Fig. 3.10 (a) with equal values of α , β and γ

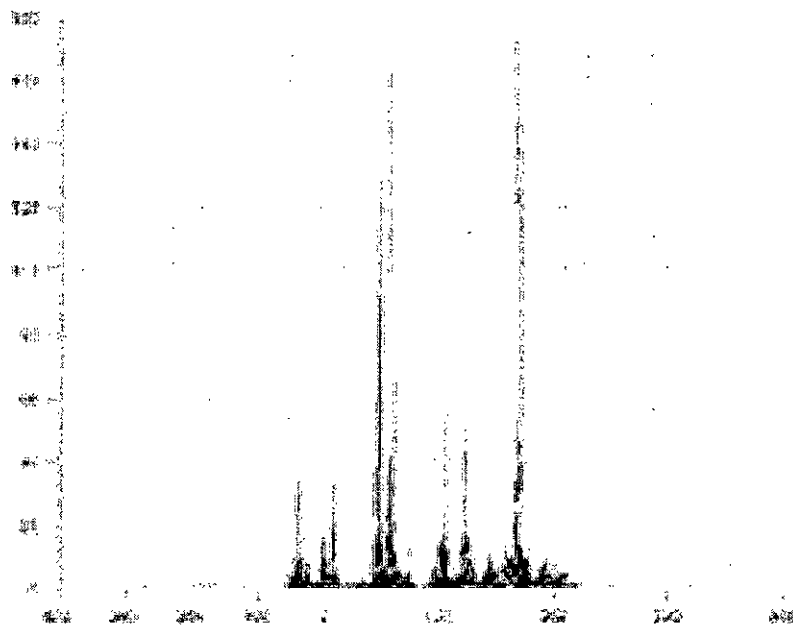


Fig. 3.10 (c): Side view of Fig. 3.10 (b)

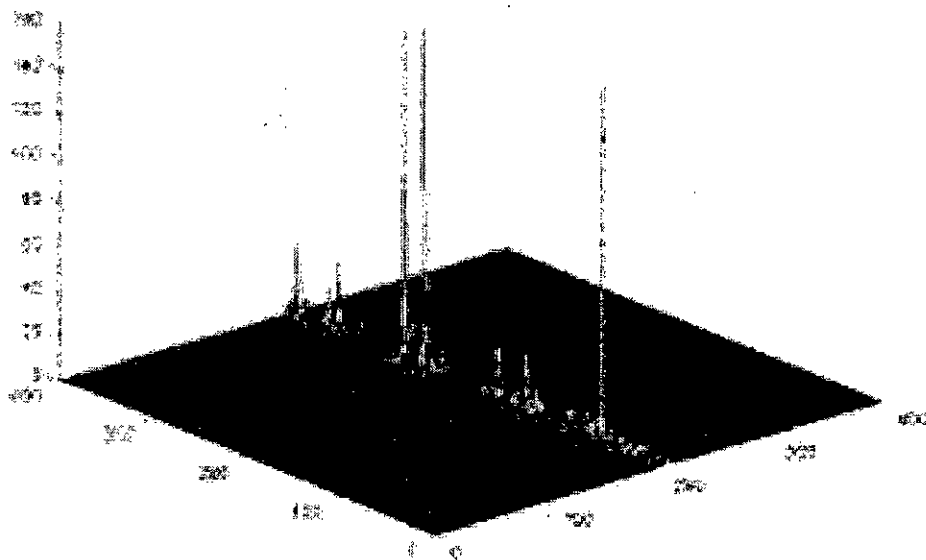


Fig. 3.10 (d): Correlation output with unequal α , β and γ

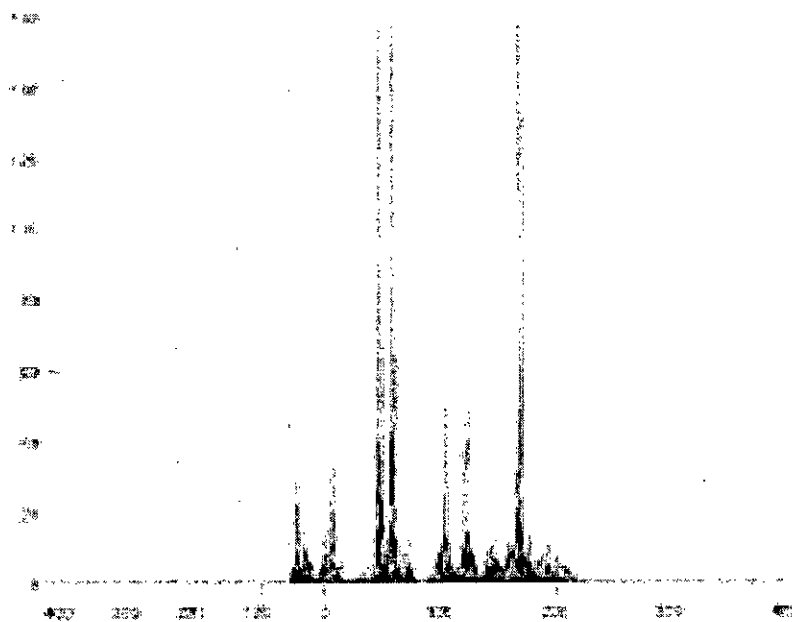


Fig. 3.10 (e): Side view of Fig. 3.10 (d)

Next we have placed these three objects in a multi object input scene and correlation output is obtained with the same values of power spectra parameters as are used in the correlation output shown in Fig. 3.10 (a) and (b). First of all, the joint image is shown in Fig. 3.11 (a) and its correlation output in Fig. 3.11 (b) and (c) where Fig. 3.11 (c) is a side view of Fig. 3.11 (b). In this case the input scene contains four target objects; therefore we should get four distinct and almost equal peaks. From Fig. 3.11 (b) and (c), four distinct target peaks are evident but they are not all equal. This is due to the random nature of phase mask used in PSPFJTC technique. The peak intensities of correlation output are 172.6194, 159.2031, 156.9767, 138.5146, 53.3552 and 51.7385; where the first four values represent the peak intensity of four target objects.

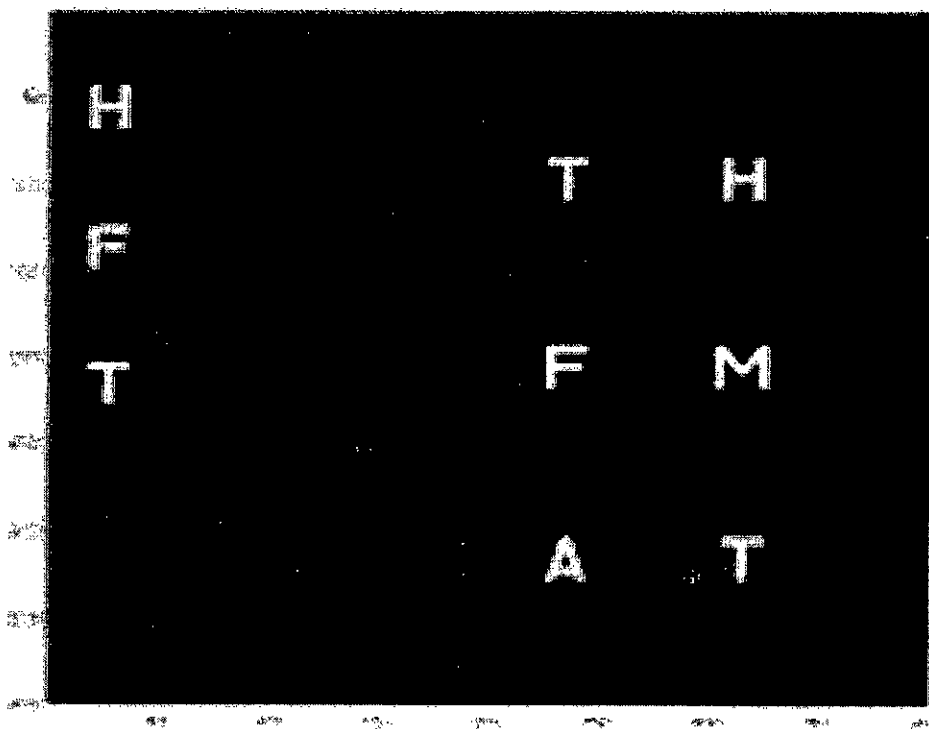


Fig. 3.11 (a): Binary joint image with 3 objects in the class and multiple objects in the input scene

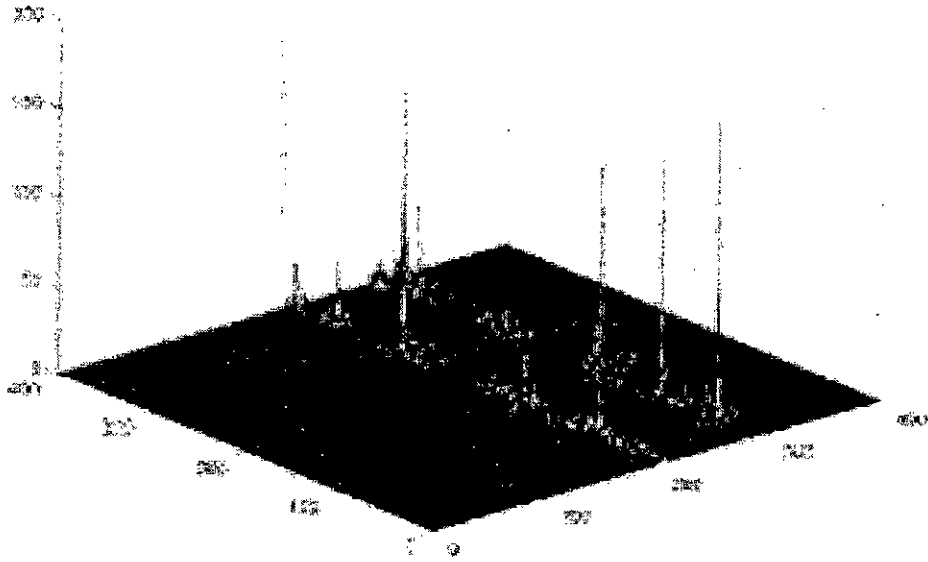


Fig. 3.11 (b): Correlation output of Fig. 3.11(a)

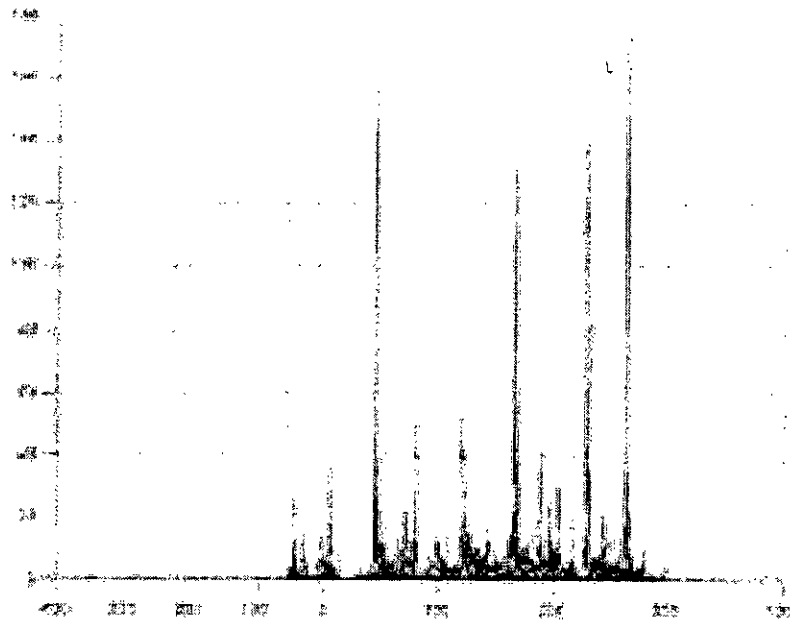


Fig. 3.11 (c): Side view of Fig. 3.11 (b)

The ratio of maximum target peak to minimum target peak is 1.2462 and the ratio of minimum target peak to maximum non-target peak is 2.5961. Since the later one is 2.0832 times higher than that of the former one, this deviation of peak intensities is quite acceptable.

3.4 Conclusion

In this chapter a novel technique named PSPFJTC is proposed for improved pattern recognition that employs phase shifting and phase encoding methods along with fringe adjusted filter for optimization of the correlation performance. The proposed PSPFJTC technique yields single correlation peak per target and thus ensures the best utilization of input and output plane SBP. This technique overcomes the limitations of currently available PFJTC technique, which is found to be very much sensitive to the number of non-targets and the presence of noise in the input scene. The proposed PSPFJTC technique can detect the targets from an input scene at any condition of non-target and also in presence of noise, with the same system parameters. Computer simulation confirms the effectiveness of the technique in multiple target detection. Next we have proposed an efficient high-speed class-associative target detection algorithm for the detection of a class of objects. The proposed algorithm utilizes multi-reference-based PSPFJTC technique along with an enhanced version of generalized fringe-adjusted filter. In this technique, the number of channels and number of processing steps are always fixed irrespective of number of objects in the class whereas in the previous methods the number of processing steps and number of channels are always increased for the same situation. Thus the proposed method ensures high-speed and low cost compared to other methods. Simulation results prove the validity of the proposed method.

Chapter 4

PROJECTION SLICE SYNTHETIC DISCRIMINANT FUNCTION-BASED JOINT TRANSFORM CORRELATION

4.1 Introduction

All JTC techniques provide shift invariance property in the input scene while suffer from high sensitivity to distortions such as scale variations and rotation. A classical JTC has a rotation invariance of ± 1 degree. To overcome this problem, several researchers have incorporated the synthetic discriminant function (SDF) formulation with JTC techniques. Distortion invariant pattern recognition using fractional power fringe adjusted joint transform correlation technique has been proposed which can successfully detect multiple distorted targets [61]. Recently, some researchers have introduced projection slice theorem with SDF formulation to achieve distortion invariance [62]. According to projection slice theorem, a set of projections is taken from each of the training images and then each projection is converted to a 1D dimensional Fourier slice. Finally, combining all of these Fourier slices forms a composite image. This composite image then can be used for detection of distorted objects. Now most of the JTC techniques produce a pair of correlation peaks for each target. Again for multiple target detection, the discrimination ratio between target and non-target objects is not so high. Here, we have introduced the projection slice synthetic discriminant function (PSDF) formulation with FAF. A phase encoded and phase shifted version of the reference image is used to produce a single correlation peak per target. Finally, a post processing technique has been adopted to increase the discrimination between similar target and non-target objects. Computer simulations show that the proposed technique can realize not only distortion invariance but also produces a sharp single correlation peak with high discrimination ratio.

4.2 Synthetic Discriminant Function

For distortion invariant pattern recognition, the reference image $r(x, y)$ is synthesized from a set of training images. Let us assume that the N training images, $r_1(x, y)$, $r_2(x, y)$, $r_n(x, y)$, containing the desired distortion invariant features, are used to construct the spatial SDF $r_{sdf}(x, y)$. Typically, the $r_{sdf}(x, y)$ is a linear combination of the training images as given by

$$r_{sdf}(x, y) = \sum_{i=1}^n a_i \times r_i(x, y) \quad (4.1)$$

where a_i represents the weights or coefficients used. This composite image is used as reference image for all types of JTC for distortion invariant recognition. In FJTC technique the GFAF is now modified as

$$H_{gfaf}^{sdf}(u, v) = \frac{C(u, v)}{D(u, v) + |R_{sdf}(u, v)|^m} \quad (4.2)$$

where $R_{sdf}(u, v)$ is the Fourier transformation of the spatial SDF $r_{sdf}(x, y)$. The SDF-based JPS is obtained by multiplying the JPS with the filter function of Eq. (4.2). This modified JPS is inverse Fourier transformed to obtain the correlation output.

4.3 Projection Slice SDF

According to projection slice theorem, the 1D Fourier transformation of the projection of an image is equivalent to a 1D Fourier slice of that image taken on the line specified.

The 1D projection of a function $f(x, y)$ along a line l and at an angle ϕ , is given by

$$g(\phi, s) = \int_l f(x, y) dl \quad (4.1)$$

All points on this line satisfy the equation

$$x \sin \phi - y \cos \phi = s \quad (4.2)$$

where s is the distance from the origin. Now the projection function can be written as

$$g(\phi, s) = \iint f(x, y) \delta(x \sin \phi - y \cos \phi - s) dx dy \quad (4.3)$$

The 1D Fourier transformation of Eq. (4.3) is given by

$$G(\phi, \omega) = \int e^{-j\omega s} g(\phi, s) ds \quad (4.4)$$

Eq. (4.4) represents the 1D Fourier slice of the 2D Fourier transform of the function $f(x, y)$. Now combining all the Fourier slices, we can get back the original function again. Substituting the value of $g(\phi, s)$ in Eq. (4.4) and using the shifting property of Dirac delta function we get

$$G(\phi, \omega) = \iint f(x, y) e^{-j\omega(x \sin \phi - y \cos \phi)} dx dy \quad (4.5)$$

If we make change of variables from rectangular to polar coordinate, then from Eq. (4.5) $f(x, y)$ can be written as

$$f(x, y) = \frac{1}{4\pi^2} \iint G(\phi, \omega) e^{j\omega(x \sin \phi - y \cos \phi)} |\omega| d\omega d\phi \quad (4.6)$$

where $|\omega|$ is the determinant of the Jacobian of the of the change of the variable from rectangular to polar coordinate.

Now if we take M equispaced projections of each of the N training images, the separation in angle of the m th slice of n th training image is expressed as

$$\phi_{mn} = \frac{m\pi}{M} + \frac{n\pi}{MN} \quad (4.7)$$

$n = 0, 1, 2, \dots, N-1$
 $m = 0, 1, 2, \dots, M-1$

The generalized projection slice synthetic discriminant function filter transfer function can be defined by

$$R(u, v) = \sum_{n=0}^N \sum_{m=0}^M a_n G_n(\omega \cos \phi, \omega \sin \phi) \delta(\phi - \phi_{mn}) |\omega| \quad (4.8)$$

where G_n is the Fourier transform of the n th training image and a_n is a normalization factor similar to that used in SDF filter design. The relation between rectangular (u, v) and polar (ω, ϕ) coordinate can be given

$$\begin{aligned} \omega &= \sqrt{u^2 + v^2}, & \phi &= \tan^{-1}\left(\frac{v}{u}\right) \\ u &= \omega \cos(\phi), & v &= \omega \sin(\phi) \end{aligned} \quad (4.9)$$

Inverse Fourier transformation of Eq. (4.8) produces a composite image. If we use this composite image as a reference image then it is possible to obtain the distortion invariant pattern recognition.

4.4 Calculation of Weighting Factor

To calculate the weighting factor a_k , $k=1,2, \dots, N$, for N training images r_k , $k=1,2, \dots, N$ respectively, we have employed a simple iterative method. First of all, all the weight factors are taken as unity. Then a fixed number of slices are taken from each of the training image and finally all the slices are placed according to the angle difference as expressed in Eq. (4.7). Finally, 2D inverse Fourier transform of all these slices give the composite image. This composite image is now used as the reference image to find the correlation with all the other training images.

First of all, we find the correlation output of the composite reference image generated previously with all the training images. In the first iteration of finding the correlation with the training set, unequal correlation peak intensities are obtained due to dissimilarity in the training images of training set. Since our main aim is to make equal correlation peaks for all the training images, the weighting factor associated with each training image is updated according the following formula

$$a_k^{i+1} = a_k^i + \eta \times (p_{\max}^i - p_k^i) \quad (4.10)$$

where a_k is the weight factor for the k^{th} image of the training image, i is the iteration number, p_{\max} is the maximum correlation heights obtained from all the training images in a particular iteration number and p_k is the correlation peak for the k^{th} image. η is a relaxation factor and the value of which has the direct effect in changing the values of weighting factor from previous iteration to the next iteration. With the new values of

weight factor, the slices of the composite reference image are updated again and the modified reference image is used for the next iteration. This iterative process is continued until the correlation peaks for all the training images are equal or the difference between the maximum and minimum peak reaches a predetermined value. This limiting value of error is expressed as

$$\xi^i = \frac{p'_{\max} - p'_{\min}}{p'_{\max}} \quad (4.11)$$

where ξ^i is the error factor obtained in the i^{th} iteration, p'_{\max} is the maximum correlation peak and p'_{\min} is the minimum correlation peak obtained in the i^{th} iteration. The composite image thus obtained is the desired projection-slice SDF (PSDF) based composite reference image.

4.5 Simulation Results

To investigate the performance of the PSDF based technique for composite image generation, we first test the projection-slice theorem for image projections and reconstruction. For simulation purpose we use an undistorted binary character 'E' of size 32×32 pixels. Then a number of projections are taken from this image and finally the image is reconstructed using the projections of the image. For image projections and reconstruction, we have used the radon and iradon routine of MATLAB software. Fig. 4.1(a) shows the binary character 'E'. Fig. 4.1 (b), (d) and (f) show the plot for different number of projections taken from the original image whereas Fig. 4.1 (c), (e) and (g) show the corresponding reconstructed image.

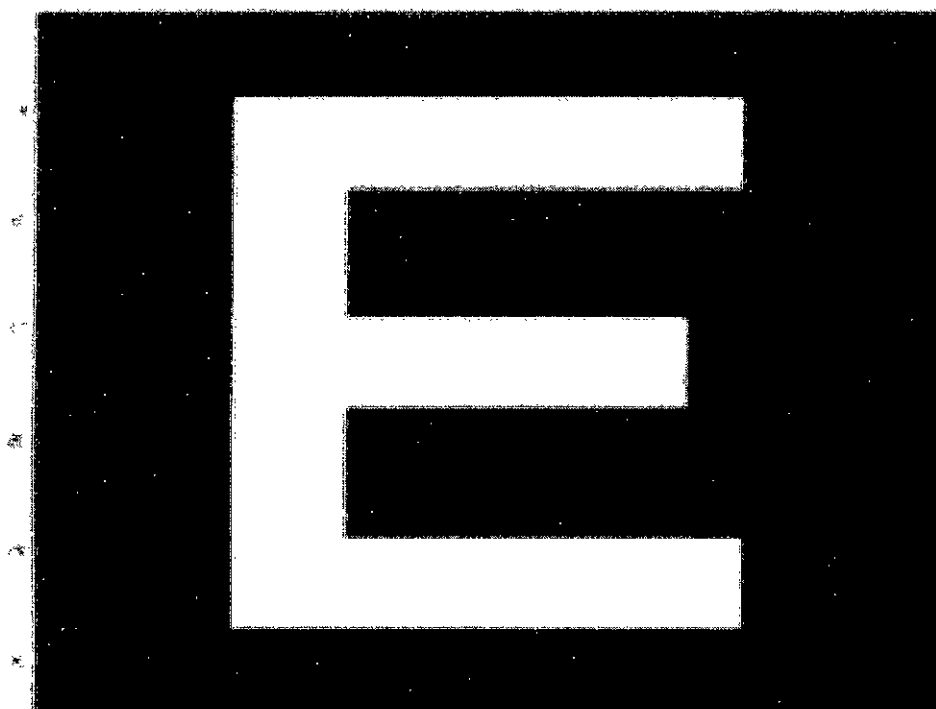


Fig 4.1 (a): Original binary image of 'E' character

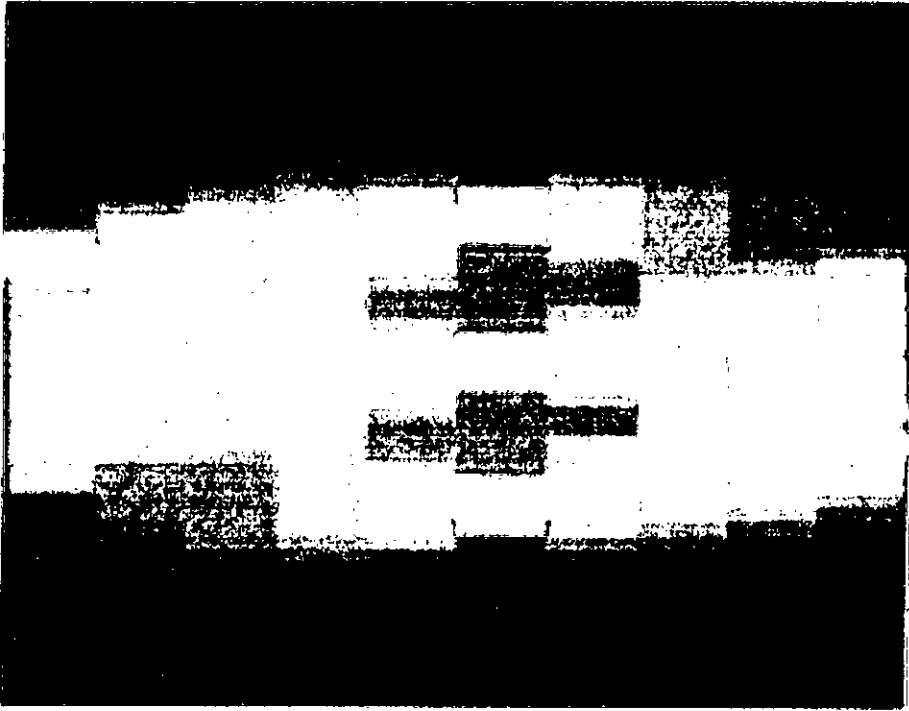


Fig 4.1 (b): 10 projections of fig. 4.1 (a)

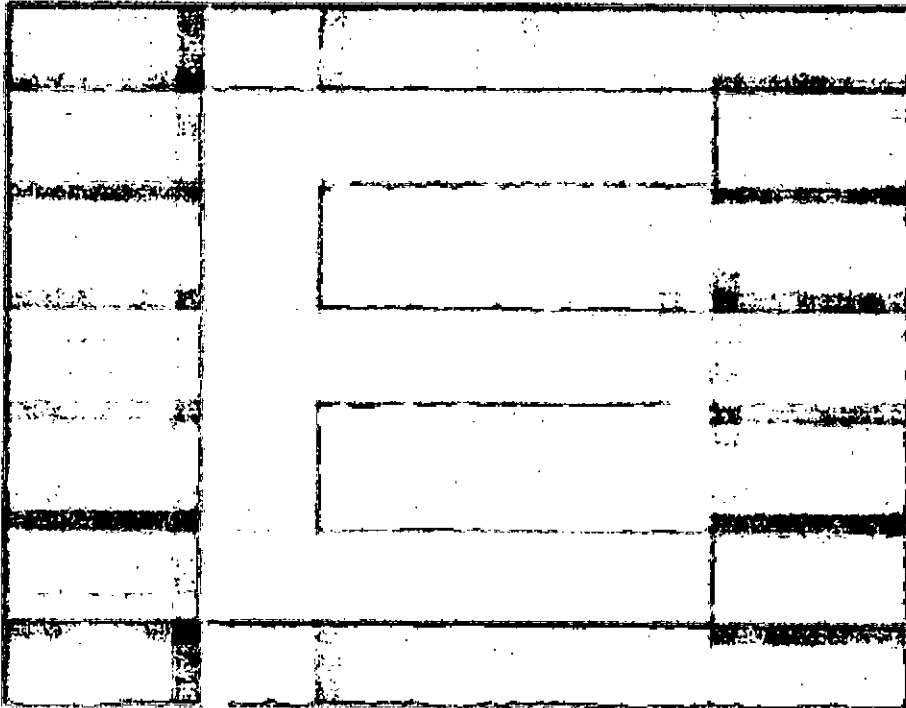


Fig 4.1 (c): Reconstructed image of fig. 4.1 (a)



Fig 4.1 (d): 60 projections of fig. 4.1 (a)

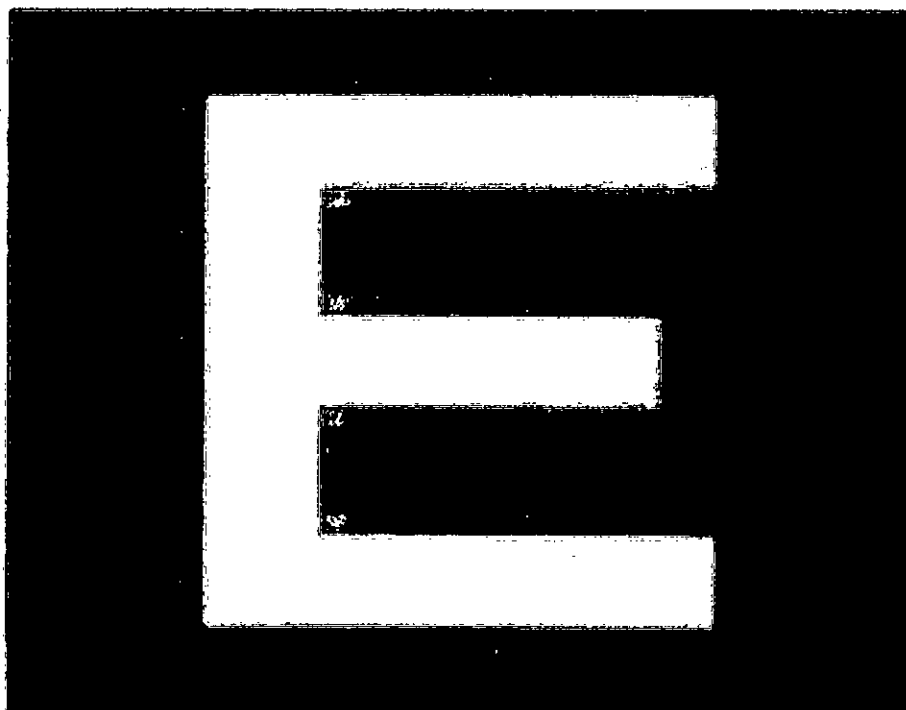


Fig 4.1 (e): Reconstructed image of fig. 4.1 (a)

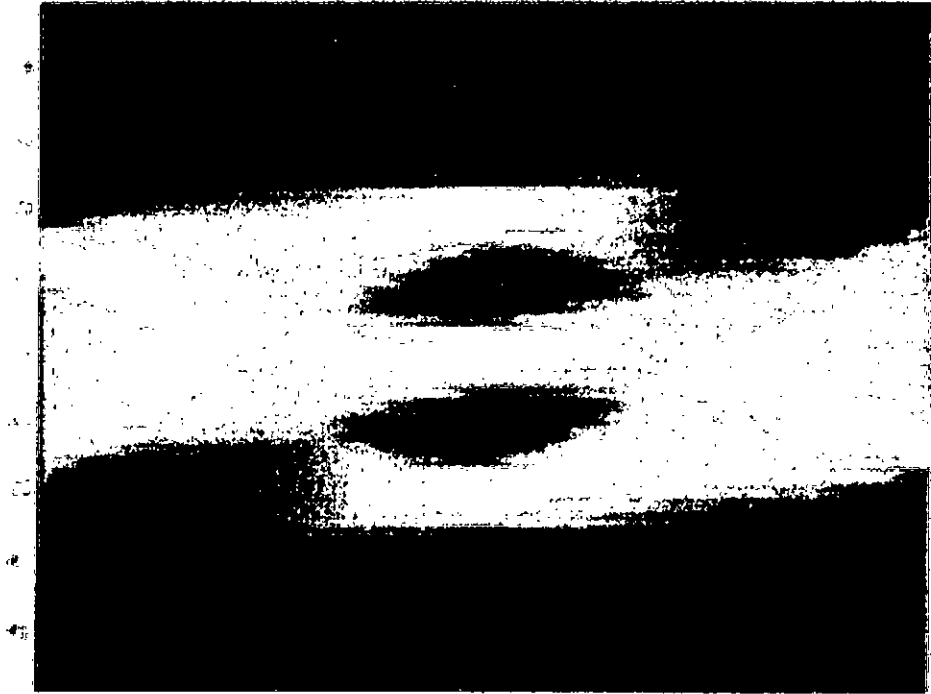


Fig 4.1 (f): 180 projections of fig. 4.1 (a)

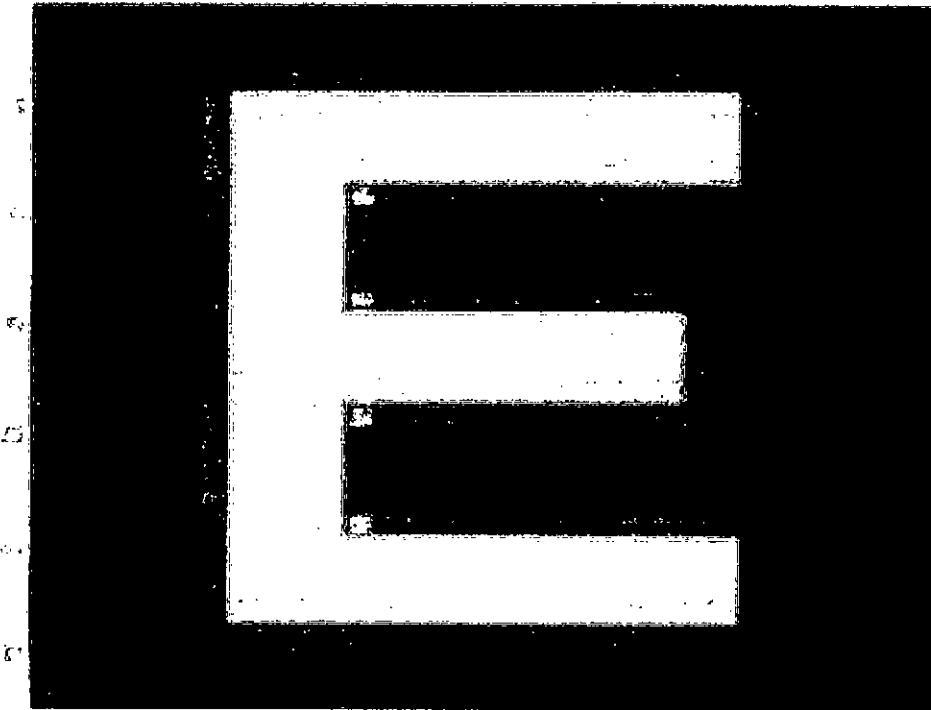


Fig 4.1 (g): Reconstructed image of fig. 4.1 (a)

From these figures, it is evident that the number of projections taken is an important factor to reconstruct the original image. Higher the number of projections, the smoother is the reconstructed image. Theoretically, an infinite no of projections is required to reconstruct the actual original image. Therefore, all the reconstructed images using projection-slice theorem contain some amount of distortion and this is an inherent property of this method. In all the practical cases, a certain amount of distortion is allowed and for this reason, the number of projections must not be less than a predefined quantity.

4.5.1 Composite Image Formation using Two Training Images

Here we form a composite image using PSDF technique using only two training images. Fig. 4.2(a) shows the two training images. Between the two training images, one is the original image and the second one is the 90^0 in-plane rotated version of the original one.

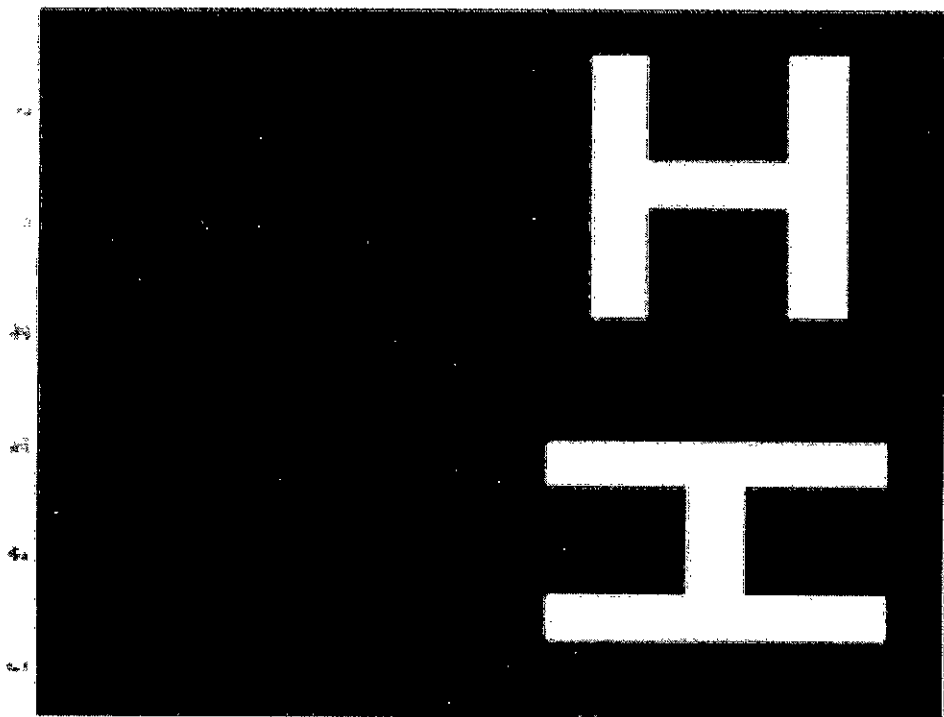


Fig 4.2(a): 2 training images

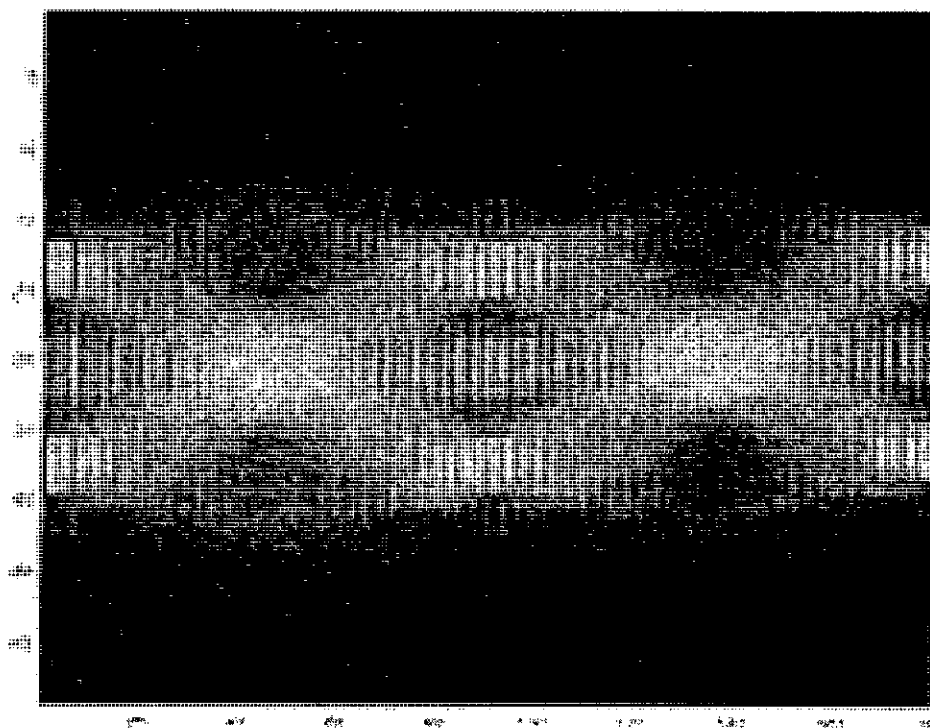


Fig 4.2 (b): 90 projections from each of the training images of fig. 4.2 (a)

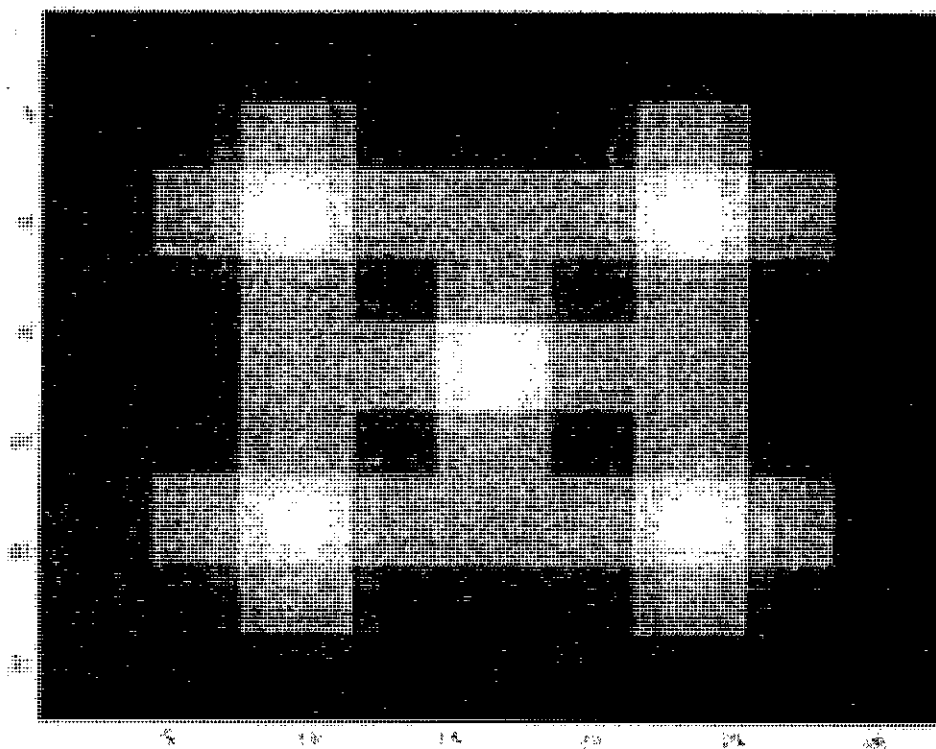


Fig 4.2 (c): Composite image using 2 training images in fig. 4.2 (a)

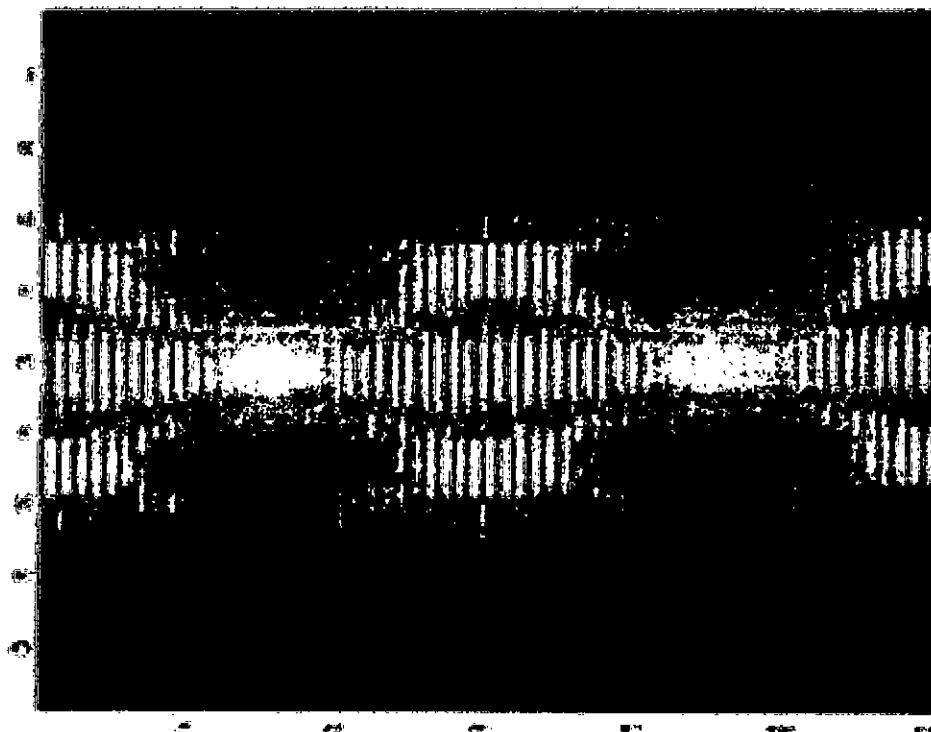


Fig 4.2 (d): 60 projections from each of the training images of fig. 4.2 (a)

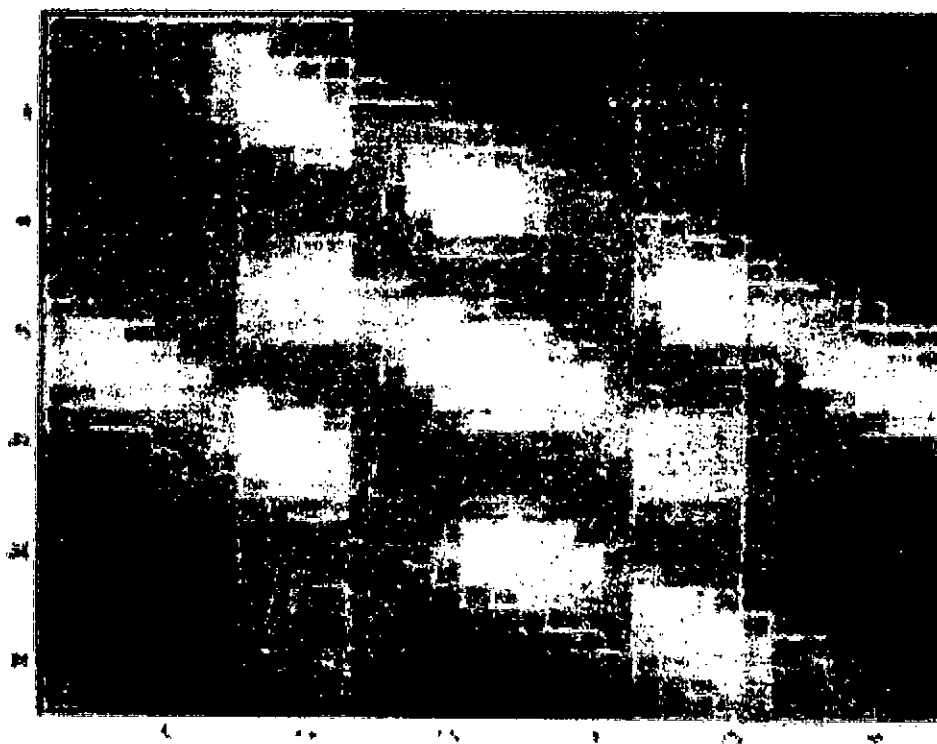


Fig 4.2 (e): Composite image using 2 training images in fig. 4.2 (a)

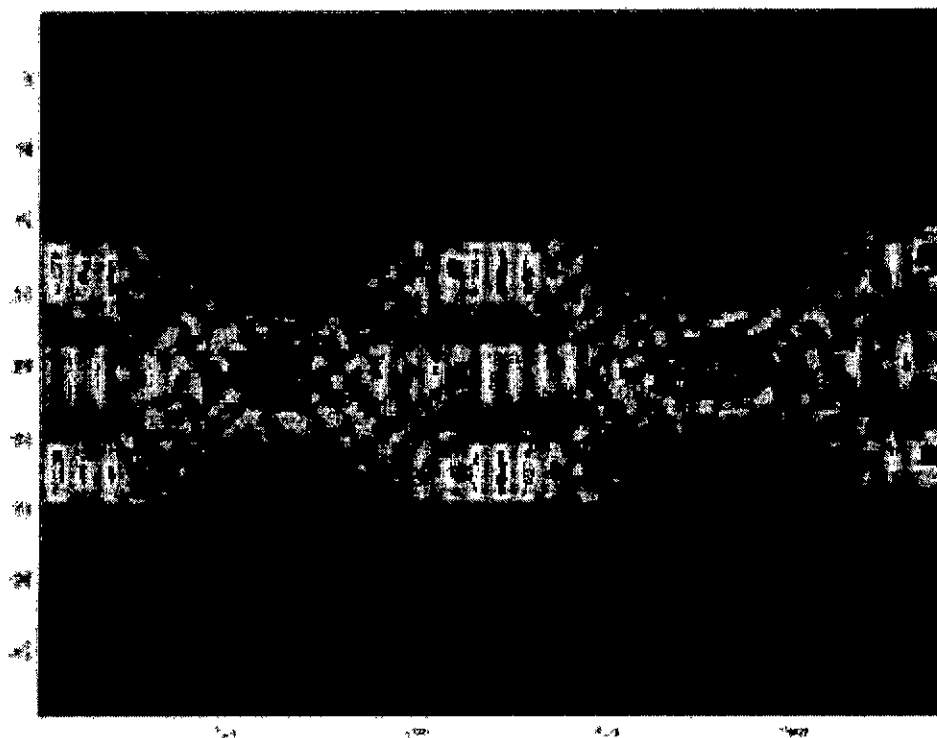


Fig 4.2 (f): 120 projections from each of the training images of fig. 4.2 (a)

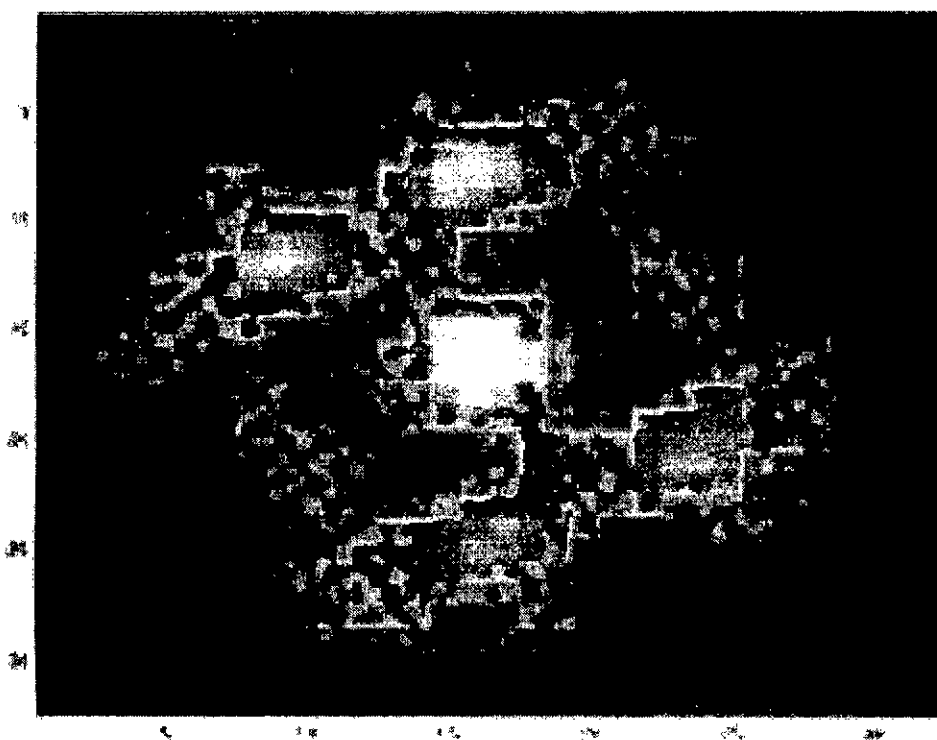


Fig 4.2 (g): Composite image using 2 training images in fig. 4.2 (a)

For composite image formation, the number of projections from each of the training image must depend on the image number of the training set and the projection number can be expressed as

$$M = \text{round}\left(\frac{180}{N}\right) \quad (4.12)$$

where M is the number of projections taken from each training image and N is the total number of training images. Fig. 4.2 (b) and (c) show the composite placement of all the projections taken from each of the training image as described in Eq. (4.7). In Fig. 4.2 (b), the number of projections taken is 90 for two training image as expressed by Eq. (4.12). But, in Fig. 4.2(d) and (f), an arbitrary number of projections are taken and in this case, the number of projection is 60 in Fig. 4.2(d) and 120 in Fig. (f). Fig. 4.2(c), (e) and (g), show the corresponding reconstructed composite image.

From the simulation output, it is obvious that only the Fig. (c) incorporates the necessary information of all the training images while the other composite images contain some erroneous information of the training set. This results due to the over or under number of slices in the composite image formation. Another important thing is that with the increase of training images, the number of slices will be reduced which may cause information loosing from each of the training image. Therefore, the number of training images and the number of slices must require an optimization for generation of a less distorted composite image.

4.5.2 Composite Image Generation for In-plane Rotation

For generation of composite image from a set of training images containing in-plane rotation, we have used binary character 'E' of 32×32 pixels. For simulation purpose, we have taken 9 training images rotated an interval of 5 degree. Therefore, the required number of slices for each training image is 20 and the maximum rotation distortion that is incorporated with this training image is 45 degree. To calculate the correlation peaks

after successive iteration, we have used phase-only JTC technique since it performs better for distorted images.

Fig 4.3 shows the nine (9) training images containing only in-plane rotation. Fig. 4.4 (a) shows the composite image where weight factors are all unity. To get an almost equal peak criterion, we have set the maximum number of iteration equal to 50 and the error limit equal to 0.01. Fig. 4.4 (b) shows the correlation output after first iteration. From figure 4.4 (b), it is observed that initially the peaks are not all equal due to dissimilarity between the training images.

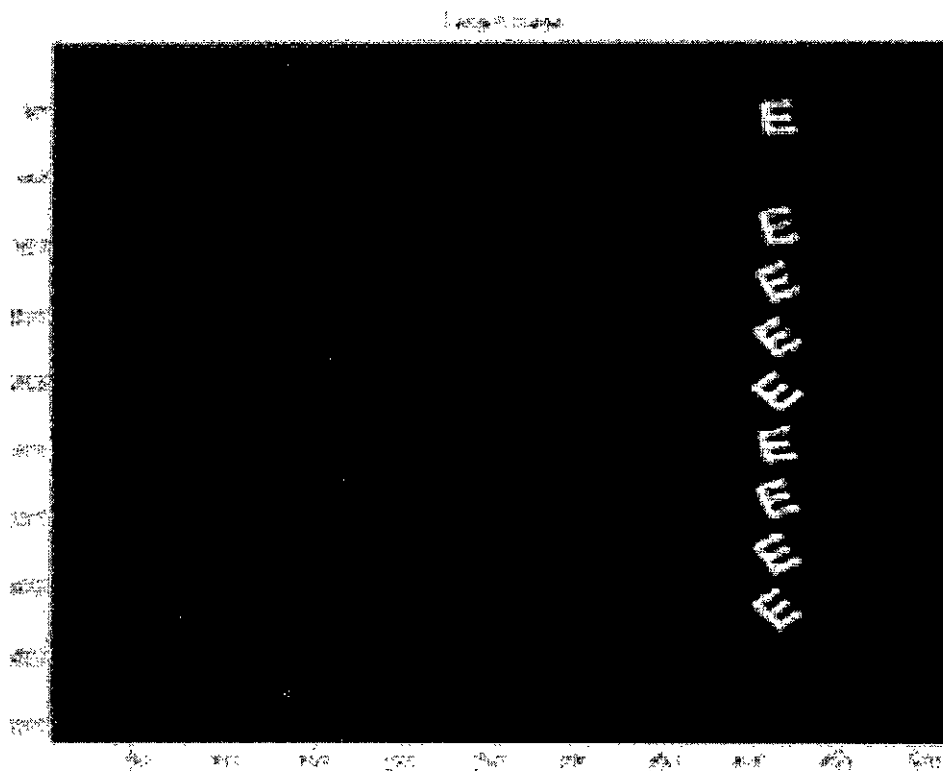


Fig 4.3: Training images

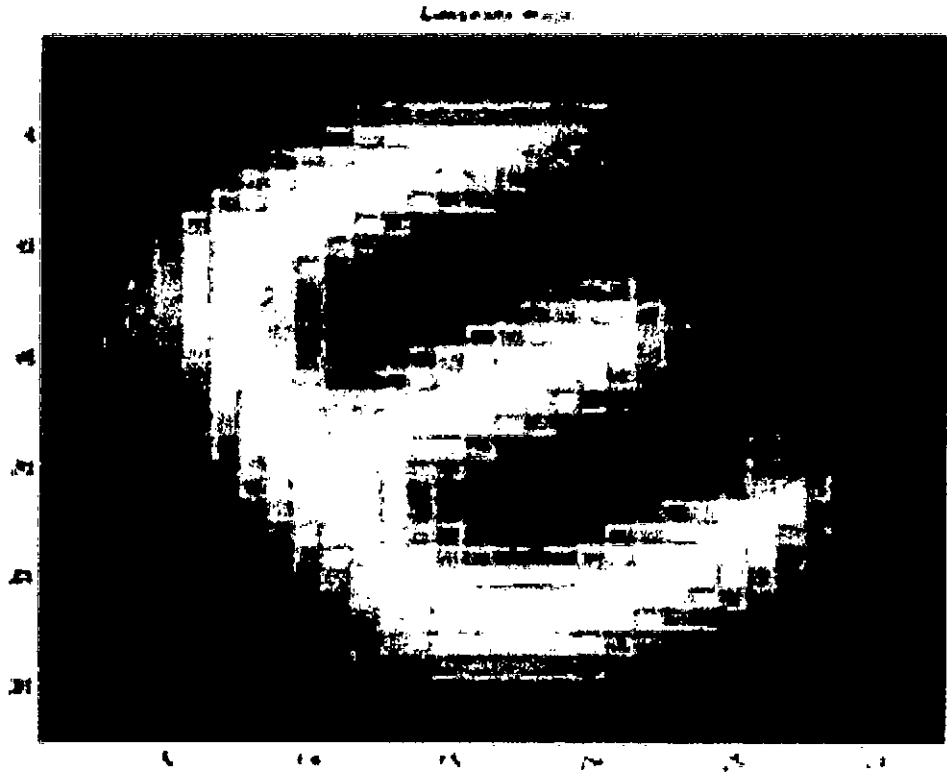


Fig 4.4 (a): Composite image using equal (=unity) weight factor

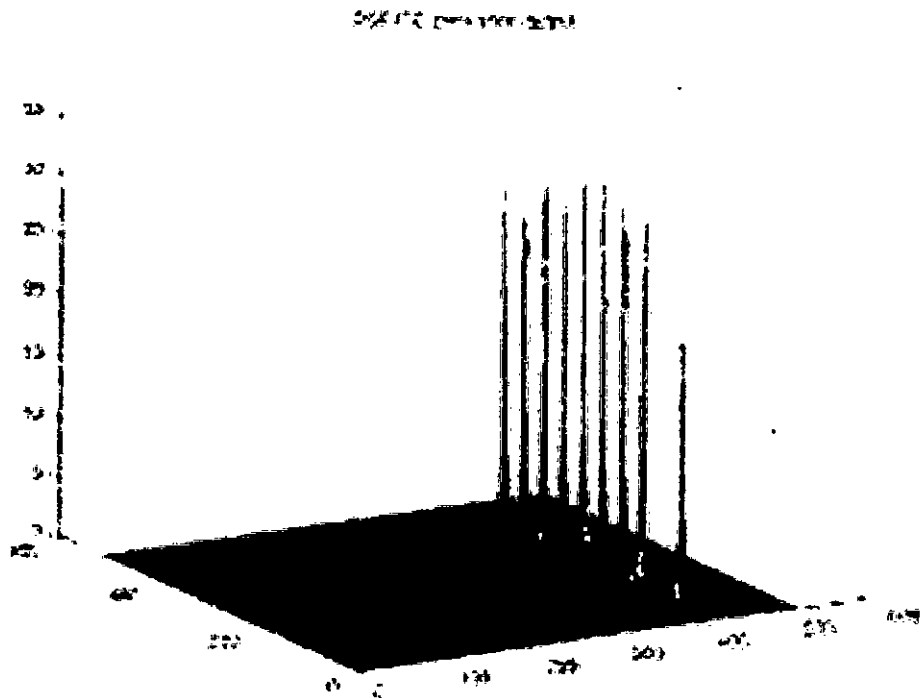


Fig 4.4(b): PSPFJTC output of all training images in fig. 4.3

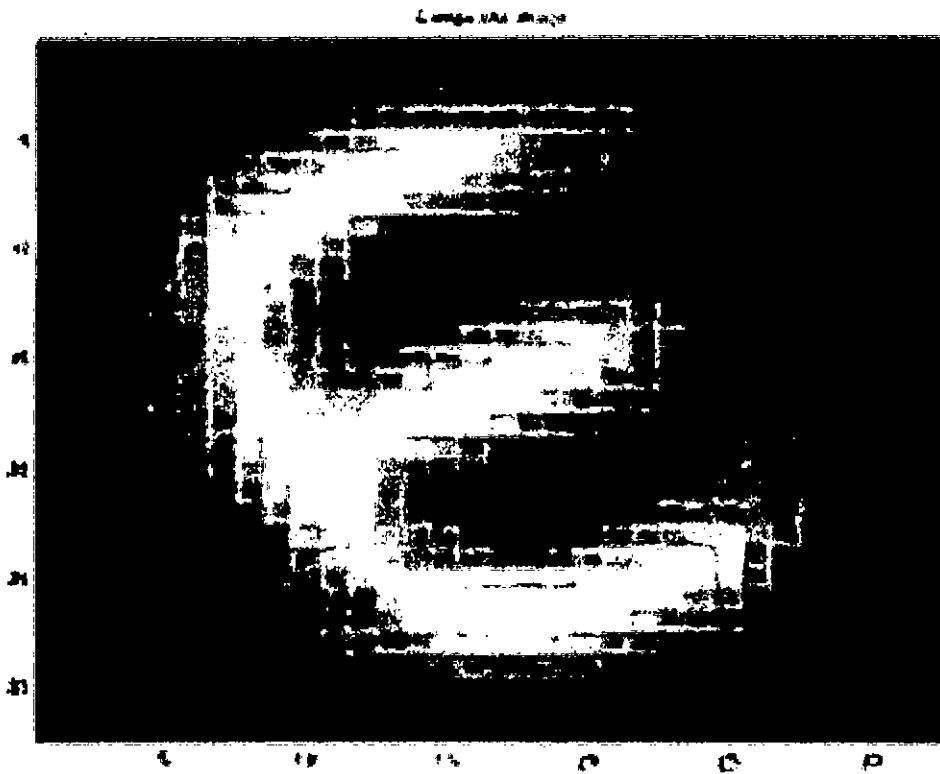


Fig 4.5 (a): Composite image after 100 iterations

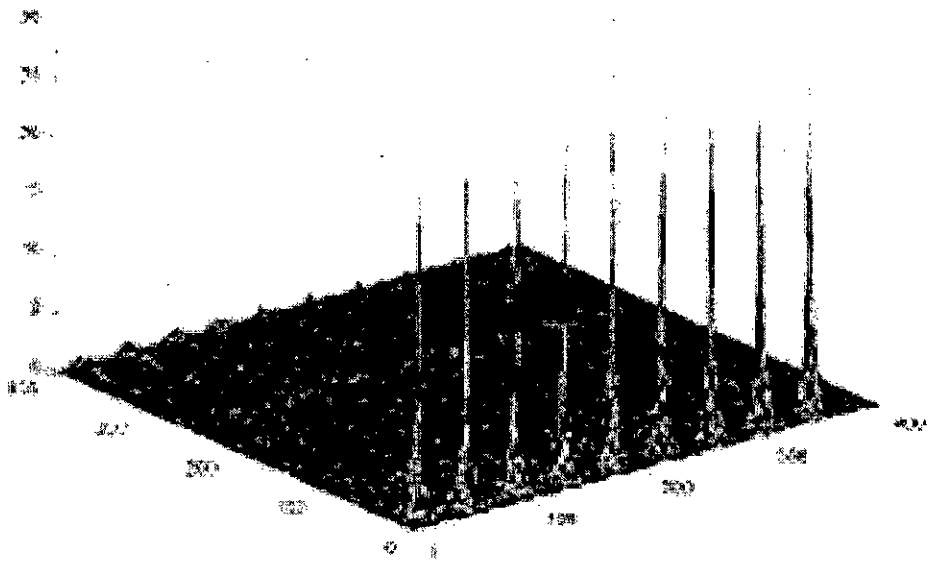


Fig 4.5 (b): PSPFJTC output of all training images in fig. 4.3

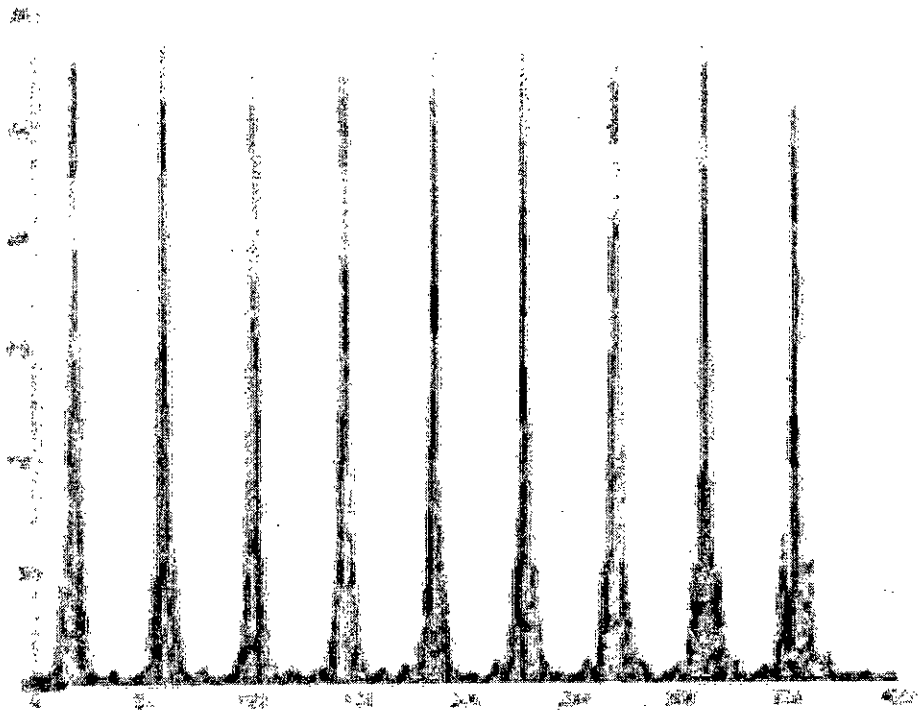


Fig 4.5 (c): Side-view of Fig 4.5 (b)

Fig 4.5 (a) shows the weighted composite image after 100 iterations. The corresponding correlation output of all the training images with the weighted composite image is shown in figure 4.5 (b). Here all most equal correlation peaks are generated for all the training images of the training set. For better visualization, the side view of figure 4.5 (b) is shown in figure 4.5 (c). The value of correlation peaks after 100 iterations is given in table 4.1 as below

Table 4.1: Peak intensity and weight factor after 100 iterations

Image Number	1	2	3	4	5	6	7	8	9
Peak Intensity	27.8128	28.6982	28.2498	28.4420	28.4112	28.4685	28.2876	28.5472	28.1385
Weight factor	14.7462	1.3918	6.8685	5.4971	4.6699	4.4365	7.2524	4.1097	7.3869

Though all the peaks are not exactly equal height, the difference between them is so small that we can easily neglect it. Therefore, the composite image generated in this way is the desired one that may be used for distortion invariant (in-plane) pattern recognition.

4.5.3 Performance Evaluation of In-plane Rotated Composite Image

To investigate the performance of the generated composite image, we have replaced some of the training images in the training set with some similar and dissimilar non-target objects or images. For similar non-target objects we have used the character 'F' and for dissimilar non-target object we have introduced the character 'M'. Both the characters 'F' and 'M' have the same size as the composite reference image of 'E'. Fig. 4.6(a) shows the joint image where the composite image of character 'E' on the upper left corner represent the reference image and all other characters on the right side



Fig 4.6 (a): Input binary joint image of characters containing in-plane rotated target and non-target images

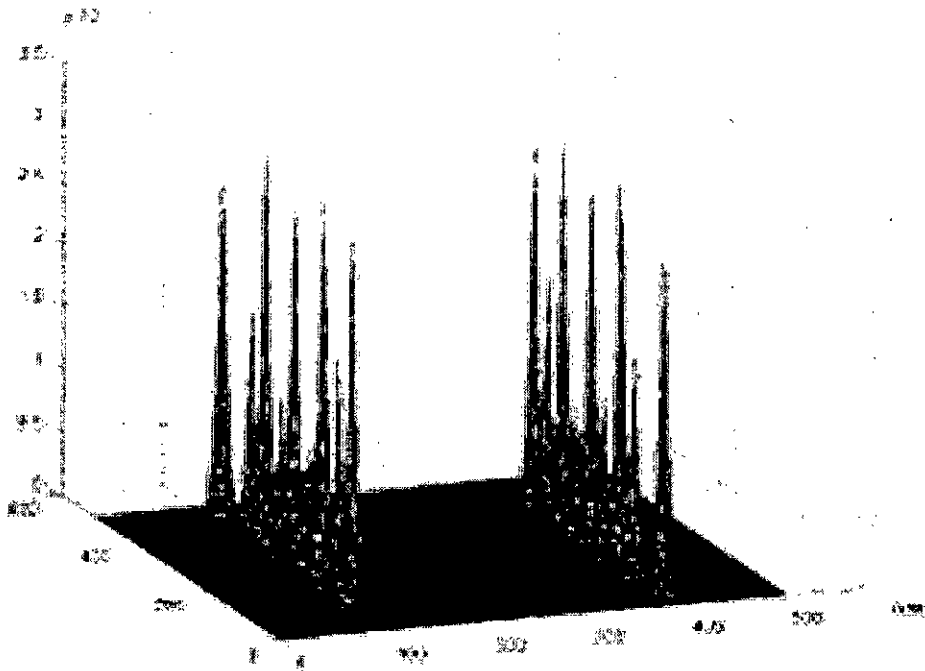


Fig 4.6 (b): CJTC output of fig. 4.6 (a) after image subtraction

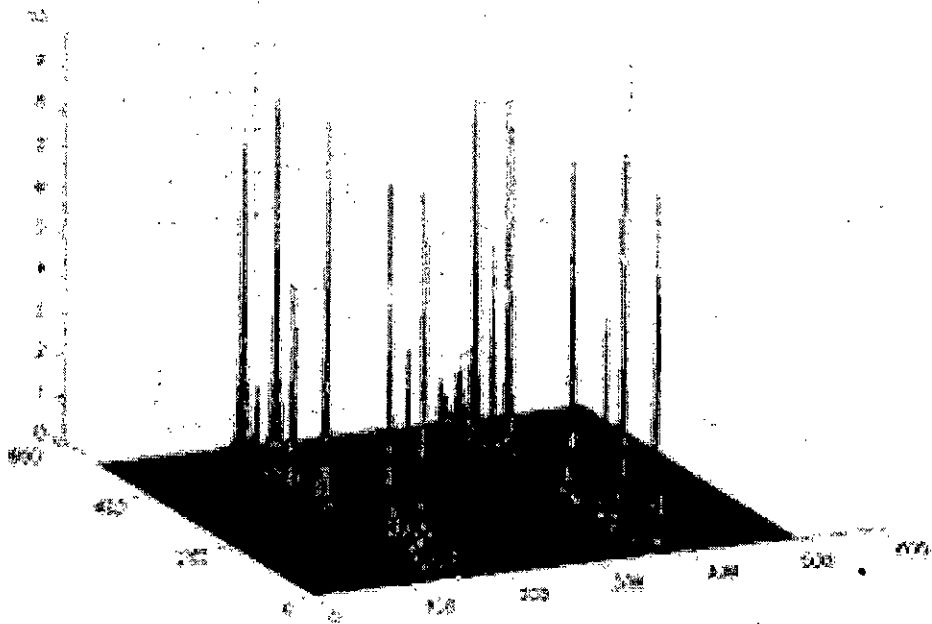


Fig 4.6 (c): POJTC output of fig. 4.6 (a)

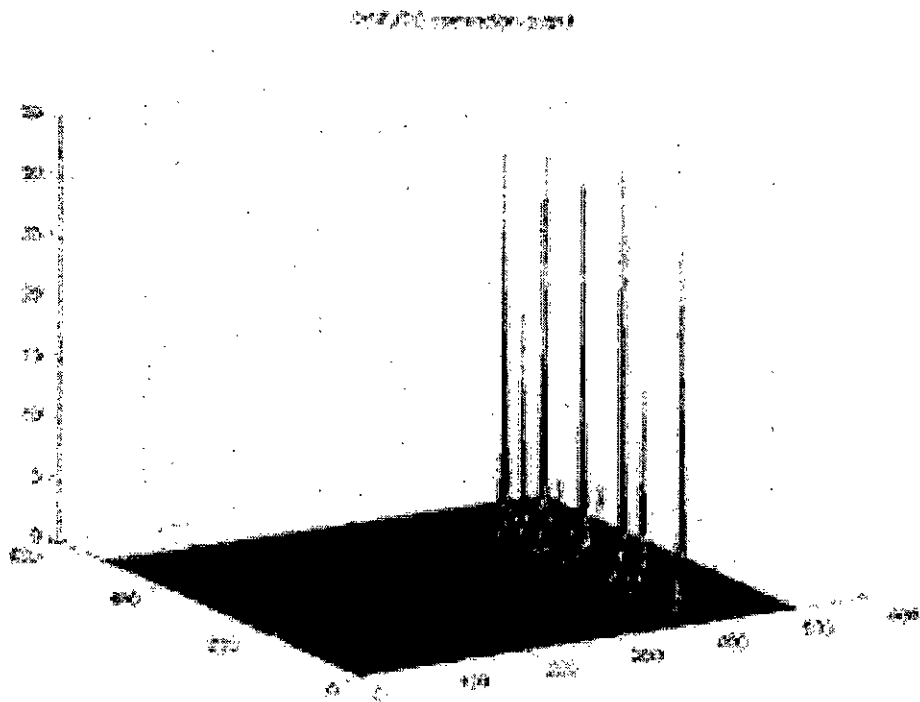


Fig 4.6 (d): PSPFJTC output of fig. 4.6 (a)

represent the input scene images. Here the input scene contains five (5) target images and four (4) non-target images.

Fig. 4.6 (b) shows the classical JTC output with image subtraction. From the figure, it is obvious that there are five pairs of high peaks for five target images while two appreciable peaks for two similar non-target images (for 'F') and two negligible peaks for two dissimilar non-target images (For 'M'). Again, all the target peaks are not almost equal and the discrimination ratio (the ratio between the minimum target peaks and maximum non-target peaks) is not so high. Fig. 4.6 (c) shows the phase only JTC output with image subtraction technique. Here a pair of sharp delta like correlation peaks is generated for each target object and the height of correlation peaks are almost equal with a slight difference in a pair of peaks. Finally, figure 4.6 (d) shows the correlation output of PSPFJTC technique. From figure 4.6 (d), it is evident that all the target peaks are almost equal. Though there is slight difference among the values of the correlation peaks, this can easily avoided for target detection purpose. The major advantage of PSPFJTC technique is that it produces a single correlation peak for each target object and therefore it can be easily incorporated in class-associative target detection purpose.

4.5.4 Composite Image Formation for Out-of-plane Rotation

The major difference between in-plane rotation and out-of-plane rotation is that in in-plane rotation, the feature of the image remains same only a rotated version of the original image. But in case of out-of-plane rotation, new features may be incorporated with the original image. In normal SDF approach, these new features may interfere with the original features and may lead to erroneous result. Since in PSDF approach, all the slices of the individual training images are placed side-by-side, there is no theoretically no chance of interference among all the features of the training set. Therefore, PSDF based approach is highly preferable one for three dimensional distortion invariant target detection process.

The out-of-plane rotation distortion can be widely classified into two categories. First, the distortion introduced due to the azimuth change and the second one is due to the change of elevation of the target objects. Here first, we have considered these two distortions individually and then a combination of both is used to verify the system.

4.5.4.1 Composite Image for Azimuthally Rotated Objects

To incorporate the three-dimensional distortion mainly out-of-plane rotation due to azimuth change, we have used gray-level car images of size 64×48 pixels. Here all the training images contain only the azimuth change and there is no variation in elevation angle. Fig. 4.7 (a) shows the PSDF composite image after 100 iterations. Fig. 4.7 (b) shows all the training images used for the composite image generation. We have used PSPFJTC technique with filter parameter $C=1$ and $D=1e-9$. The error limit is set 0.01 and the maximum iteration number equal to 100. After 100 iterations, the PSPFJTC output of all the training images is shown in Fig. 4.7 (c). Fig. 4.7 (d) shows the side view of Fig. 4.7 (c) for better visualization of the height of correlation peaks. Table 4.2 shows the numerical values of correlation peaks of all the training images and the corresponding weight factor associated for the generation of the composite image.

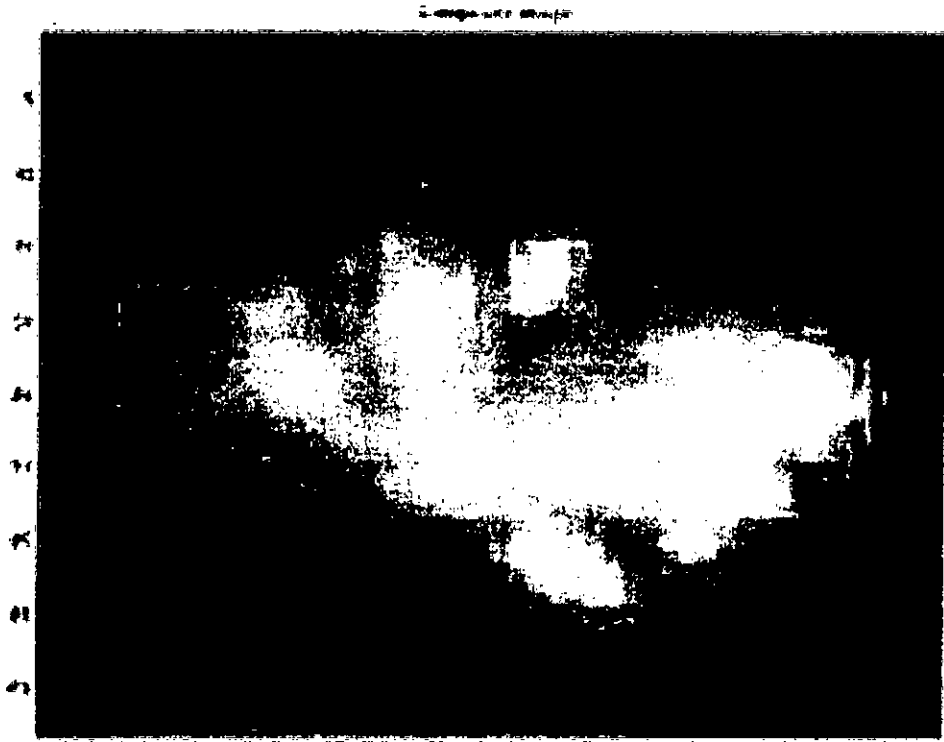


Fig. 4.7 (a): Composite gray-level car image for azimuth change

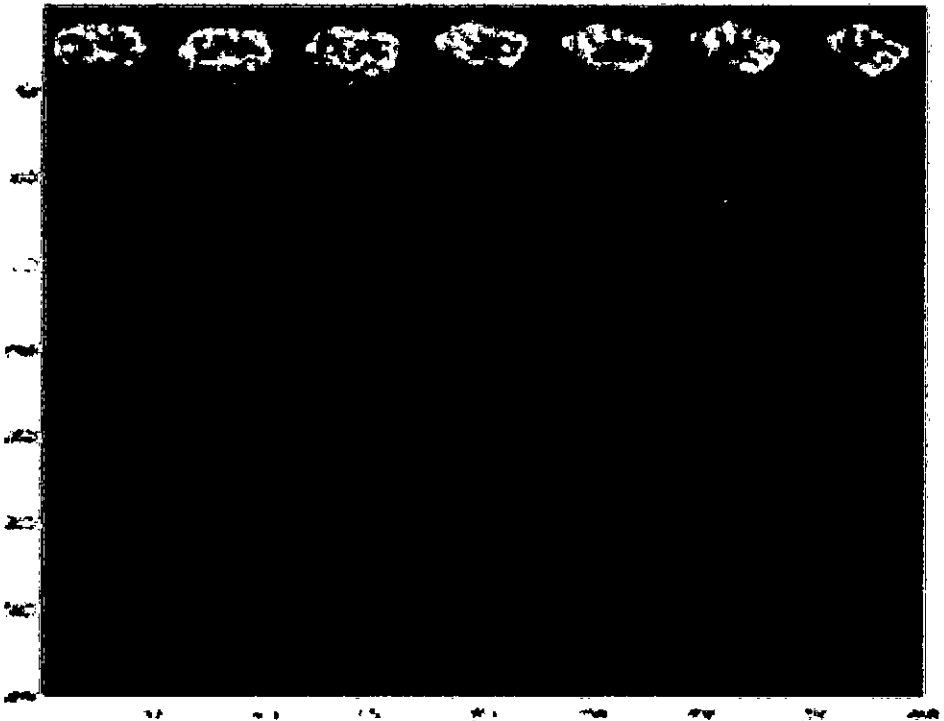


Fig. 4.7 (b): Training images

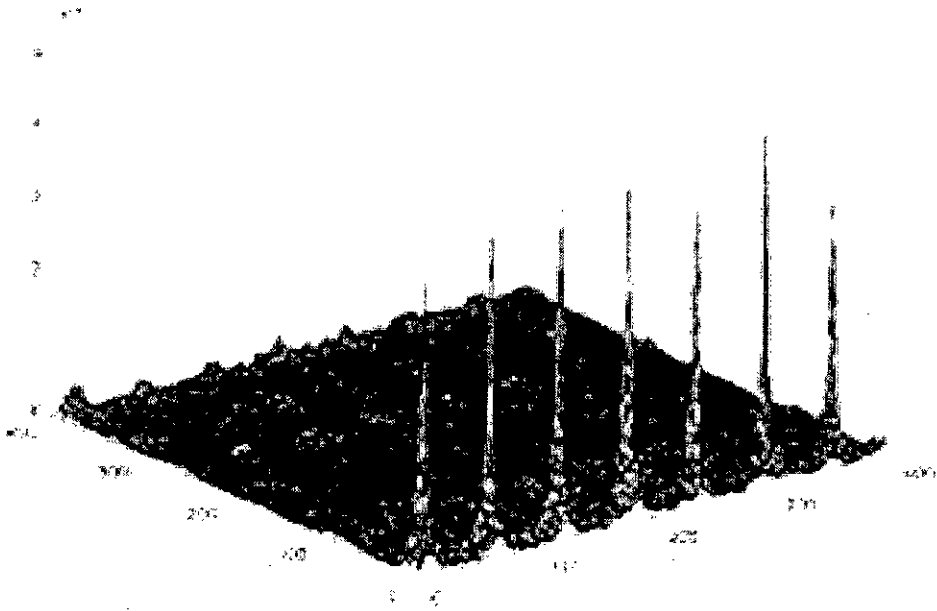


Fig 4.7 (c): PSPFJTC output of all training images in fig. 4.7 (b)

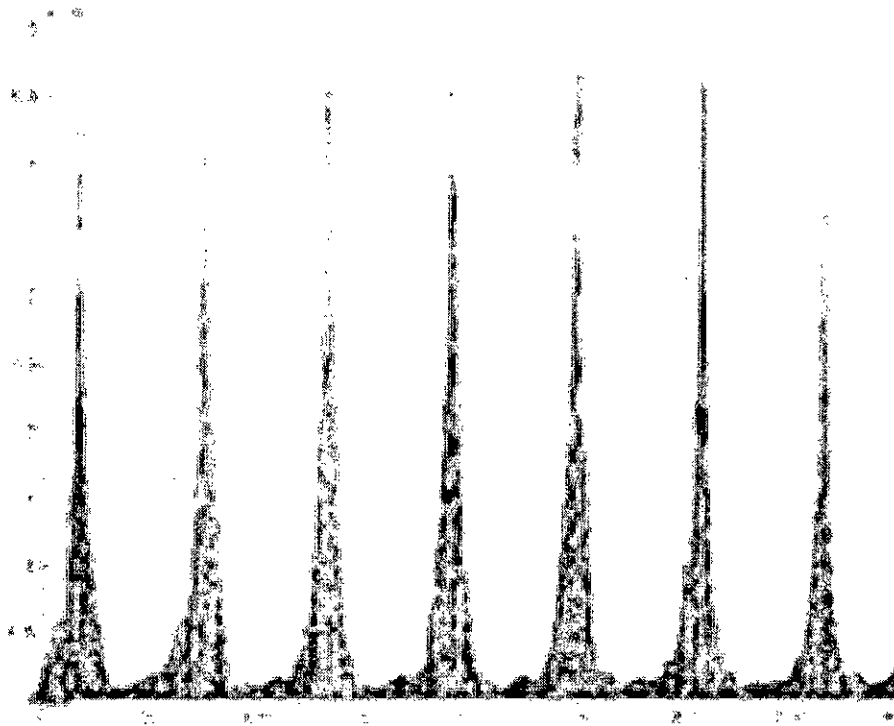


Fig. 4.7 (d): Side view of Fig. 4.7 (c)

Table 4.2: Peak intensity and weight factor after 100 iterations

Image number	1	2	3	4	5	6	7
Peak intensity	10.7485	10.9860	11.0965	11.5194	12.0820	11.5771	8.6587
Weight factor	84.3795	69.9793	68.5629	39.0600	1.3340	53.1598	204.0730

From fig. 4.7 (c), (d) and Table 4.1, it is evident that though the peaks are not equal, the difference among them is quite small and therefore they are quite acceptable for target detection purposes.

4.5.4.2 Performance Evaluation of the Composite Image for Azimuthally Rotated Objects

To investigate the performance of the generated composite image, we have introduced both target and non-target objects of same size in the input scene. First of all, Fig. 4.8 (a) depicts the joint image containing the composite reference image and input scene images of target objects only. Fig. 4.8 (b) and (c) show the phase only JTC (POJTC) and PSPFJTC outputs of the joint image of Fig.4.8 (a).

In POJTC technique, four distinct pair of peaks is generated for four target objects while in PSPFJTC technique, four single correlation peaks is generated for the same number of target objects. The variation of correlation peak height is slightly less than that of the PSPFJTC technique. In case of target detection process, the number of distinct peaks is important rather than the equal peaks of all the target objects. In that reason, both the systems are quite suitable for distortion invariant target detection process.

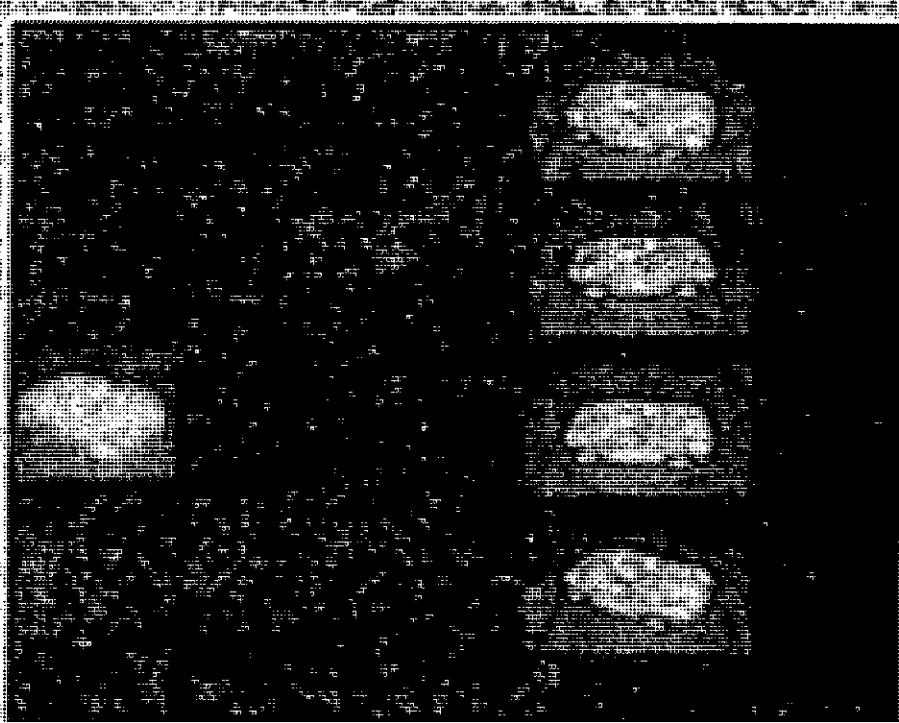


Fig. 4.5 (a). Joint image containing all target images.

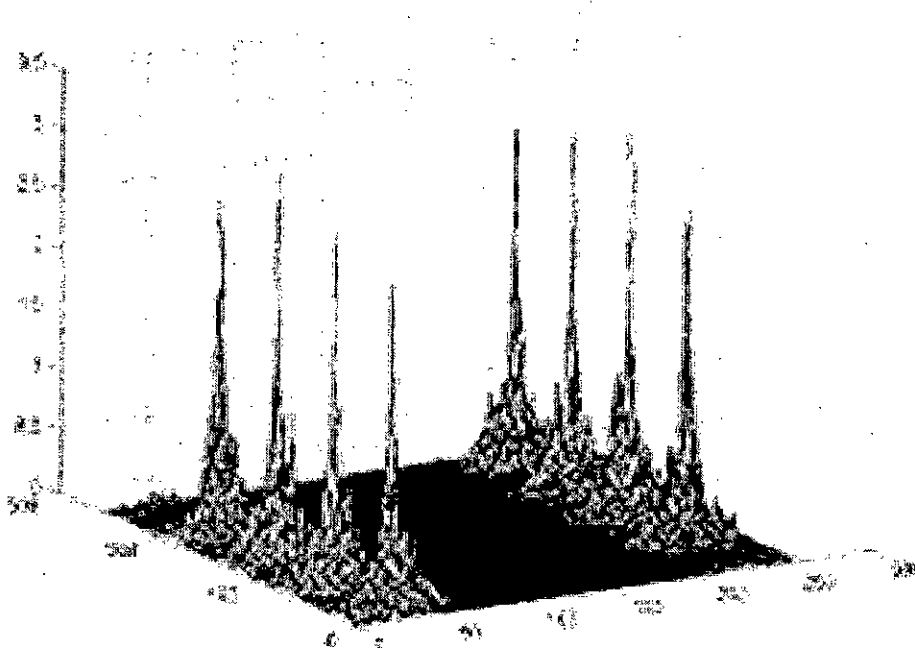


Fig. 4.8 (b): POJTC output of Fig. 4.8 (a)

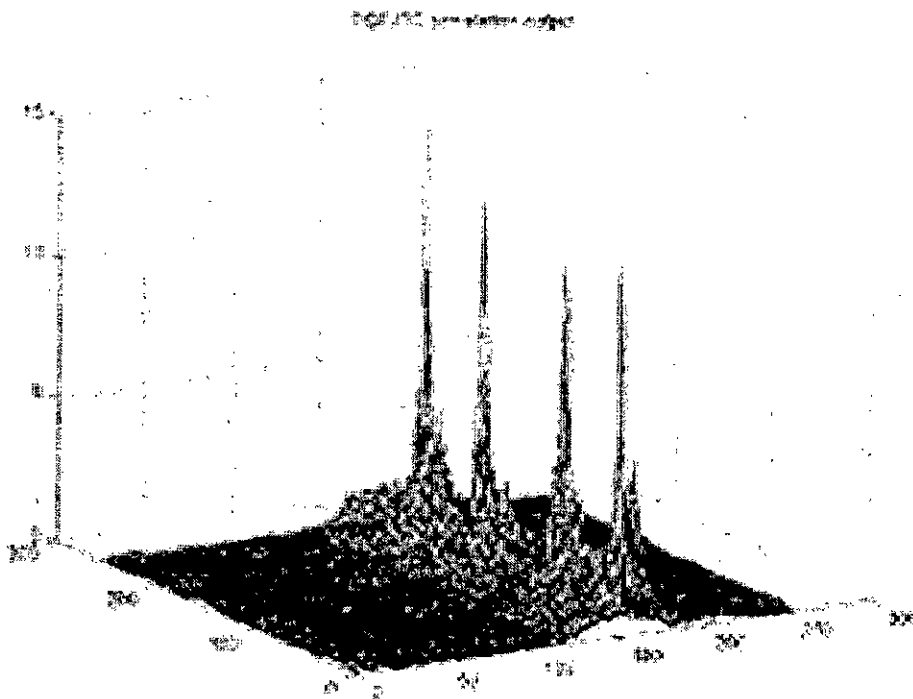


Fig 4.8 (c): PSPFJTC output of Fig. 4.8 (a)

To compare the formation of target and non-target peaks and their relative height, we have placed non-target objects of different car model of same size in the input scene. Fig. 4.9 (a) shows the joint image where the input scene contains two target images and two non-target images of almost same orientation. Fig. 4.9 (b) and (c) show the POJTC and PSPFJTC output of the joint image of Fig. 4.9 (a). Both the techniques produce distinct target peaks for two target objects and almost negligible non-target peaks for non-target objects. Therefore, both the techniques can successfully detect the target objects from the input scene.

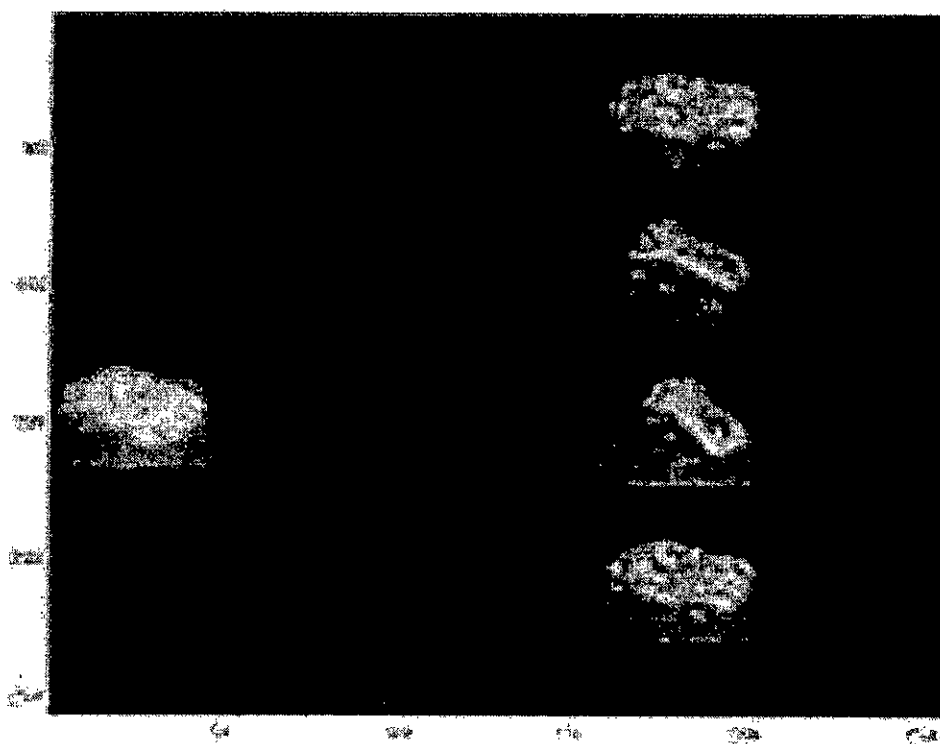


Fig 4.9 (a): Joint image containing two target and two non-target images

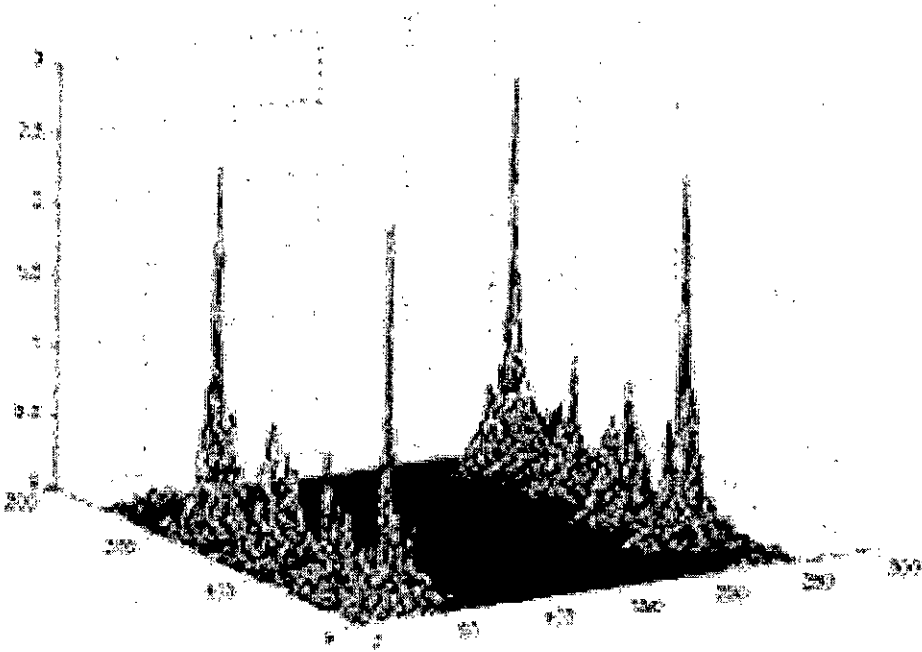


Fig. 4.9 (b): POJTC output of Fig. 4.9 (a)

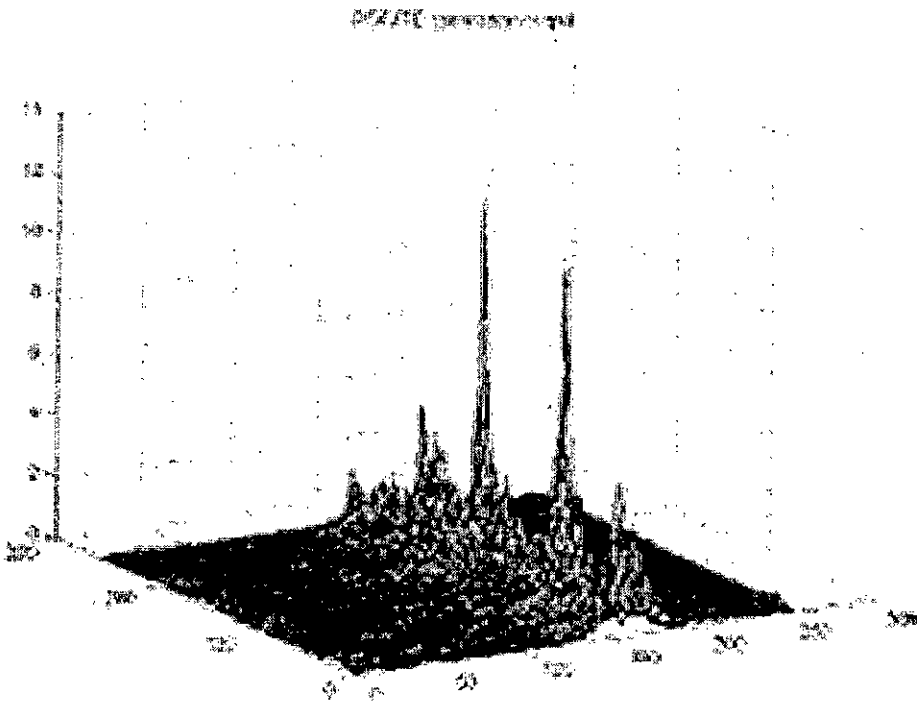


Fig. 4.9 (c): PSPFJTC output of Fig. 4.9 (a)

Finally, to compare the performance of the system with the increasing number of non-target objects, we have placed three non-target objects with only one target object as shown in Fig. 4.10 (a). As previous, the POJTC and PSPFJTC outputs of the Fig. 4.10 (a) are shown in Fig. 4.10 (b) and (c) respectively. It is evident from both the figures that there is distinct pair of target peaks (POJTC technique) or a single target peak (PSPFJTC technique) and almost none or negligible peaks for non-target objects. This certainly ensures the better performance of the generated composite image and effectiveness of both the techniques.

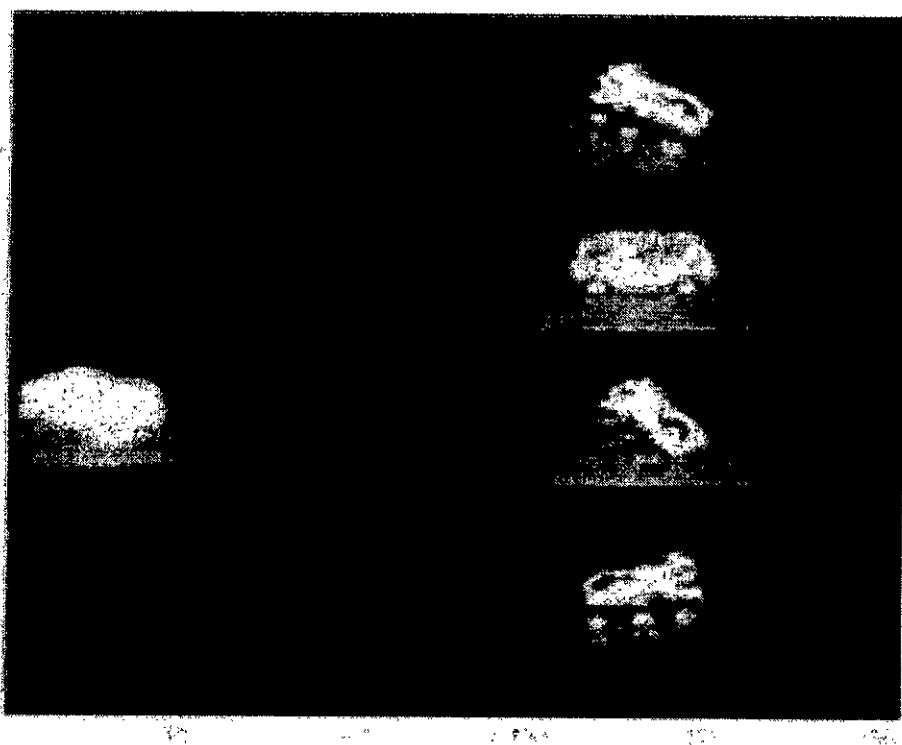


Fig. 4.10 (a): Joint image containing one target and three non-target images

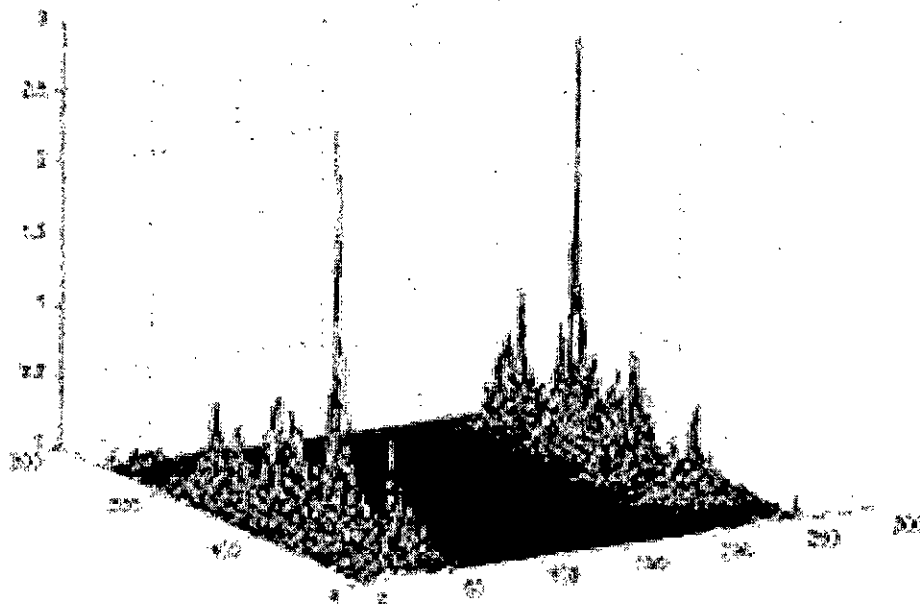


Fig. 4.10 (b): POJTC output of Fig. 4.10 (a)

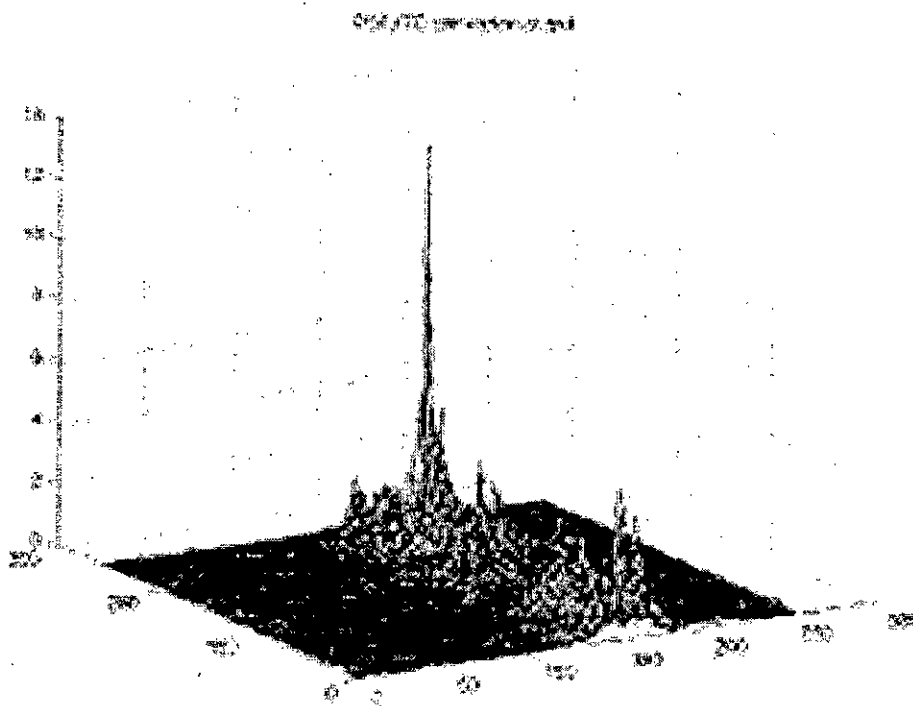


Fig. 4.10 (c): PSPFJTC output of Fig. 4.10 (a)

4.5.4.3 Composite Image for Change of Elevation Angle of Objects

To incorporate the distortion due to change of elevation angle, we have used the same gray-level car images of size 64×48 pixels. Here all the training images contain only the change of elevation angle and there is no variation in azimuth. Fig. 4.11 (a) shows the generated composite whereas the Fig. 4.11(b) shows all the training images used for the composite image generation.

We have used PSPFJTC technique with filter parameter $C=1$ and $D=1e-9$. The error limit is set 0.01 and the maximum iteration number equal to 50. After 50 iterations, the PSPFJTC output of all the training images is shown in Fig. 4.11 (c). Fig. 4.7 (d) shows the side view of Fig. 4.11 (c) for better visualization of the height of correlation peaks. Table 4.3 shows the numerical values of correlation peaks of all the training images and the corresponding weight factor associated for the generation of the composite image.

Table 4.3: Peak intensity and weight factor after 50 iterations

Image number	1	2	3	4	5	6	7
Peak intensity	9.7964	9.8940	9.8138	9.8377	9.8761	9.9300	9.7889
Weight factor	17.7293	7.4464	16.1234	11.0276	9.5799	7.1354	15.1704

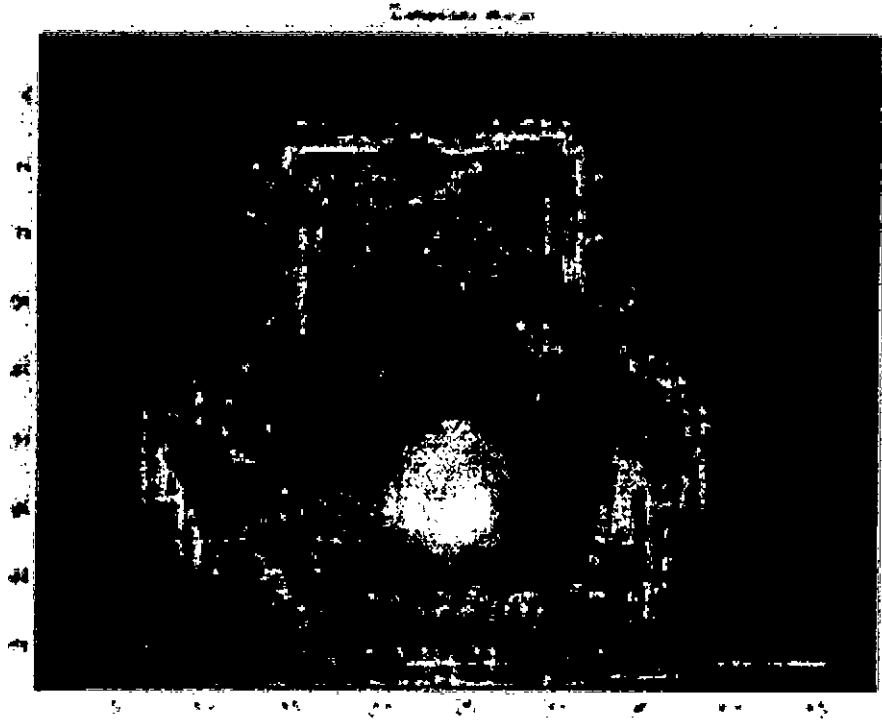


Fig. 4.11 (a): Composite image of gray-level car image for change in Elevation angle



Fig. 4.11 (b): Training images

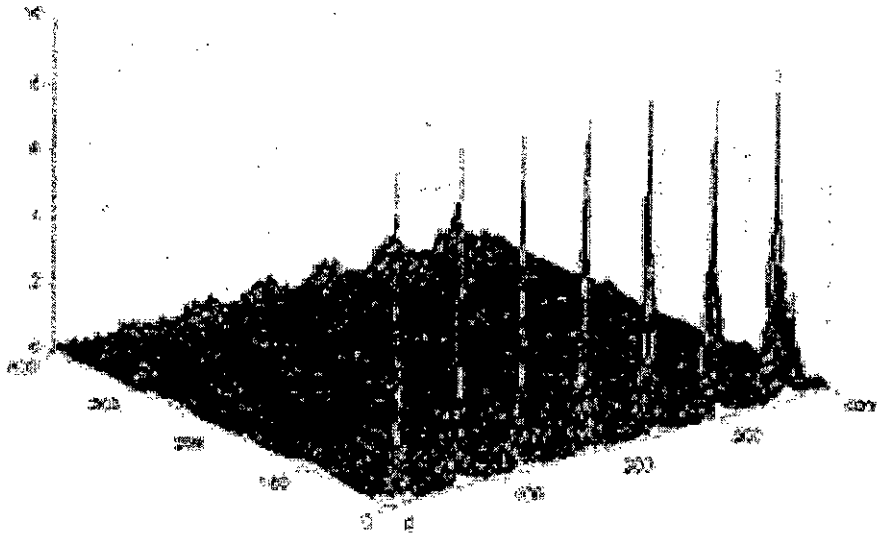


Fig. 4.11 (c): PSPFJTC output of all training images in fig. 4.11 (b)

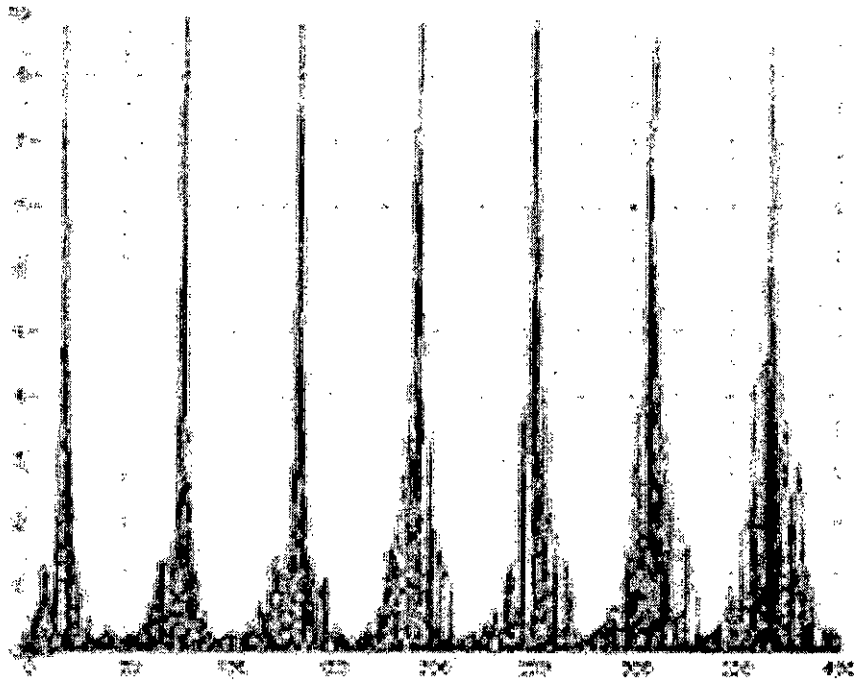


Fig. 4.11 (d): Side view of Fig. 4.11 (c)

4.5.4.4 Performance Evaluation of the Composite Image for Change of Elevation Angle of Objects

To investigate the performance of the generated composite image as shown in Fig 4.11(a), we have placed non-target images of almost same orientation and size with the target images and correlation operation is performed by both the POJTC and PSPFJTC techniques. Fig. 4.12 (a) shows the joint image where the image at the lower half portion represents the composite reference image and the images on the upper half portion represent the input scene images. Here the input scene contains three target images of different elevation angle. It also contains four non-target images of almost same orientation and size of the target images. Fig. 4.12 (b) and (c) show the POJTC and PSPFJTC output of the figure 4.12 (a). From Fig. 4.12 (b) and (c), it is obvious that POJTC technique produces three distinct pair of peaks for three target objects while PSPFJTC technique produces only three peaks for the same number of target objects.

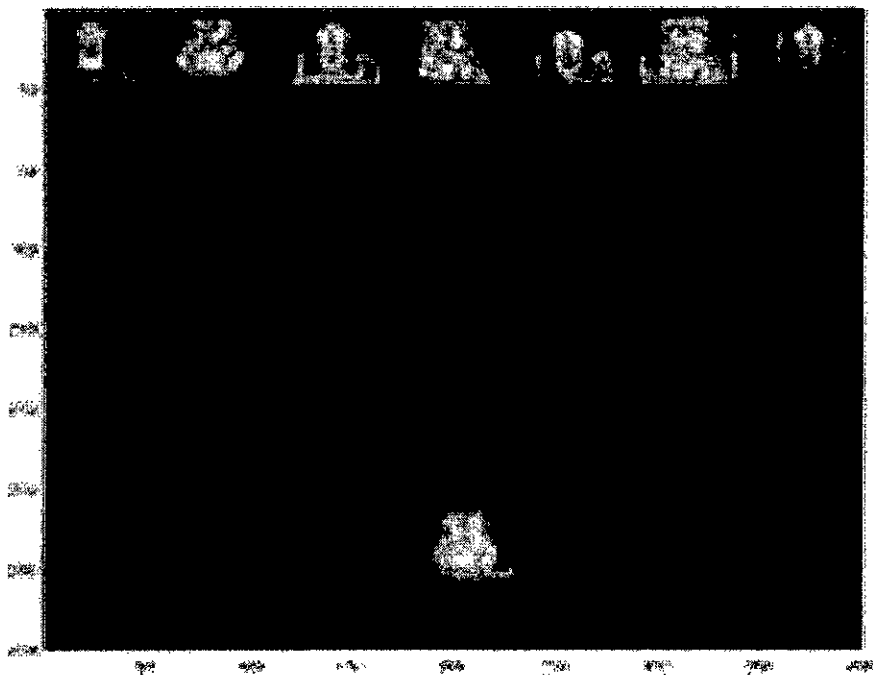


Fig. 4.12 (a): Joint image containing three target and four non-target images

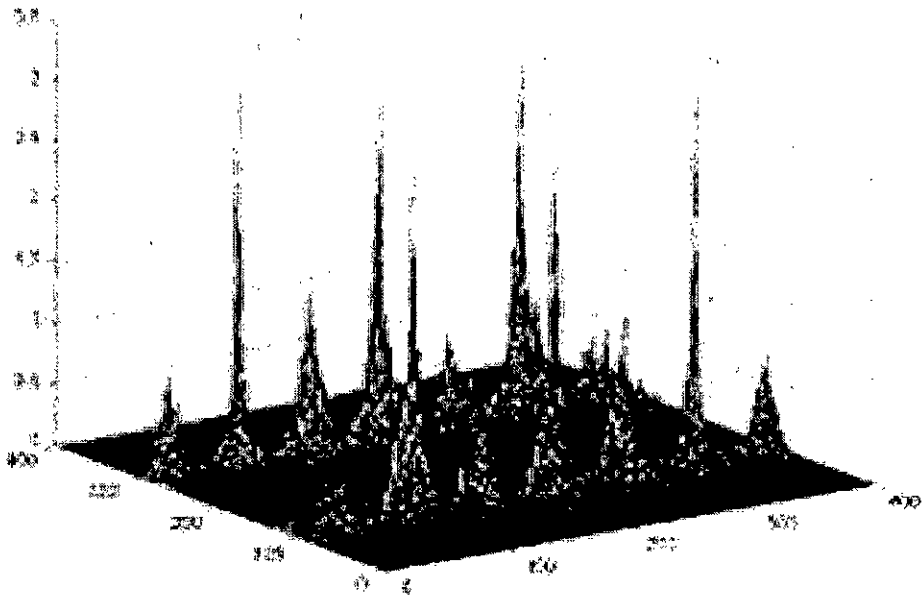


Fig. 4.12 (b): POJTC output of Fig. 4.12 (a)

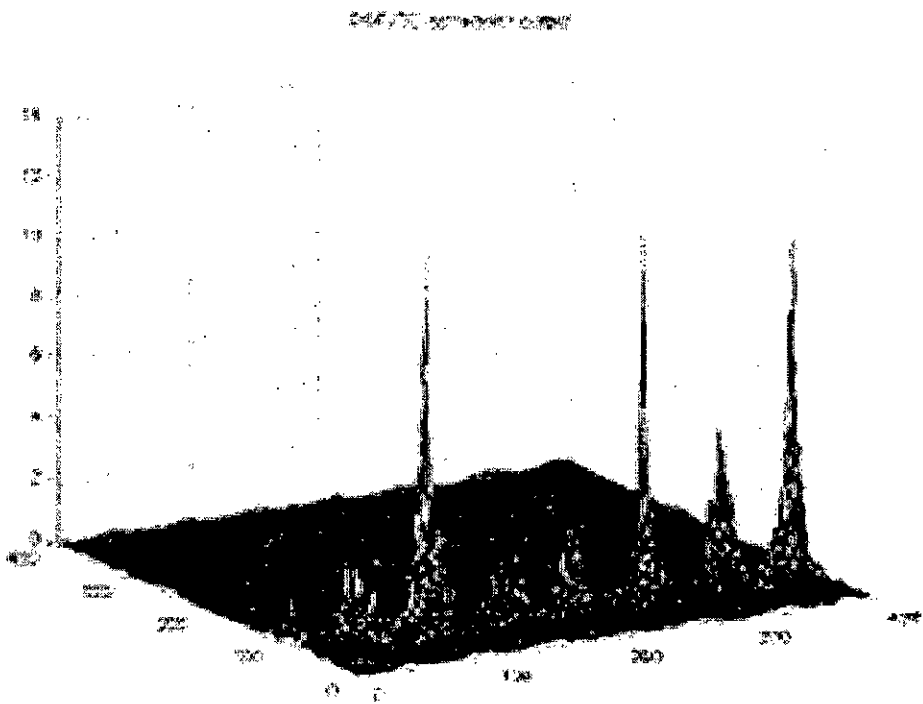


Fig. 4.12 (c): PSPFJTC output of Fig. 4.12 (a)

Next the performance of the system is investigated with the increased number of non-target objects. Fig. 4.13 (a) shows the joint image where the input scene contains six non-target images with only one target image. As usual Fig. 4.13 (c) and (d) show the POJTC and PSPFJTC outputs of the joint image respectively. Here again both the techniques can produce distinct correlation peaks for the target objects. Though there is presence of some non-target peaks, the discrimination between them is appreciable and therefore, we can easily avoid them.

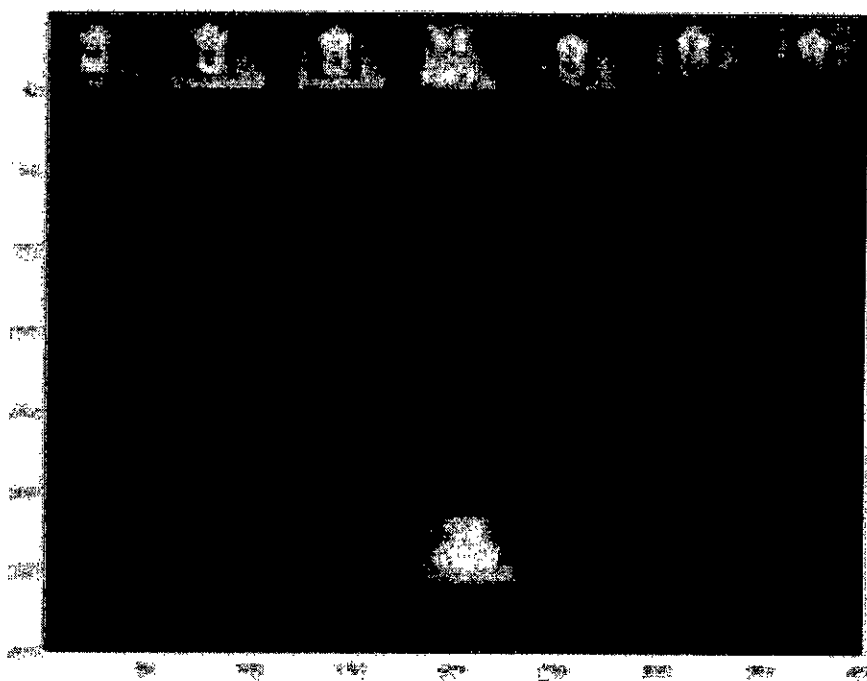


Fig. 4.13 (a): Joint image containing one target and six non-target images

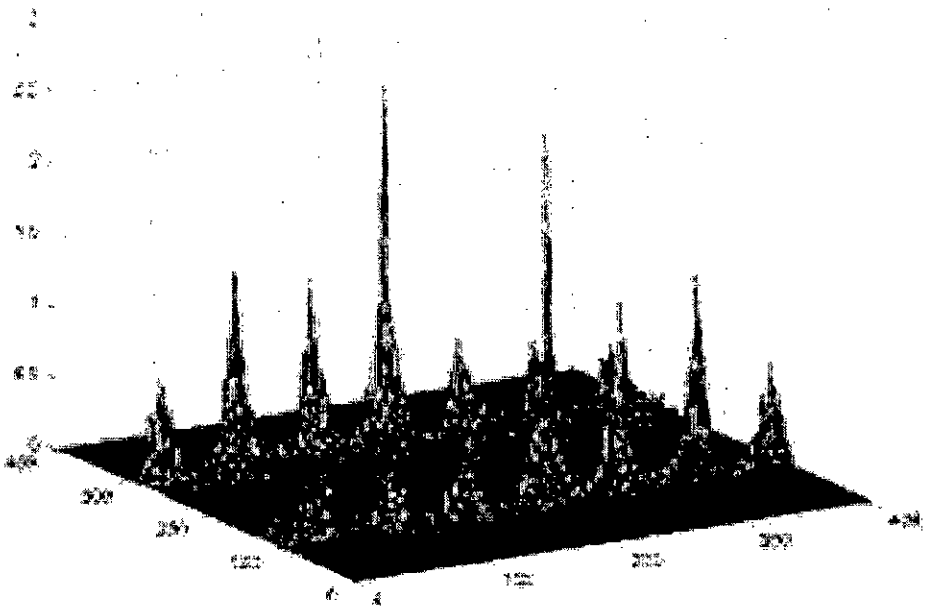


Fig. 4.13 (b): POJTC output of Fig. 4.13 (a)

POJTC output of Fig. 4.13 (a)

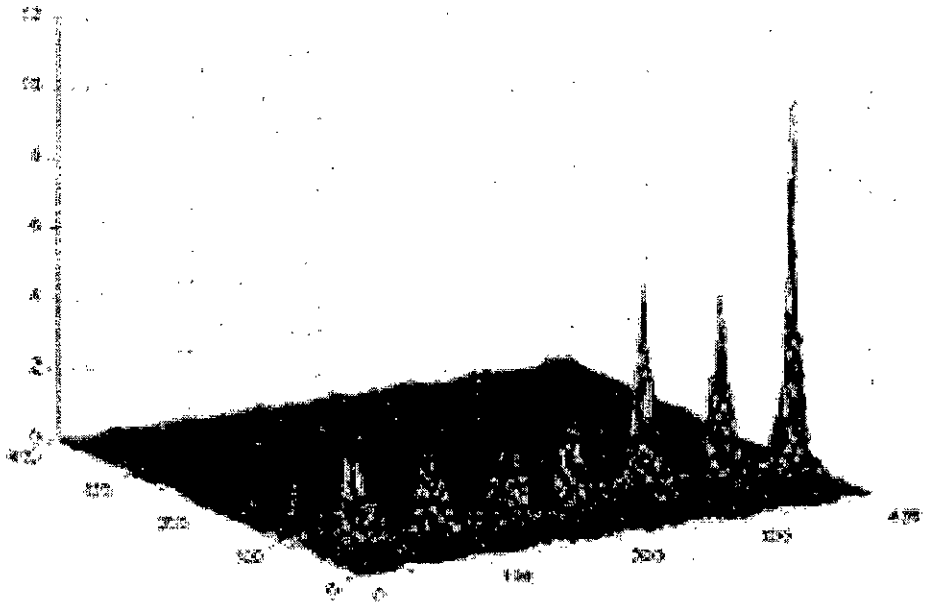


Fig. 4.13 (c): PSPFJTC output of Fig. 4.13 (a)

4.5.4.5 Performance Evaluation of the Composite Images for 3-D Distortions

To investigate the performance of the composite images for 3D distortions containing both azimuth change and change of elevation angle, we have placed the two composite reference images as shown in Fig. 4.7 (a) and Fig. 4.11 (a), side-by-side in the joint image. These two reference images now form the joint composite reference image where both azimuth and change of elevation angle are incorporated. Next the class-associative target detection algorithm is applied for 3D distortion invariant target detection purpose. Here the FAF filter as expressed in Eq. 3.16 is used with filter parameters $C=1$, $D=1e-9$ and $m=1$. The value of α and β are set by trial and error method to get appreciable correlation height for each of the target objects. Fig. 4.14 (a) shows the joint image where two images on the lower half portion represent two

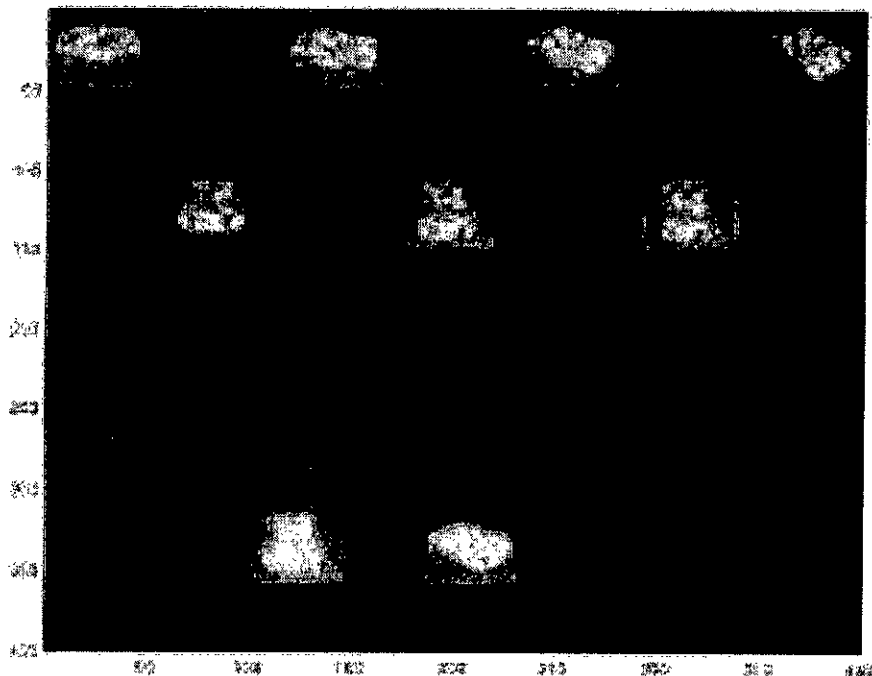


Fig. 4.14 (a): Gray-level joint image containing out-of-plane rotated target images only

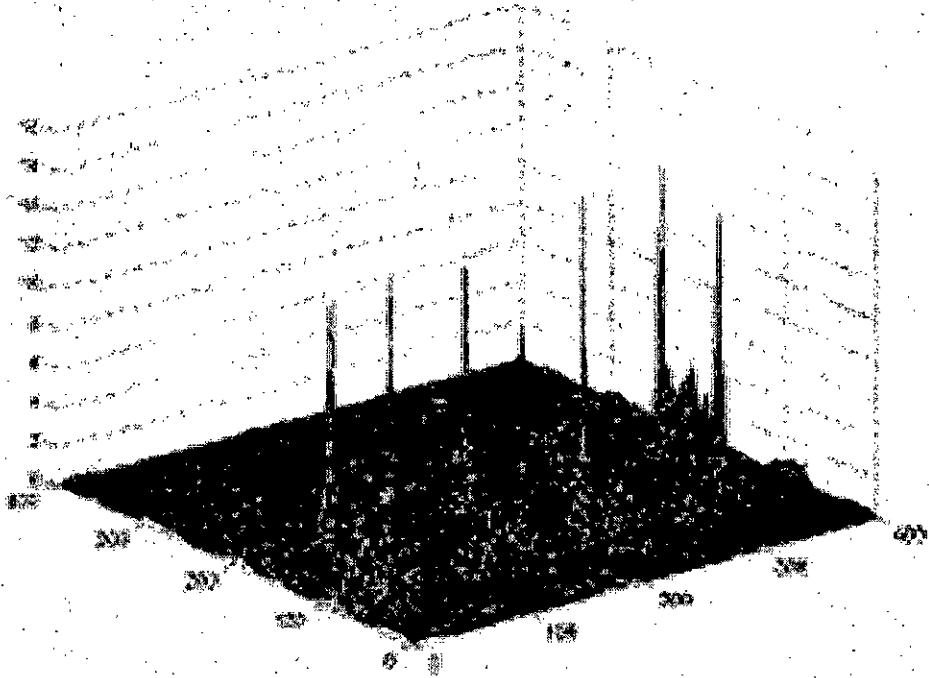


Fig. 4.14 (b): PSpFJTC output of Fig. 4.14 (a)

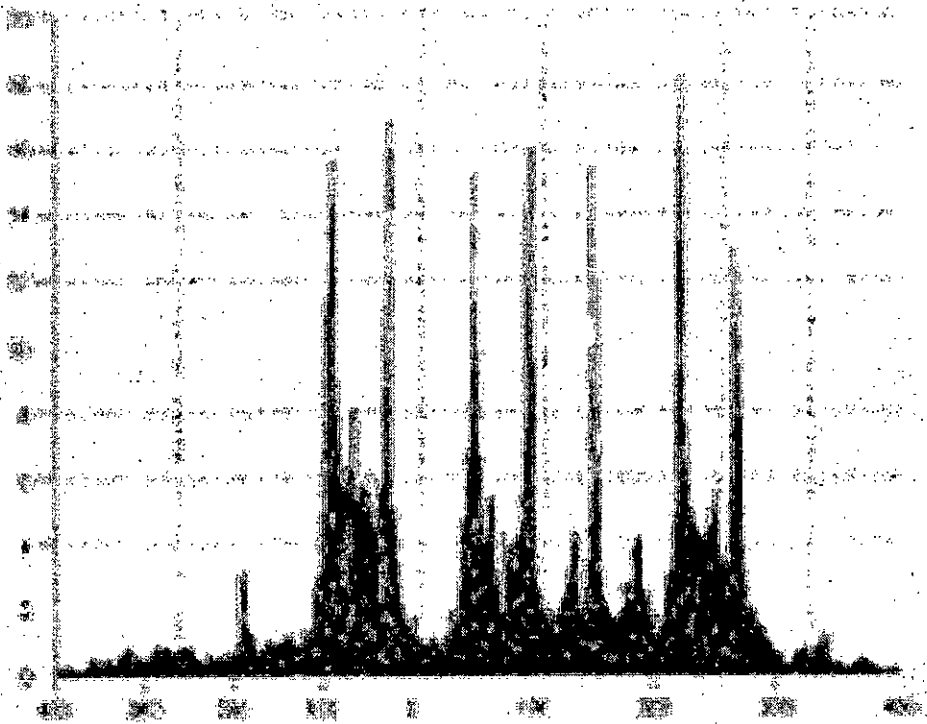


Fig. 4.14 (c): Side view of Fig. 4.14 (b)

composite reference images and they form the joint composite reference image. The images on the upper half portion represent the input scene images. Here the input contains seven 3D distorted target images and their distortion ranges mainly from azimuth change to change of elevation angle. Fig. 4.14 (b) shows the PSPFJTC output of Fig. 4.14 (a). For better visualization the side view of Fig. 4.14 (b) is shown in Fig. 4.14 (c). To get almost equal correlation height for all the target objects, we set $\alpha=0.82$ and $\beta=0.18$. The value of all the peaks are given in table 4.4 as below

Table 4.4: Detection performance for 3D distorted images

Image No	1	2	3	4	5	6	7
Peak intensity	18.8991	18.4032	18.1739	17.2327	16.9701	15.1008	14.0482

From table 4.4, the ratio of maximum target peak to minimum target peak is equal to 1.3453 and error form the exact equal peak criterion is 34.37%. Though this is a relatively high percentage error, we can accept it for pattern recognition purpose.

Next, we have placed 3D distorted target and non-target images in the input scene of joint image as shown in Fig. 4.15 (a). In this case, the input scene contains five out-of-plane rotated target images and two non-target images of almost same orientation. The lower half portion of the joint image represents two composite reference images for 3D distortion invariant target detection. These two reference images again represent two members of a class. Finally, the class-associative target detection algorithm is applied with FAF parameters as set previously, i.e. $C=1$, $D=1e-9$, $m=1$, $\alpha=0.82$ and $\beta=0.18$. The corresponding PSPFJTC output is shown in Fig. 4.15 (b). Fig. 4.15 (c) shows the side view of Fig. 4.15 (b) for clarity of peak heights. From Fig. 4.15 (b) and 4.15 (c), it is evident that there are four distinct peaks for four target objects. Again, all the peaks contain side-lobes that certainly reduce the optical efficiency but there is no deterioration of the visibility of target peaks. Table 4.5 summarizes the detection

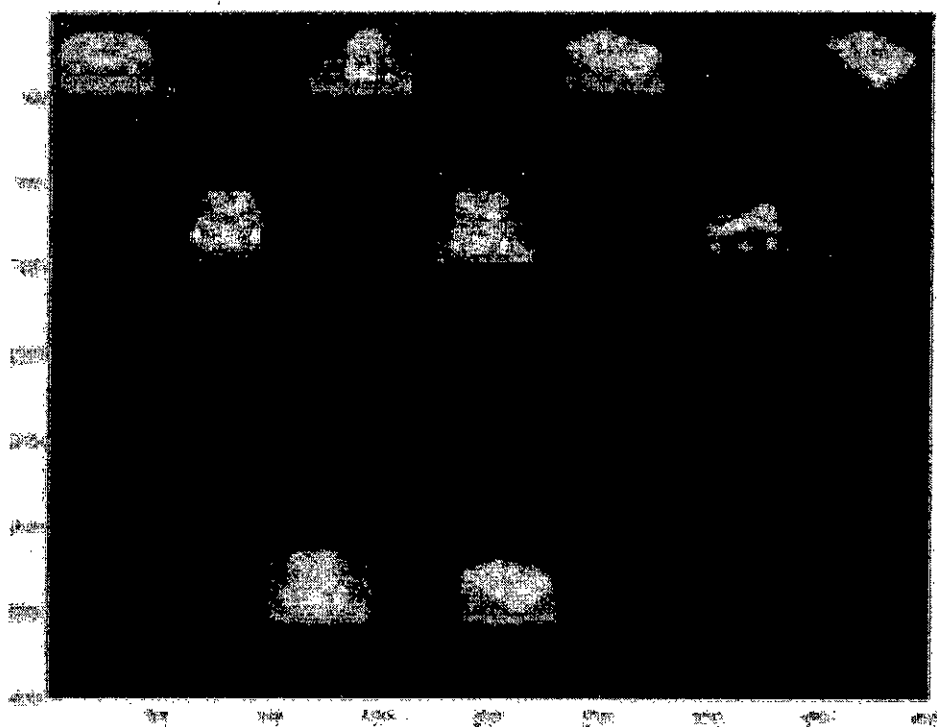


Fig. 4.15 (a): Gray-level joint image containing out-of-plane rotated target and non-target images

RM

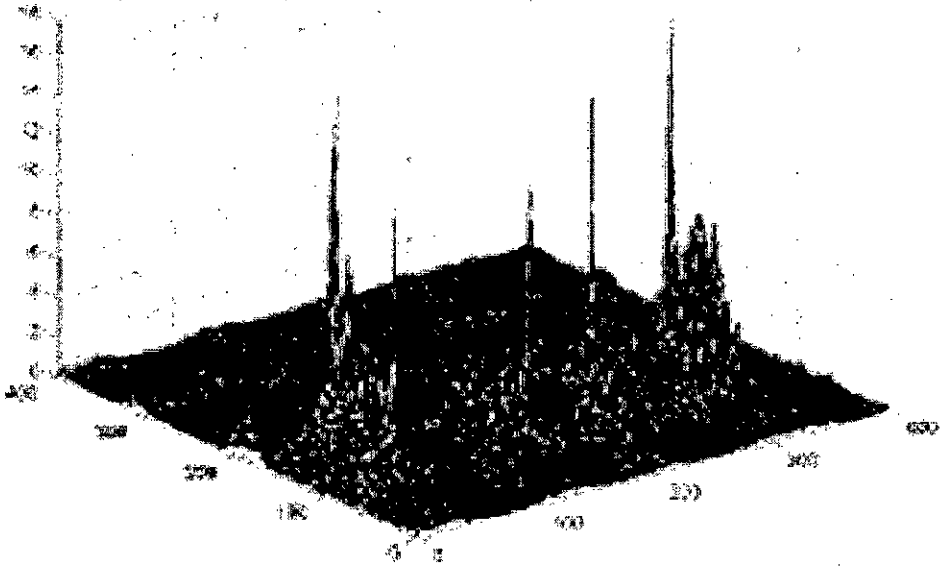


Fig. 4.15 (b): PSPFJTC output of Fig. 4.15 (a)

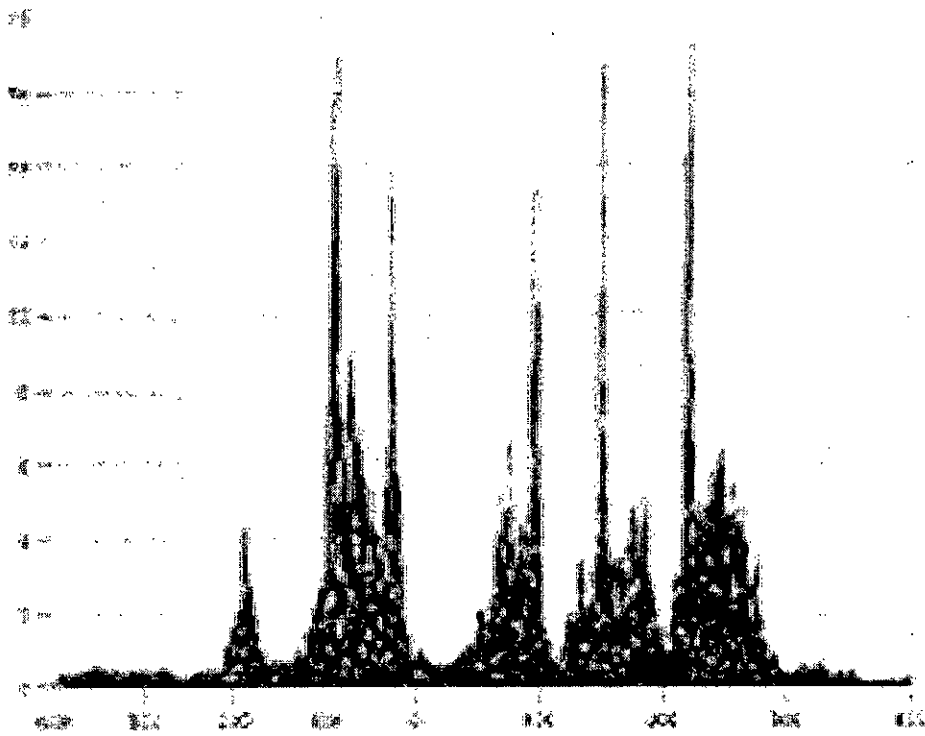


Fig. 4.15 (c): Side view of Fig. 4.15 (b)

performance in terms of ratio of maximum target peak to minimum target peak and maximum non-target peaks.

Table 4.5: Detection performance of 3D distorted objects

Maximum Target Peak (T_{\max})	Minimum Target Peak (T_{\min})	Maximum Non-target Peak (N_{\max})	$\frac{T_{\max}}{T_{\min}}$	$\frac{T_{\max}}{N_{\max}}$
17.3242	13.4567	8.0502	1.2874	2.1520

From table 4.5, we can summarize that there is appreciable discrimination between target and non-target peaks and therefore, the proposed system is feasible for target detection purposes.

Next to investigate the noise robustness of the system, additive gaussian white noise is added in the input scene. Fig. 4.16 (a) shows the joint image with an additive white noise of 5 dB in the input scene. Fig. 4.16 (b) and (c) show the corresponding PSPFJTC output of the Fig. 4.16 (a).

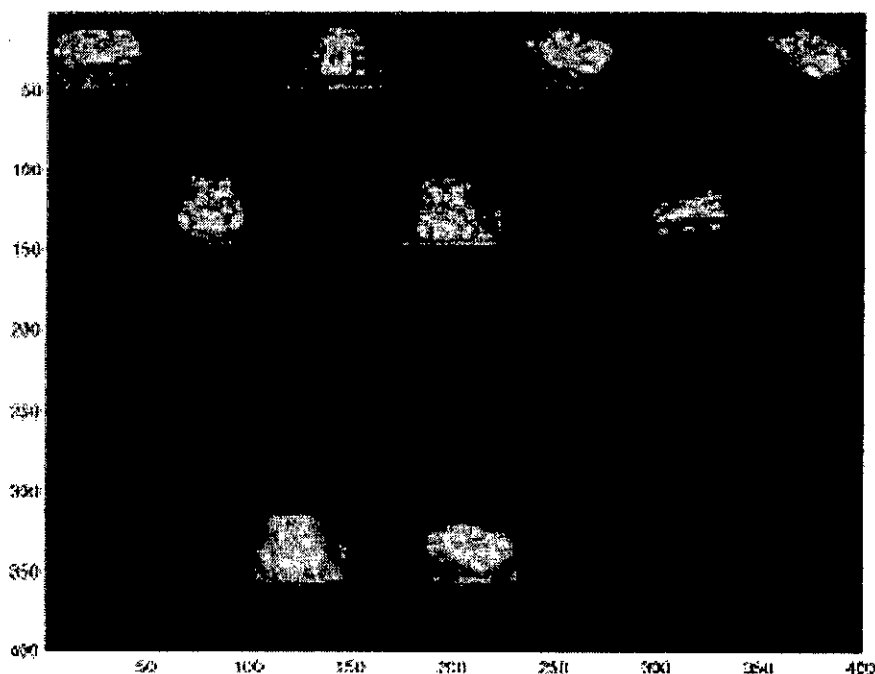


Fig. 4.16 (a): Joint image with an additive white noise of 5 dB

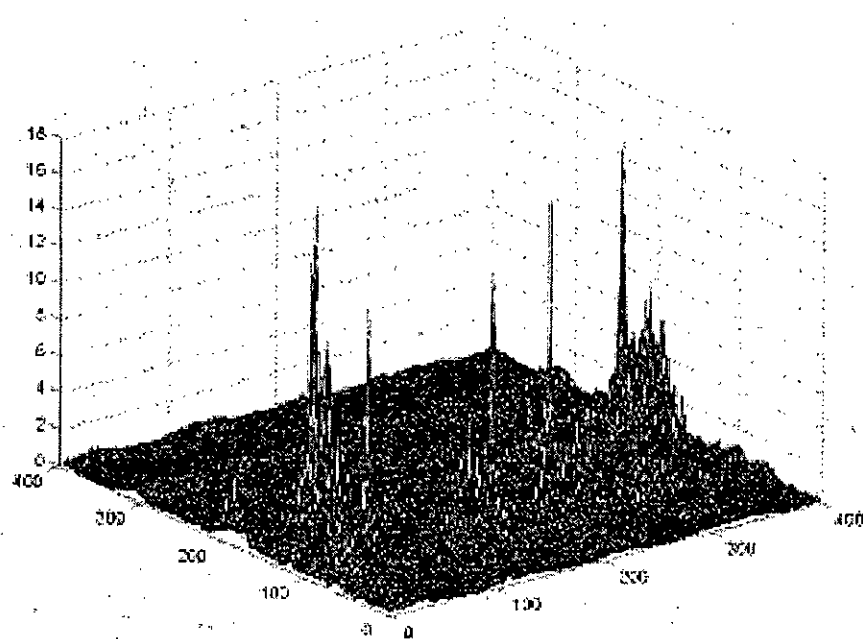


Fig. 4.16 (b): PSPFJTC output of Fig. 4.16 (a)

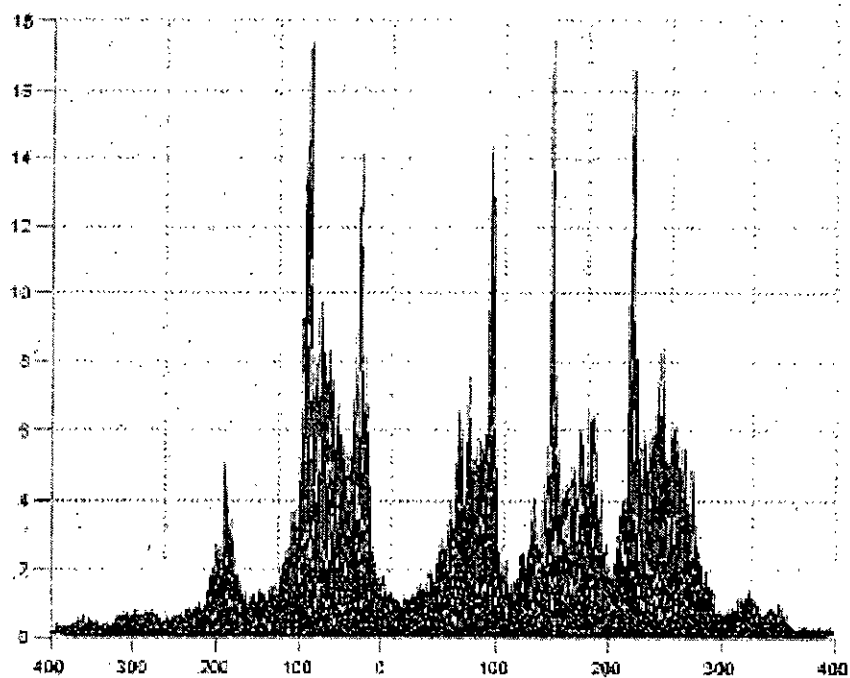


Fig. 4.16 (c): Side view of Fig. 4.16 (b).

Then the sample simulation is performed with an additive noise of 0 dB and the corresponding joint image and output are given in Fig. 4.17 (a), (b) and (c). From Fig. 4.16 (c) and Fig. 4.17 (c), it is obvious that with the addition of noise, the peak intensities are varied but do not hamper the target detection purpose so much. Table 4.6 gives the summary of detection performance of the PSPFJTC technique in noisy conditions.

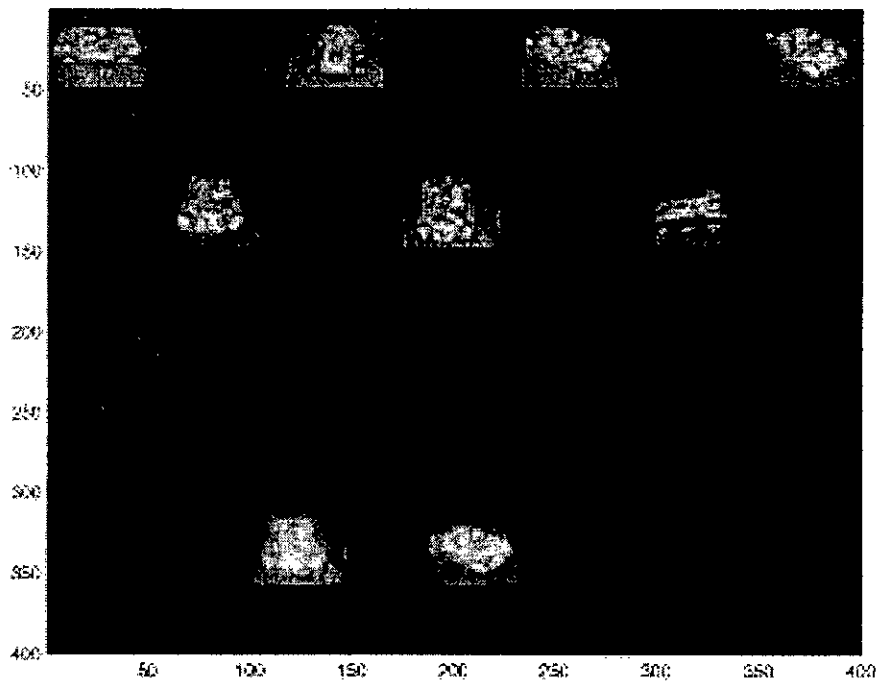


Fig. 4.17 (a): Joint image with an additive white noise of 0 dB

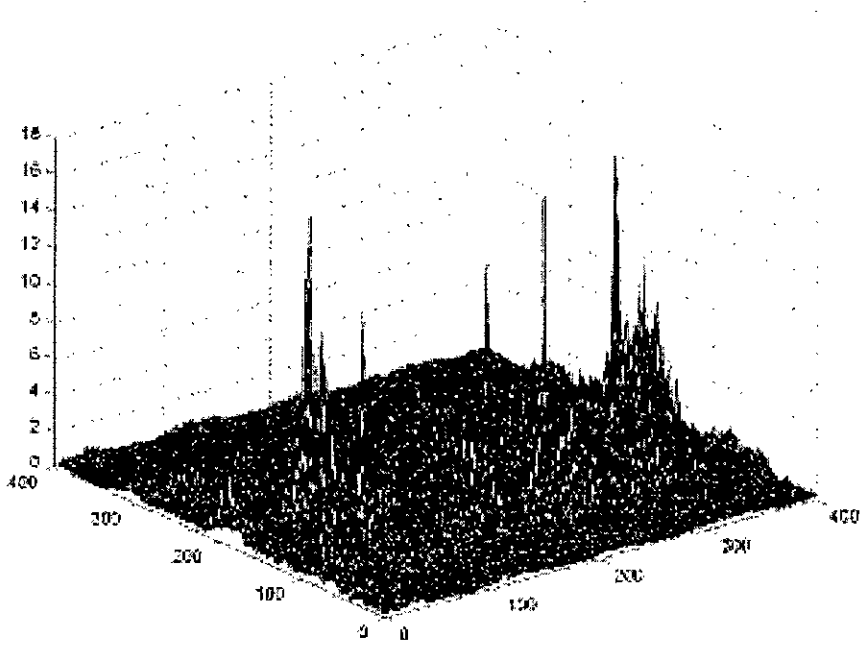


Fig. 4.17 (b): PSPFJTC output of Fig. 4.17 (a)

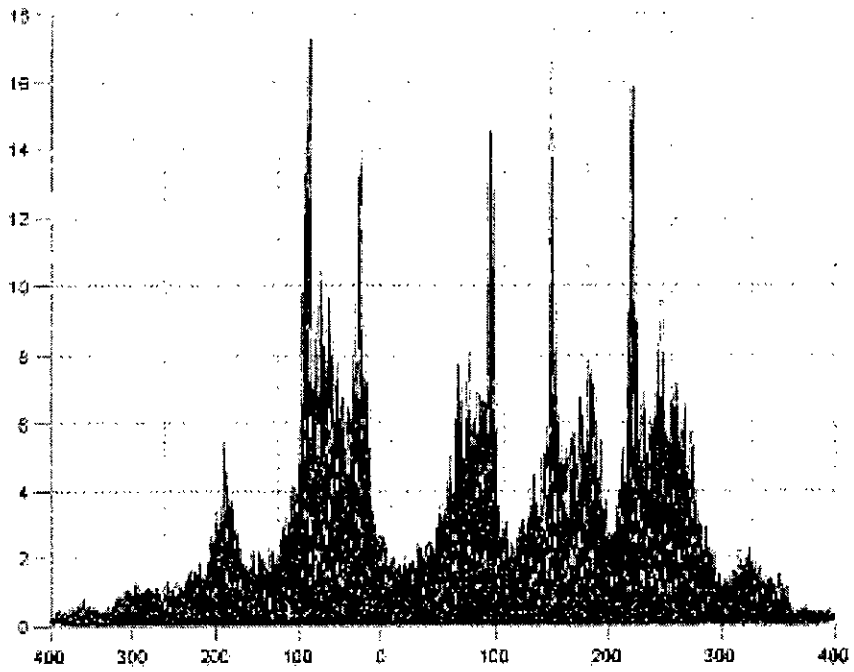


Fig. 4.17 (c): Side view of Fig. 4.17 (b)

Table 4.6: Detection performance of 3-D distorted objects in noisy conditions

Noise in dB	Max. Target peak (Tmax)	Min. Target peak (Tmin)	Max. non-target peak (Nmax)	$\frac{T_{\max}}{T_{\min}}$	$\frac{T_{\min}}{N_{\max}}$
5	17.5036	14.1366	8.3825	1.24	1.69
0	17.6687	14.0694	9.7018	1.26	1.45

From table 4.6, the discrimination ratios between target and non-target peak (i.e. ratio between minimum target peak and maximum non-target peak) are successively lowered with the addition of more and more noise in the input scene. With noise of 0 dB, the ratio becomes 1.45. Since in comparative judgment, here the target peak is at least 1.45 times higher than that of non-target peak and this is an appreciable limit for threshold. Therefore, by using suitable threshold value, the target peaks can be easily extracted in the midst of non-target peaks and other spurious signal generated due to noisy signal.

In the thesis work, a post-processing technique has been developed to increase the discrimination ratio. Here, the normalized correlation output is first squared by a CCD and is then divided by the negative exponential of the correlation output as expressed by

$$O_p(x, y) = \frac{|O(x, y)|^2}{e^{-O(x, y)}} \quad (4.13)$$

where $O(x, y)$ is the normalized actual correlation output and $O_p(x, y)$ is the output after post-processing. Fig. 4.18 (a) and (b) show the after effect of post-processing technique of the correlation outputs as shown in fig. 4.16 (c) and 4.17 (c) respectively.

Table 4.7: Detection performance of 3-D distorted objects in noisy conditions after post-processing technique

Noise in dB	Max. target peak (Tmax)	Min. Target peak (Tmin)	Max. non-target peak (Nmax)	$\frac{T_{\max}}{T_{\min}}$	$\frac{T_{\min}}{N_{\max}}$
5	2.7183	1.9839	0.9567	1.37	2.07
0	2.7183	1.9436	1.1520	1.40	1.69

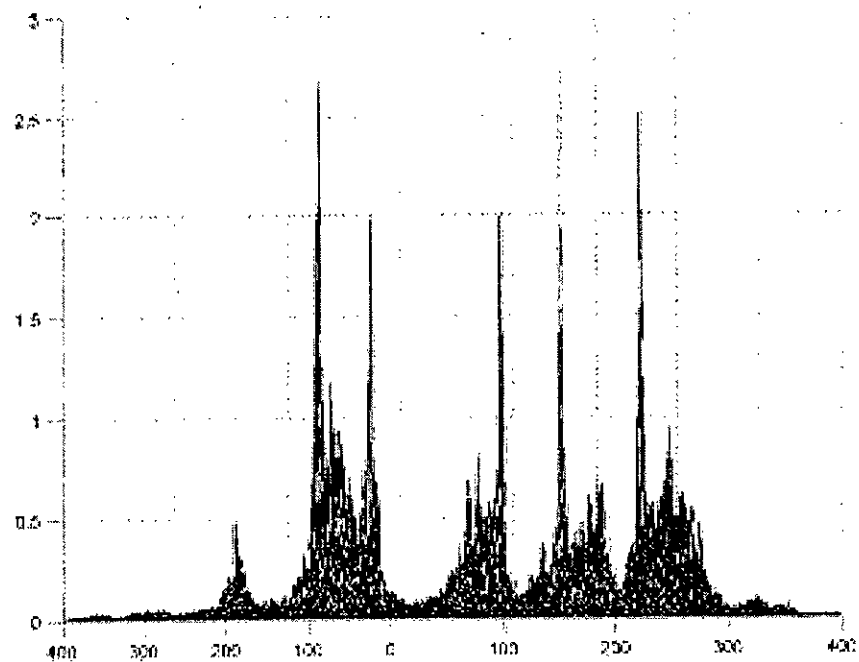


Fig. 4.18 (a): Correlation output of fig. 4.16 (c) after post-processing

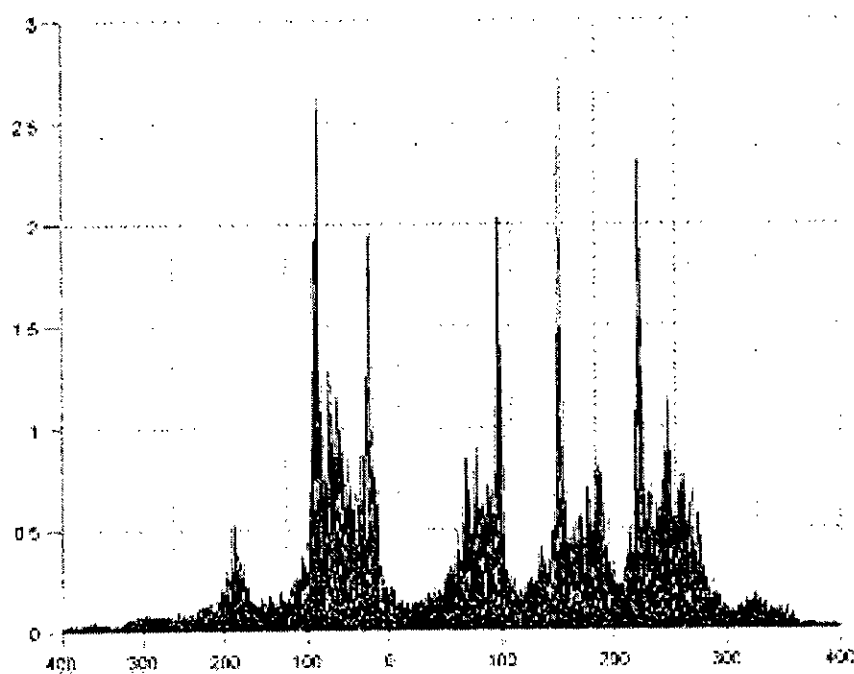


Fig. 4.18 (b): Correlation output of fig. 4.17 (c) after post-processing

This technique certainly increases the target peak intensity variation but the major advantage is that it gives much higher discrimination between target and non-target objects. Table 4.7 gives the peak intensity and corresponding ratios after the post-processing technique for figure 4.16 (c) and 4.17 (c).

4.6 Conclusion

This chapter is mainly concerned about the distortion-invariant pattern recognition using PSDF. All JTC suffers from high sensitivity to distortions such as scale variations and rotation. Again such distortions may be in-plane and out-of-plane distortions. SDF based JTC provides better distortion invariance for in-plane distortions. To achieve out-of-plane distortion invariance, projection-slice theorem is incorporated with SDF. Here the out-of-plane rotation of an object is widely divided into two category- Distortion due to azimuth change and distortion due to change of elevation angle. Then using PSDF, two composite images are formed for two different distortion types. These composite images are used as members of a class and application of efficient class-associative target detection algorithm and PSPFJTC technique, produce a sharp single correlation peak for each target object in a class. Finally, to improve the discrimination ratio between look-alike target and non-target object, a post-processing technique is proposed here. Simulation results show that the proposed system performs well in the midst of target and non-target objects even with the presence of noise.

Chapter 5

CONCLUSIONS

5.1 Conclusion

Optical correlators, especially joint transform correlators are versatile tool for pattern recognition or pattern classification. They provide real time detection of the targets with very high optical efficiency and reduced cost. Since the advent of joint transform correlators, various works have been performed on it and various modifications or improvements have been proposed. Among the various forms of joint transform correlation techniques, the fringe-adjusted joint transform correlation and recently devised multi-target detection algorithm are suitable for successful detection of objects or images when there are multiple targets and non-targets in the input scene. The later one facilitates class-associative target detection by using multiple references whereby two or more dissimilar objects representing a class can be detected with equal correlation peaks. All these techniques produce a pair of peaks for each target object. Phase-encoding with fringe-adjusted filter technique produce a single peak per target object. Again, this technique is very much sensitive to the number of objects and noise in the input scene.

In this work, a phase-shifted phase-encoded fringe-adjusted technique has been adopted for detection of multiple targets and with better utilization of space bandwidth product. In the proposed method, the phase-shifting and phase-encoding principle is applied to the reference image since this can be done prior to real time operation and therefore does not hamper the system processing speed. The inclusion of phase-shifting principle with phase-encoding principle provides the robustness of the system. Simulation results suggest that the phase-shifted phase-encoded fringe-adjusted technique is much better than phase-encoded fringe-adjusted technique in terms of fringe-adjusted filter parameters variations. The proposed method also performs well even in the midst of noise in the input scene whereas the phase-encoded fringe-adjusted technique may fail

to detect target in those cases. Furthermore, it facilitates the class-associative target detection.

For detection of a class of dissimilar objects, a new fast and efficient multiple target detection method has been proposed. In the proposed scheme, the number of processing steps is always fixed irrespective of the number of objects in the class whereas the number of processing steps is increased in the earlier methods. To achieve this goal, instead of calculating the separate power spectra from each of the reference images, a multi-reference joint image with an enhanced version of fringe-adjusted filter is used to detect multiple dissimilar target objects with almost equal correlation peaks. Inclusion of phase-shifting and phase-encoding principle with the newly proposed class-associative target detection technique provides better performance of the system with the generation of single correlation peak per target object.

However, the proposed schemes are sensitive to target distortion as all other JTC techniques. Synthetic discriminant function (SDF) and projection-slice synthetic discriminant function (PSDF) are the two tools in the areas of joint transform correlation for distortion invariant target detection. Between these two techniques, the PSDF technique is the preferred one for detection of 3-D distorted objects. In this thesis work, mainly the out-of-plane rotation of an object is considered as the 3-D distortion. For the detection of distorted objects, a composite image is formed first from a set of training or distorted images that specify the range of distortion to be detected. In PSDF technique, inclusion of higher range of distortion in a single composite image causes loss of important information. To get rid of the problem, the newly proposed class-associative target detection technique is incorporated with PSDF technique where different PSDF-based composite images are formed for different distortion range and are used as members of a class.

In this thesis work, the binary images of English block-lettered alphabets are used to show and testify the performance of the proposed schemes for in-plane rotation

distortion. For in-plane rotation invariance, the binary images are rotated by MATLAB software and up to 45° rotations invariance has been shown. For out-of-plane rotation of an object, two different gray-level car images are used. The images contain the actual out-of-plane rotation of the objects. The training set contains an azimuth change of almost 45° and a change in elevation angle of almost 90° . Simulation results show that the proposed schemes can successfully detect the 3-D distorted (mainly out-of-plane rotation) targets. Though the target peaks are not all equal, the discrimination between target and non-target peaks are so high that the inequality of target peak heights does not hamper the target detection purpose. Therefore, it can be said that the proposed scheme is a successful one though more works can be done on it for further excellence.

5.2 Future Works

In this thesis work, mainly out-of-plane rotation of a still object is considered. But in real world, an object may be moving and in that case the image of a real 3-D object contains much more distortion other than the out-of-plane rotation. The other types of distortions may be distortion of a portion of image rather than the whole image, the distortion of radial components of the image. In these cases, features extraction using wavelet transform may be useful. The feature extraction of an image using wavelet involves the convolution of the image with a dilated and translated version of suitable mother wavelet. Extracting the basic features of the training images, a composite image can be formed by the weighted sum of all these features.

Again, in this work, the target detection operation is performed on a stationary image. Here the object is assumed to be still. In case of moving objects, another important application arises that is target tracking from moving objects. In this case frame-based analysis is important. A set images or frames are taken and the object is identified from each of this frame. Finally, a trajectory of the moving object can be plotted to trace the future position of the object.

Another important work of pattern recognition is the remote sensing using radar images. In these cases thermal images are used instead of images of line-of-sight of vision. MACH (Maximum average correlation height) and MACE (Maximum average correlation energy) filters are the important types of filters used for the above purposes.

All real objects are color objects and contain more information than binary or gray level images. Therefore, to deal with color images, three basic colors i.e. RED, GREEN and BLUE have to be extracted from the actual image and would have to be processed using three channels. For distortion invariant target detection process, projection-slice SDF based formulation may be adopted on individual channels. At the output plane, the output from the three channels may be fused to get a single correlation output.

Another important application of projection-slice theorem may be the multi-input single output system design. Here the outputs from different input channels are sliced and then these slices are arranged to produce a single output.

All of these are an overview of future works in this field. Further study and detailed analysis is required in every case to implement the purpose.

REFERENCES

- [1] Robert Schalkoff, "Pattern Recognition, Statistical Structural and Neural Approaches," John Wiley & Sons, New York, 1992.
- [2] Richard O. Duada, Peter E. Hart, David G. Stork, "Pattern Classification," Second Edition, Wiley-Interscience, 2000.
- [3] B. Yegnarayana, "Artificial Neural Networks," Prentice-Hall, India, 2003.
- [4] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing," Addison-Wesley Publishing Company, Inc., 1992.
- [5] S. M. Attaullah Bhuiyan, "Distortion invariant class-associative target detection using joint transform correlation," M. Sc. Thesis, Department of EEE, BUET, 2003.
- [6] J. W. Goodman, "Introduction to Fourier Optics," McGraw Hill, New York, 1968.
- [7] Eugene Hecht and Alfred Zajac, "Optics," Addison-Wesley Publishing Company, Philippines, 1979.
- [8] A. VanderLugt, "Signal detection by complex spatial filtering," *IEEE Trans. Inf. Theory* **IT-10**, pp. 139-146, 1964.
- [9] H. J. Caulfield and R. Haimes, "Generalized matched filtering," *Appl. Opt.*, Vol. 18, No. 2, pp. 181-183, 1980.
- [10] D. A. Gregory, "Real-time pattern recognition using a modified LCTV in a coherent optical correlator," *Appl. Opt.*, Vol. 25, pp. 467-469, 1986.
- [11] A. D. Gara, "Real-time tracking of moving objects by optical correlation," *Appl. Opt.*, Vol. 18, pp. 172-172, 1989.
- [12] A. Mahalanobis and D. Casasent, "Performance evaluation of minimum average correlation filters," *Appl. Opt.*, Vol. 30, pp. 561-571, 1991.
- [13] R. R. Kallman and D. H. Goldstein, "Phase-encoding input image for optical pattern recognition," *Opt. Eng.*, Vol. 33, pp. 1806-1813, 1994.
- [14] D. Casasent and A. Furman, "Sources of correlation degeradation," *Appl. Opt.*, Vol. 16, pp. 1652-1661, 1977.
- [15] X. J. Lu, F. T. s. Yu and D. A. Gregory, "Comparison of VanderLugt and joint transform correlator," *Appl. Phys. B. Photophys. Laser Chem.*, Vol. 51, pp. 153-164, 1990.
- [16] Purwardi Purwosumarto and Francis T. S. Yu, "Robustness of joint transform correlator versus VanderLugt correlator," *Opt. Eng.*, Vol. 36, No. 10, pp. 2775-2780, 1997.
- [17] C. S. Weaver and J. W. Goodman, "A technique for optically convolving two functions," *Opt. Commun.*, Vol. 52, No. 7, pp. 1248-1249, 1966.
- [18] F. T. S. Yu and X. J. Lu, "A real-time programmable joint transform correlator," *Opt. Commun.*, Vol. 52, No. 10, pp. 10-16, 1984.
- [19] F. T. S. Yu, S. Jutamulia, T. W. Lin and D. A. Gregory, "Adaptive real-time pattern recognition using a liquid crystal TV based joint transform correlator," *Appl. Opt.*, Vol. 26, pp. 1370-1372, 1988.

- [20] B. Javidi and C. Kuo, "Joint transform image correlation using a binary spatial light modulator at the Fourier plane," *Appl. Opt.*, Vol. 27, pp. 663-665, 1988.
- [21] S. K. Rogers, J. D. Cline and M. Kabrisky, "New binarization techniques for joint transform correlation," *Opt. Eng.*, Vol. 29, No. 9, pp. 1018-1023, 1990.
- [22] B. Javidi, J. Wang and Q. Tang, "Multiple-object binary joint transform correlation using multiple-level threshold crossing," *Appl. Opt.*, Vol. 31, No. 29, pp. 4816-4822, 1992.
- [23] M. S. Alam and M. A. Karim, "Improved correlation discrimination in a multiobject bipolar joint transform correlator," *Opt. & Laser Tech*, Vol. 24, pp. 45-50, 1992.
- [24] W. B. Hahn and D. L. Flannery, "Design elements of a binary joint transform correlator and selected optimization techniques," *Opt. Eng.*, Vol. 31, No. 5, pp. 896-905, 1992.
- [25] F. T. S. Yu, F. Cheng, T. Nagata and D. A. Gregory, "Effects of fringe binarization of multi-object joint transform correlation," *Opt. Eng.*, Vol. 28, No. 5, pp. 2988-2990, 1989.
- [26] Q. Tang and B. Javidi, "Multiple-object detection with a chirp-encoded joint transform correlator," *Appl. Opt.*, Vol. 32, pp. 4344-4350, 1993.
- [27] S. Jutamulia, G. M. Storti, D. A. Gregory and J. C. Kirsch, "Illumination-independent high-efficiency joint transform correlation," *Appl. Opt.*, Vol. 30, No. 29, pp. 4173-4175, 1991.
- [28] C. Li, S. Yin and F. T. S. Yu, "Nonzero-order joint transform correlator," *Opt. Eng.*, Vol. 37, No. 1, pp. 58-65, 1998.
- [29] M. S. Alam and Y. A. Gu, "Improving multi-object correlation discrimination using data fusion in the correlation plane," *Optik*, Vol. 106, No. 1, pp. 1-6, 1997.
- [30] G. Lu, Z. Zhang, S. Wu and F. T. S. Yu, "Implementation of a non-zero order joint transform correlator by use of phase shifting techniques," *Appl. Opt.*, Vol. 36, No. 2, pp. 470-483, 1997.
- [31] S. Jutamulia and D. A. Gregory, "Soft blocking of the dc term in Fourier optical systems," *Opt. Eng.*, Vol. 37, No. 1, pp. 49-51, 1998.
- [32] F. T. S. Yu, C. Li and S. Yin, "Comparison of detection efficiency of nonzero-order and conventional joint transform correlation," *Opt. Eng.*, Vol. 37, No. 1, pp. 52-57, 1998.
- [33] M. S. Alam and M. A. Karim, "Fringe-adjusted joint transform correlator," *Appl. Opt.*, Vol. 32, No. 23, pp. 4344-4350, 1993.
- [34] M. S. Alam and M. A. Karim, "Joint transform correlation under varying illumination," *Appl. Opt.*, Vol. 32, pp. 4344-4350, 1993.
- [35] M. S. Alam, "Fractional power fringe-adjusted joint transform correlator," *Opt. Eng.* Vol. 34, pp. 3208-3216, 1995.
- [36] M. S. Alam, "Phase-encoded fringe-adjusted joint transform correlation," *Opt. Eng.*, Vol. 39, No. 5, pp. 1169-1176, 2000.

- [37] J. Khoury, G. Asimellis and C. Woods, "Incoherent-erasure nonlinear joint transform correlator," *Opt. Lett.*, Vol. 20, pp. 2321-2323, 1995.
- [38] M. S. Alam and J. Khoury, "Fringe-adjusted incoherent-erasure joint transform correlator," *Journal of Optical Engineering*, Vol. 37, pp. 75-82, 1998.
- [39] M. S. Alam and M. A. Karim, "Multiple target detection using a modified fringe-adjusted joint transform correlator," *Opt. Eng.*, Vol. 33, pp. 1610-1617, 1994.
- [40] M. S. Alam, "Multi-target photo refractive fringe-adjusted joint transform correlation," *Journal of Optical memory and Neural Networks*, Vol. 6, pp. 287-296, 1998.
- [41] A. A. S. Awwal, M. A. Karim and S. R. Jahan, "Improved correlation discrimination using an amplitude modulated phase-only filter," *Appl. Opt.*, Vol. 29, pp. 233-236, 1990.
- [42] T. Nomura, "Phase-encoded joint transform correlator to reduce the influence of extraneous signals," *Appl. Opt.*, Vol. 37, pp. 3651, 1998.
- [43] G. Lu and F. T. S. Yu, "Performance of a phase transformed input joint transform correlator," *Appl. Opt.*, Vol. 35, pp. 304, 1996.
- [44] M. K. Hu, "Visual pattern recognition by moment invariance," *IRE Trans. Inf. Theory*, IT-8, pp. 179-187, 1962.
- [45] D. Casasent and D. Psaltis, "Optical pattern recognition using normalized invariant moments," *Proc. SPIE*, Vol. 201, pp. 107-114, 1979.
- [46] Y. Li, "Reforming the theory of invariant moments of pattern recognition," *Pattern Recog.*, Vol. 25, No. 7, pp. 723-730, 1992.
- [47] A. Ghostasby, "Template matching in rotated images," *IEEE Trans. Pattern Anal. Mach. Intel.*, PAMI, Vol. 7, No. 3, pp. 338-344, 1985.
- [48] Y. Sheng and H. H. Arsenault, "Object detection from a real scene using the correlation peak coordinates of a multiple circular harmonic filters," *Appl. Opt.*, Vol. 28, No. 2, pp. 245-249, 1989.
- [49] M. T. Manry, "Compound filter using the circular harmonic expansion," *Appl. Opt.*, Vol. 26, No. 17, pp. 3622-3627, 1987.
- [50] D. Mendlovic, E. Moran and N. Konforti, "Shift and scale invariant pattern recognition using Mellin radial harmonics," *Opt. Commun.*, Vol. 67, No. 3, pp.172-176, 1988.
- [51] Y. Sheng, C. Lejeune and H. H. Arsenault, "Frequency-domain Fourier Mellin descriptors for invariant pattern recognition," *Opt. Eng.*, Vol. 27, No. 5, pp. 354-357, 1988.
- [52] M. Fang and G. Hausler, "Class of transforms invariant under shift, rotation and scaling," *Appl. Opt.*, Vol. 29, No. 5, pp. 704-708, 1990.
- [53] W. S. Shao and Y. S. Chen, "Pattern analysis on shift, rotation and scaling," *Electron Lett.*, Vol. 28, No. 25, pp. 2271-2272, 1992.
- [54] D. Casasent, "Unified synthetic discriminant function computational formulation," *Appl. Opt.*, Vol. 23, pp. 1620-1627, 1984.

- [55] B. V. K. V. Kumar, "Minimum variance synthetic discriminant function," *J. Opt. Soc. Am. A*, 3, pp. 1579-1584, 1986.
- [56] A. Mahalanobis, B. V. K. V. Kumar and D. Casasent, "Minimum average correlation energy filters," *Appl. Opt.*, Vol. 26, pp. 3633-3640, 1987.
- [57] D. Casasent and W. T. Chang, "Correlation synthetic discriminant functions," *Appl. Opt.*, Vol. 25, pp. 2343-2350, 1986.
- [58] Z. Bahri and B. V. K. Kumar, "Generalized synthetic discriminant functions," *J. Opt. Soc. Am. A*, 5, pp. 562-571, 1988.
- [59] B. Javidi, "Synthetic discriminant function based binary nonlinear optical correlator," *Appl. Opt.*, Vol. 28, pp. 2490-2493, 1989.
- [60] R. Wang, C. R. Chatwin and L. Shang, "Synthetic discriminant function fringe-adjusted joint transform correlator," *Opt. Eng.*, Vol. 34, No. 10, pp. 2935-2944, 1995.
- [61] X. W. Chen, M. A. Karim and M. S. Alam, "Distortion-invariant fractional power fringe-adjusted joint transform correlation," *Opt. Eng.*, Vol. 37, No. 1, pp. 138-143, 1998.
- [62] V. R. Riasati and M. Abushagur, "Projection-slice synthetic discriminant functions for optical pattern recognition," *Appl. Opt.*, Vol. 36, pp. 3022-2034, 1997.
- [63] D. Casasent, J. Smokelin and A. Ye, "Wavelet and Gabor transforms for detection," *Opt. Eng.*, Vol. 31, pp. 1893-1898, 1992.
- [64] F. Ahmed, M. S. Alam and M. A. Karim, "Wavelet transform based correlator for the recognition of rotationally distorted images," *Opt. Eng.*, Vol. 34, pp. 3187-3192, 1995.
- [65] B. Javidi, J. Li and Q. Tang, "Optical implementation of neural networks for face recognition by use of nonlinear joint transform correlator," *Appl. Opt.*, Vol. 34, pp. 3950-3962, 1995.
- [66] A. R. Alsamman and M. S. Alam, "Invariant face recognition using fringe-adjusted joint transform correlator based neural network," in *Optical Pattern Recognition XII*, D. P. Casasent and T. H. Chao, Eds., Proc. SPIE, Vol. 4387, pp. 164-175, 2001.
- [67] D. Casasent and B. Tefler, "Key and recollection vector effect on heteroassociative memory performance," *Appl. Opt.*, Vol. 28, pp. 272-283, 1989.
- [68] J. Khoury, P. D. Gianino and Charles L. Woods, "Class associative correlation filter using cross correlation enhancement," *Optical Pattern Recognition XI*, D. P. Casasent and T. H. Chao, eds., Proc. SPIE, Vol. 4043, pp. 7456-7463, 2000.
- [69] M. S. Alam and M. M. Rahman, "Class associative multiple target detection by use of fringe-adjusted joint transform correlation," *Appl. Opt.*, Vol. 41, No. 35, pp. 1-8, 2002.

