

BANDPASS FILTERS FOR UNCONSTRAINED TARGET RECOGNITION AND THEIR IMPLEMENTATION IN COHERENT OPTICAL CORRELATORS

A Thesis
submitted to the
Faculty of Engineering
of the University of Glasgow
for the Degree of
Doctor of Philosophy
by
Rupert Charles David Young

Dedicated to Dorothy V Young for many things but in particular taking me to see the Engine Hall of the Science Museum, London, in 1965.

There is a mask of theory over the whole face of nature.

William Whewell

The Philosophy of the Inductive Sciences

Summary

The design of bandpass filters and their assessment through both simulation and experiment is studied, with particular emphasis on a coherent optical implementation. Their performance in an industrial component recognition task and for the orientation independent recognition of a motor vehicle is investigated. The origin of current filter designs from the original work on the matched filter is discussed, together with essential differences arising from the application of the filter to image data, in contrast to the original radar signal processing task. Emphasis is placed on the fundamental requirement of striking a balance between the discriminatory ability of the filter against inter-class objects and background clutter on the one hand, while on the other the requirement for tolerance to intra-class variations such as orientation changes of a target object. The conclusion is reached that a bandpass filter is required, low spatial frequencies being removed to enhance discrimination, while high spatial frequencies are attenuated to avoid excessive sensitivity.

An up-dateable correlator is simulated which is based on the non-degenerate four wave mixing (NDFWM) interaction in the photorefractive material bismuth silicon oxide ($\text{Bi}_{12}\text{SiO}_{20}$). Specifically, it is shown that variable bandpass filters can be implemented directly in the correlator by adjusting the relative strengths of the signal and reference beams used to write the Fourier transform hologram into the photorefractive. The synthetic discriminant function (SDF) method of grey-level multiplexing is reviewed. A bandpass modification of this technique is used in the design of a multiplexed filter for the recognition of an industrial test component from a limited number of known stable state orientations when viewed from an overhead camera position. Its performance in this task when implemented in the up-dateable correlator is assessed through simulation. The conclusion of this work is that filter multiplexing must be used judiciously for orientation invariant recognition. Only a limited number of images, typically under ten, may be multiplexed into each filter since correlation peak heights and peak-to-sidelobe ratios inevitably progressively deteriorate as images are added to the filter.

The effect of severe amplitude disruptions in the frequency plane on correlation peak localisation is examined. In two or higher dimensions simulations show the localisation is very robust to this disruption; an analysis is developed to indicate the reason for this. The effect is exploited by the implementation of an algorithm that locally removes the spatial frequencies that exhibit close phase matching between intra- and inter-class images. The inter-class response can be forced to zero while simultaneously improving the intra-class tolerance to orientation changes. The technique is assessed through simulation with

images of two types of motor vehicle, in a variety of orientations, and shown to be effective in improving discrimination and intra-class tolerance for examples in which these were initially very poor.

Bandpass filters are experimentally implemented in a joint transform correlator (JTC) based on a NDFWM interaction in $\text{Bi}_{12}\text{SiO}_{20}$. The JTC is described and its full bandwidth performance initially assessed. As anticipated from the previous considerations, inter-class discrimination was high but the intra-class tolerance very poor due to the high sensitivity of the filter. The difference of Gaussian approximation to a Laplacian of a Gaussian filter is described and its experimental implementation in the JTC detailed. Experimental results are presented for the orientation independent recognition of a car while maintaining discrimination against another car. An intra-class to inter-class correlation ratio of 7.5 dB was obtained as a best case and 3.6 dB as a worst case, the intra-class variation being at $11\frac{1}{4}^\circ$ increments in orientation at zero elevation angle. The results are extrapolated to estimate that approximately 80 filters would be required for a full 2π steradian orientation coverage. The implementation of the frequency removal technique and the Wiener filter in the JTC is briefly considered in conclusion to this work.

To maintain a correlation response in situations in which the orientation of the target object is liable to abrupt change it is thus necessary to perform a template search at a rate of at least two orders of magnitude faster than the input frame rate. To accomplish this an optical/digital hybrid correlator system is proposed which provides the potential for a multi-kilohertz reference template search rate on image data acquired at video rates. This is achieved by storing the reference templates as angle multiplexed volume holograms in a fixable photorefractive material. By exploiting the large storage capacity and capability for rapid optical read-out that this affords, many correlation searches per input frame may be performed. The initial Fourier transform is calculated digitally, the phase portion being applied to an electrically addressed phase modulating spatial light modulator. During an interval in which the display is held stable the reference templates are reconstructed and optically correlated with the current input scene. Energy losses through the system are critical to the obtainable template search rate. However, it is estimated that with system components currently available, a 3 kHz search rate can be achieved from a volume holographic memory containing 1000 templates.

Acknowledgements

I would like to thank the following for the financial support given that allowed this work to be undertaken :

Dorothy V Young for one year of study;

Isabella W Phemister via a Scholarship endowed to the memory of Andrew G Phemister for two years of study.

I would like to thank the following for help given :

Dr Christopher R Chatwin for his supervision of the research, his friendship and honesty;

Irene Tait for typing the bulk of the manuscript;

Dr Hussein Abdullah for help with micro-computing;

Yasmine Mather for assistance with SUN workstations;

Isabelle Lawson for drafting diagrams;

Riu Kang Wang for assistance with computer plots;

The many technicians who helped with the experimental work, particularly Mr Bob MacCletchy for his skilled construction of optical mounts of the required rigidity.

I would also like to thank the Head of Department, Professor Brian F Scott, for his interest in the research and support of the experimental work through scarce resources from the Departmental equipment fund.

List of Contents

Summary	i
Acknowledgements	iii
Figures and Tables	vii
Notation	xii
Abbreviations	xviii
Introduction	1
References to Introduction	5
Chapter 1	
Development of the bandpass filter from the matched filter	6
1.1 Historical development	6
1.2 Variations from the matched filter	15
1.2.1 The inverse filter	15
1.2.2 Reduction of filter sensitivity	18
1.2.3 The phase only filter	19
1.2.4 The Wiener filter	22
References to Chapter 1	27
Chapter 2	
Design and simulation of a bandpass synthetic discriminant function	29
2.1 Introduction	29
2.2 The up-dateable correlator	31
2.2.1 System configuration	31
2.2.2 Variation of response with beam ratio	34
2.3 The synthetic discriminant function filter	45
2.3.1 Background	45
2.3.2 SDF filter design considerations	51
2.4 Simulation results of bandpass SDF	61
2.5 Conclusion	65
References to Chapter 2	67

Chapter 3

Discrete frequency removal technique and its application to filter performance enhancement	69
3.1 Introduction	69
3.2 Robustness of correlation peak localisation to frequency plane amplitude discontinuities	71
3.2.1 Demonstration by simulation	71
3.2.2 Analysis of correlation localisation by consideration of diffraction by a random set of delta functionals	83
3.2.3 Analysis of correlation peak localisation using the central-slice theorem	86
3.3 Use of discrete frequency removal to enhance the correlation filter performance	93
3.3.1 Simulation of the joint transform correlator	93
3.3.2 Joint transform correlator response to the test inputs	103
3.3.3 Application of the discrete frequency removal technique	107
3.4 Conclusions	152
References to Chapter 3	156

Chapter 4

Implementation and experimental assessment of bandpass filters	157
4.1 Introduction	157
4.2 Joint transform correlator experimental implementation	157
4.3 Joint transform correlator full bandwidth performance	175
4.4 Joint transform correlator bandpass performance	184
4.4.1 The difference of Gaussian filter	184
4.4.2 Experimental implementation of the difference of Gaussian filter in the joint transform correlator	192
4.5 Intra-class and inter-class performance of the bandpass joint transform correlator	207
4.6 Conclusions	244
References to Chapter 4	247

Chapter 5

A proposed hybrid correlator system and its performance assessment	248
5.1 Introduction	248
5.2 Digital Fourier transform	249
5.3 Phase modulating spatial light modulator	253
5.4 Volume hologram template memory	255
5.5 Beam energy losses through the system	256
5.5.1 Source power and component losses	256
5.5.2 Energy losses due to frequency plane filtering	257
5.6 Correlation plane detection and read-out	259
5.7 Filter multiplexing	261
5.8 Conclusion	262
References to Chapter 5	263

Figures and Tables

Chapter 1

Figure 1.1	Schematic of frequency plane regions used in defining the modified inverse filter	17
------------	---	----

Chapter 2

Figure 2.1	Schematic of up-dateable photorefractive correlator	32
Figure 2.2	Cross-sections through edge reconstructions for the beam ratio values indicated	37
Figure 2.3	Auto-correlation functions for TC3, view 3 for the indicated beam ratios	41
Figure 2.4	Auto-correlation of TC3, view 3	41
Figure 2.5	Auto-correlation peak intensity variation with beam ratio	42
Figure 2.6	Correlation plane energy variation with beam ratio	42
Table 2.1	Summary of auto-correlation properties for indicated beam ratios	43
Table 2.2	Summary of correlation peak fall-off with rotation angle for indicated beam ratios	43
Figure 2.7	Fall in correlation with rotation of TC3, view 3 for the indicated beam ratio values	44
Figure 2.8	Stable state views of TC3 and one view of TC6	46
Figure 2.9	Field amplitudes, $U_4(\mathbf{k})$, directly behind filter plane	57
Table 2.3	Cross-correlation maxima (amplitude) for TC3 stable state views	58
Table 2.4	Cross-correlation peak locations	58
Table 2.5	On-axis contributions of un-aligned cross-correlations	59
Figure 2.10	Cross-correlation of TC3, view 3 and TC3, view 1 normalised to the ACF of TC3, view 1	59
Figure 2.11	SDF space domain function	60
Table 2.6	Calculated SDF coefficients for TC3	62
Table 2.7	SDF correlation responses to training set images – TC3 stable state views	62
Figure 2.12	SDF correlation response to two members of the training set	63
Figure 2.13	SDF correlation response to TC6	64

Chapter 3

Figure 3.1	Comparison of the auto-correlation functions generated by a matched spatial filter and phase only filter	70
Figure 3.2	The discrete approximation to a δ -functional and its discrete Fourier transform	73
Figure 3.3	Cross-sections through random data array with threshold set at one standard deviation	76
Figure 3.4	Cross-sections through Fourier transform of random data array shown in Figure 3.3	77
Figure 3.5	Cross-sections through intensity values of the simulated correlation plane	78
Figure 3.6	Cross-sections through maximum noise spike of the simulated correlation plane real field values	79
Figure 3.7	Isometric of random data array with threshold set at two standard deviations	80
Figure 3.8	Cross-sections through Fourier transform of random data array shown in Figure 3.7	81
Figure 3.9	Depiction of the central-slice theorem with notations	88
Figure 3.10	The Radon transform	89
Figure 3.11	Application of the central-slice theorem to a random frequency domain function	92
Figure 3.12	Example views of the Escort and Mazda cars used as test imagery in both simulation and experimental work	95
Figure 3.13	Auto-correlation function generated by the joint transform correlator	100
Figure 3.14	Frequency plane Gaussian apodisation and resulting auto-correlation function	101
Figure 3.15	Cross-correlation performance of joint transform correlator for the indicated car orientations	105
Figure 3.16	Intra-class cross-correlation of E000 and E349	110
Figure 3.17	Inter-class cross-correlation of E000 and M000	112
Figure 3.18	Intra-class cross-correlation of E000 and E349 after discrete frequency removal	114
Figure 3.19	Inter-class cross-correlation of E000 and M000 after discrete frequency removal	116
Figure 3.20	Auto-correlation function of E000 after discrete frequency removal	119

Figure 3.21	Plot of the relative intra-class and inter-class cross-correlation magnitudes of E000 with E349 and M000 as discrete frequency components are removed	122
Figure 3.22	Intra-class cross-correlation of E180 and E191	124
Figure 3.23	Inter-class cross-correlation of E180 and M180	125
Figure 3.24	Intra-class cross-correlation of E180 and E191 after discrete frequency removal	128
Figure 3.25	Inter-class cross-correlation of E180 and M180 after discrete frequency removal	130
Figure 3.26	Plot of the relative intra-class and inter-class cross-correlation magnitudes of E180 with E191 and M180 as discrete frequency components are removed	133
Figure 3.27	Plot of the relative intra-class and inter-class cross-correlation magnitudes of E270 with E281 and M270 as discrete frequency components are removed	133
Figure 3.28	Inter-class cross-correlation of E270 and M270	134
Figure 3.29	Inter-class cross-correlation of E270 and M270 after discrete frequency removal	135
Figure 3.30	Field components directly behind modified E000/E349/M000 filter plane with E000 as input	137
Figure 3.31	Field components directly behind modified E180/E191/M180 filter plane with E180 as input	140
Figure 3.32	Field components directly behind unmodified E000 filter with M000 as input	141
Figure 3.33	Field components directly behind unmodified E000 filter with E349 as input	144
Figure 3.34	Field components directly behind modified E000/E349/M000 filter with M000 as input	145
Figure 3.35	Field components directly behind modified E000/E349/M000 filter with E349 as input	146
Figure 3.36	Field components directly behind unmodified E180 filter with M180 as input	148
Figure 3.37	Field components directly behind unmodified E180 filter with E191 as input	149
Figure 3.38	Field components directly behind modified E180/E191/M180 filter with M180 as input	150
Figure 3.39	Field components directly behind modified E180/E191/M180 filter with E191 as input	151

Chapter 4

Figure 4.1	Schematic of photorefractive based joint transform up-dateable correlator	158
Figure 4.2	Experimental implementation of joint transform correlator	159
Table 4.1	Summary of equipment used in the experimental implementation of the joint transform correlator	160
Figure 4.3	Examples of the rigid optical mounts employed in the joint transform correlator	163
Figure 4.4	Example views of car test imagery	170
Figure 4.5	Power spectra of the indicated car orientations	172
Figure 4.6	Power spectrum of Escort sideview at various exposures	174
Figure 4.7	Full bandwidth auto-correlation responses of the joint transform correlator	178
Figure 4.8	Full bandwidth intra-class and inter-class response of the joint transform correlator	180
Table 4.2	Full bandwidth joint transform correlator response to in-plane rotation of car test imagery	182
Figure 4.9	Fall in full bandwidth correlation with in-plane rotation	183
Figure 4.10	Response of the difference of Gaussian filter to a step change in intensity	185
Figure 4.11	The difference of Gaussian function and its spectrum	187
Figure 4.12	The effect on edge localisation produced by changing the size of the convolving difference of Gaussian kernel	188
Figure 4.13	The amplitude transmittance-exposure curve for films of different photographic γ	194
Figure 4.14	Experimentally generated difference of Gaussian approximation	196
Figure 4.15	Impulse response of difference of Gaussian filter as implemented in the joint transform correlator	197
Figure 4.16	Image reconstructions of difference of Gaussian filtered car test imagery	202
Figure 4.17	Under-exposed image reconstructions to demonstrate energy localisation effect of difference of Gaussian filter	205
Figure 4.18	Auto-correlation response of bandpass joint transform correlator	209
Figure 4.19	Auto-correlation and cross-correlation responses of bandpass JTC to E090 and M090	210
Figure 4.20	Auto-correlation and cross-correlation responses of bandlimited and bandpass JTC to E180 and M180	214

Figure 4.21	Auto-correlation and cross-correlation responses of bandlimited and bandpass JTC to E000 and M000	216
Figure 4.22	Auto-correlation and cross-correlation responses of bandlimited and bandpass JTC to E045 and M045	218
Figure 4.23	Isometric views of auto-correlation and cross-correlation responses of bandlimited and bandpass JTC to E045 and M045	220
Figure 4.24	Auto-correlation and cross-correlation responses of bandlimited and bandpass JTC to E135 and M135	222
Figure 4.25	Auto-correlation and cross-correlation responses of bandlimited and bandpass JTC to E011 and M011	224
Figure 4.26	Auto-correlation and cross-correlation responses of bandlimited and bandpass JTC to E079 and M079	226
Figure 4.27	Auto-correlation and cross-correlation responses of bandlimited and bandpass JTC to E124 and M124	228
Table 4.3	Quantitative summary of bandlimited and bandpass JTC inter-class cross-correlation responses	230
Figure 4.28	Auto-correlation and cross-correlation responses of bandpass JTC to orientation changes around E090	231
Figure 4.29	Auto-correlation and cross-correlation responses of bandpass JTC to orientation changes around E180	233
Figure 4.30	Auto-correlation and cross-correlation responses of bandpass JTC to orientation changes around E000	234
Figure 4.31	Auto-correlation and cross-correlation responses of bandpass JTC to orientation changes around E045	235
Figure 4.32	Auto-correlation and cross-correlation responses of bandpass JTC to orientation changes around E135	238
Table 4.4	Quantitative summary of bandpass JTC intra-class cross-correlation responses	240
 Chapter 5		
Figure 5.1	Schematic of the hybrid optical/digital correlator	250
Figure 5.2	Layout of the two-dimensional fast Fourier transform processor	251

Notation

	Space domain		Frequency domain	
Scalar constants	$a, b...$		$A, B...$	
Scalar functions	$g(\mathbf{r}), h(\mathbf{r})...$		$G(\mathbf{k}), H(\mathbf{k})...$	
Vectors	$\mathbf{g}, \mathbf{h}...$		$\mathbf{G}, \mathbf{H}...$	
Matrices	$h, d...$		$H, D...$	
Unit vectors	(\mathbf{i}, \mathbf{j})		(\mathbf{i}, \mathbf{j})	
Vector variables	$\mathbf{r} \equiv (i\mathbf{x} + j\mathbf{y})$ eg. $h(\mathbf{r})$		$\mathbf{k} \equiv (ik_x + jk_y)$ eg. $H(\mathbf{k})$	
Concatonated image vectors (of dimension d^2 derived from an image array (d,d))	$\mathbf{h}(\mathbf{r})$		$\mathbf{H}(\mathbf{k})$	
Array subscripts	(i,j)		(i,j)	
	Continuous	Discrete	Continuous	Discrete
Filter transfer function	$g(\mathbf{r})$	$\mathbf{g}(\mathbf{r})$	$G(\mathbf{k})$	$\mathbf{G}(\mathbf{k})$
Approximation to $\hat{f}(\mathbf{r}), f(\mathbf{r})$	$\hat{f}(\mathbf{r})$	$\hat{\mathbf{f}}(\mathbf{r})$	$\hat{F}(\mathbf{k})$	$\hat{\mathbf{F}}(\mathbf{k})$
Wiener filter transfer function	$w(\mathbf{r})$	$\mathbf{w}(\mathbf{r})$	$W(\mathbf{k})$	$\mathbf{W}(\mathbf{k})$
Input functions to a system :				
Signal and noise	$f(\mathbf{r})$	$\mathbf{f}(\mathbf{r})$	$F(\mathbf{k})$	$\mathbf{F}(\mathbf{k})$
Signal	$h(\mathbf{r})$	$\mathbf{h}(\mathbf{r})$	$H(\mathbf{k})$	$\mathbf{H}(\mathbf{k})$
System output functions	$s(\mathbf{r})$	$\mathbf{s}(\mathbf{r})$	$S(\mathbf{k})$	$\mathbf{S}(\mathbf{k})$

System point spread function	$p(\mathbf{r})$	$\mathbf{p}(\mathbf{r})$	$\mathbf{P}(\mathbf{k})$	$\mathbf{P}(\mathbf{k})$
Noise (clutter)	$n(\mathbf{r})$	$\mathbf{n}(\mathbf{r})$	$\mathbf{N}(\mathbf{k})$	$\mathbf{N}(\mathbf{k})$

Abbreviated mathematical notations

Fourier transform operator $F\{\dots\}$

Inverse Fourier transform operator $F^{-1}\{\dots\}$

Transpose of a matrix $[\dots]^T$

Complex conjugate of a matrix $[\dots]^*$

Complex conjugate transpose of a matrix $[\dots]^+$

Convolution
ie.

$$s(\mathbf{r}) = \int \int_{-\infty}^{\infty} h(\boldsymbol{\rho}) g(\mathbf{r} - \boldsymbol{\rho}) d\rho_x d\rho_y$$

$$s(\mathbf{r}) = h(\mathbf{r}) * g(\mathbf{r})$$

Correlation
ie.

$$s(\mathbf{r}) = \int \int_{-\infty}^{\infty} h^*(\boldsymbol{\rho}) g(\boldsymbol{\rho} + \mathbf{r}) d\rho_x d\rho_y$$

$$s(\mathbf{r}) = h^*(\mathbf{r}) \otimes g(\mathbf{r})$$
or

$$s(\mathbf{r}) = \int \int_{-\infty}^{\infty} h(\boldsymbol{\rho}) g^*(\boldsymbol{\rho} - \mathbf{r}) d\rho_x d\rho_y$$

$$s(\mathbf{r}) = h(\mathbf{r}) * g^*(-\mathbf{r})$$

Vector inner product $\mathbf{h}(\mathbf{r}) \cdot \mathbf{g}(\mathbf{r})$
ie.

$$c(o) = \sum_{\mathbf{r}}^{\mathbf{d}^2} \mathbf{h}(\mathbf{r}) \mathbf{g}(\mathbf{r})$$

Linear, shift-invariant system :

$$\begin{array}{ll} \text{Input} & f(\mathbf{r}) = h(\mathbf{r}) + n(\mathbf{r}) \\ \text{Output} & s(\mathbf{r}) = y(\mathbf{r}) + n_g(\mathbf{r}) \end{array}$$

where :

$$n_g(\mathbf{r}) = g(\mathbf{r}) * n(\mathbf{r})$$

and

$$y(\mathbf{r}) = g(\mathbf{r}) * h(\mathbf{r})$$

$$\text{Auto-correlation of } h(\mathbf{r}) \quad R_{hh}$$

$$\text{Cross-correlation of } h(\mathbf{r}) \text{ and } g(\mathbf{r}) \quad R_{hg}$$

$$\text{Power spectrum of } h(\mathbf{r}) \quad S_{hh}$$

(Power spectral density function) ie.

$$S_{hh} = \int \int_{-\infty}^{\infty} R_{hh} \exp(j2\pi \mathbf{k} \cdot \mathbf{r}) d\mathbf{x} d\mathbf{y}$$

(Wiener-Khintchine theorem)

$$\begin{array}{ll} \text{Noise power spectral} & \\ \text{density function} & S_{nn} \end{array}$$

$$\begin{array}{ll} \text{Correlation plane on-axis} & \\ \text{value ie. vector inner product} & c(\mathbf{o}), c(\mathbf{o}) \end{array}$$

$$\begin{array}{ll} \text{On-axis system output} & s(\mathbf{o}), s(\mathbf{o}) \end{array}$$

Statistical notation

$$\langle h(\mathbf{r}) \rangle = \lim_{N \rightarrow \infty} \int \int_{\frac{N}{2}}^{\frac{N}{2}} h(\mathbf{r}) d\mathbf{x} d\mathbf{y}$$

$$\langle h(\mathbf{r}) \rangle = \frac{1}{d^2} [h(r_1) + h(r_2) + \dots + h(r_{d^2})]$$

Alternatively, written as :

$$E[h(\mathbf{r})], \overline{h(\mathbf{r})}, \mu_h, \text{ mean value or expected value.}$$

$$\langle h^2(\mathbf{r}) \rangle = \lim_{N \rightarrow \infty} \int \int_{\frac{N}{2}}^{\frac{N}{2}} h^2(\mathbf{r}) dx dy$$

$$\langle h^2(\mathbf{r}) \rangle = \frac{1}{d^2} [h^2(\mathbf{r}_1) + h^2(\mathbf{r}_2) + \dots + h^2(\mathbf{r}_{d^2})]$$

Alternatively, written as :

$\overline{h^2(\mathbf{r})}$, σ_h^2 , mean square value or variance.

Also,

$$\left(\langle h^2(\mathbf{r}) \rangle \right)^{\frac{1}{2}} = \text{root mean square}$$

and

$$\begin{aligned} \left\langle \left(h(\mathbf{r}) - \overline{h(\mathbf{r})} \right)^2 \right\rangle &= \lim_{N \rightarrow \infty} \int \int_{\frac{N}{2}}^{\frac{N}{2}} \left(h(\mathbf{r}) - \overline{h(\mathbf{r})} \right)^2 dx dy \\ &= E \left[\left(h(\mathbf{r}) - \overline{h(\mathbf{r})} \right)^2 \right] \end{aligned}$$

$$\|h(\mathbf{r})\| = \left[\int \int_{-\infty}^{\infty} |h(\mathbf{r})| dx dy \right]^{\frac{1}{2}}$$

Norm of $h(\mathbf{r})$

$$\langle h(\mathbf{r}), g(\mathbf{r}) \rangle = \lim_{N \rightarrow \infty} \int \int_{\frac{N}{2}}^{\frac{N}{2}} h(\mathbf{r}) g(\mathbf{r}) dx dy$$

$$\langle h(\mathbf{r}), g(\mathbf{r}) \rangle = \frac{1}{d^2} [h(\mathbf{r}_1)g(\mathbf{r}_1) + h(\mathbf{r}_2)g(\mathbf{r}_2) + \dots + h(\mathbf{r}_{d^2})g(\mathbf{r}_{d^2})]$$

Vector inner product of $h(\mathbf{r})$ and $g(\mathbf{r})$

$$\langle h^2(\mathbf{r}) \rangle = \int \int_{-\infty}^{\infty} |H(\mathbf{k})|^2 dk_x dk_y$$

(Parseval's Theorem)

Commonly used criteria in assessing correlation filter performance

Image plane energy
$$IPE = \int \int_{-\infty}^{\infty} |f(\mathbf{r})|^2 dx dy$$

Correlation plane energy
$$CPE = \int \int_{-\infty}^{\infty} |s(\mathbf{r})|^2 dx dy$$

Correlation peak to CPE ratio
$$PCE = \frac{|s(\mathbf{o})|^2}{CPE}$$

Matched filter SNR criterion
$$SNR_{MSF} = \frac{|s(\mathbf{o})|^2}{\sigma_n^2}$$

Equivalently,
$$SNR_{MSF} = \frac{\left| \int \int_{-\infty}^{\infty} G(\mathbf{k}) H(\mathbf{k}) dk_x dk_y \right|^2}{\int \int_{-\infty}^{\infty} |G(\mathbf{k})|^2 |N(\mathbf{k})|^2 dk_x dk_y}$$

Original Horner efficiency
$$\eta_H = \frac{\int \int_{-\infty}^{\infty} |s(\mathbf{r})|^2 dx dy}{\int \int_{-\infty}^{\infty} |f(\mathbf{r})|^2 dx dy}$$

Equivalently,
$$\eta_H = \frac{CPE}{IPE}$$

Modified Horner efficiency
$$\eta_{H_m} = \frac{|s(\mathbf{o})|^2}{\int \int_{-\infty}^{\infty} |f(\mathbf{r})|^2 dx dy}$$

Kallman's SNR criterion
$$SNR(g(\mathbf{r})) = \frac{\left(\min_{1 \leq i \leq l} \max_{\mathbf{r} \in B_i} |\langle g(\mathbf{r}), h_i(\mathbf{r}) \rangle|^2 \right)}{\left(\min_{1 \leq i \leq l} \max_{\mathbf{r} \in B_i} |\langle g(\mathbf{r}), h_i(\mathbf{r}) \rangle|^2 \right)}$$

where,

$h_1, \dots, h_p, \dots, h_m$ are training set images

h_1, \dots, h_l are in-class objects

h_{l+1}, \dots, h_m are out-of-class objects

B_i is a box of chosen size around the origin

$\left| \langle g(\mathbf{r}), h_i(\mathbf{r}) \rangle \right|^2 \quad (1 \leq i \leq l)$ are intensities in the output plane
(large inside B_i and small outside B_i
for an effective filter)

$\left| \langle g(\mathbf{r}), h_i(\mathbf{r}) \rangle \right|^2 \quad (l+1 \leq i \leq m)$ are intensities in the output plane
(small inside B_i and small outside B_i
for an effective filter)

The SNR as defined above is a measure of the smallest true signal given by $g(\mathbf{r})$ to the largest false signal generated by $g(\mathbf{r})$. As such it is similar to the signal-to-clutter ratio (SCR) used in radar. The name peak-to-secondary ratio (PSR) has been suggested for this measure.

Since Kallman's definition of SNR does not have a functional form, it cannot be used in analytical optimisation procedures as can the CPE and PCE measures. However, it is the *only* intrinsically useful measure of filter performance for real object data which, by its very nature, cannot be specified analytically.

Abbreviations

ACF	Auto-correlation function
CCF	Cross-correlation function
A/O	Acousto-optic
Ar ⁺	Argon ion (laser)
HeNe	Helium Neon (laser)
BSO	Bismuth Silicon oxide (Bi ₁₂ SiO ₂₀)
BR	Beam ratio
CAD	Computer aided drafting
CCD	Charge coupled device
CCIR	European video standard
RS170	American video standard
CPU	Central processor unit
n-D	n-dimensional
DFWM	Degenerate four wave mixing
NDFWM	Non-degenerate four wave mixing
DOG	Difference of Gaussian
LOG	Laplacian of Gaussian
DSP	Digital signal processing
η	efficiency
Exxx	Escort car orientation
Mxxx	Mazda car orientation
FT	Fourier transform
FFT	Fast Fourier transform
FWHM	Full width at half maximum
γ	Photographic gamma (slope of Hurter-Driffield curve)
JTC	Joint transform correlator
LCD	Liquid crystal display
MACE	Minimum average correlation energy (filter)
SDF	Synthetic discriminant function (filter)
MVSDF	Minimum variance SDF
MSF	Matched spatial filter
MTF	Modulation transfer function
OTF	Optical transfer function
NAG	Numerical algorithms group
PAN	Pan-chromatic (photographic film)
PC	Personal computer

POF	Phase only filter
SLM	Spatial light modulator
SNR	Signal-to-noise ratio (see text for specific definitions)
TEM	Transverse electric mode
TCxx	Industrial test component

In addition, standard abbreviations have been used throughout the text with which the reader will be familiar such as, for example, kHz for kilo-Hertz etc.

Introduction

The thesis is concerned with the design of bandpass filters and their assessment through both simulation and experiment, specifically for implementation in coherent optical correlator systems. Their application to unconstrained, or limited constraint, problems is examined for two different tasks: recognition of an industrial component, from an overhead camera position, known *a priori* to be resting in a number of well defined stable states; also the orientation independent recognition of a motor vehicle while maintaining discrimination against a similar vehicle.

In Chapter 1, the origin of the ideas currently employed in the design of filters for coherent optical correlators in the work of North [1] on the matched filter is discussed. In its application by Vander Lugt [2] to image recognition in a coherent optical implementation the background scene to the object to be identified is treated as a noise term. However, for the filter to be effective the power spectral density of the noise must be included in its design. Unfortunately, this is not readily achieved in an optical implementation which has given rise to the inappropriate use of white noise models in much of the subsequent filter modelling work. This has led to poor results in the discrimination ability of such filters. The matched filter is shown to be at the low sensitivity end of a range of filters. At the other extreme is the inverse filter which maximizes the discriminatory ability of the filter but gives poor tolerance to distortions of the in-class object. The filter sensitivity is dependent on the relative weighting given to high spatial frequencies as compared to the low frequencies in the filter transfer function. For the tasks being examined a bandpass filter is shown to be appropriate. Thus high frequencies are removed since they lead to over-sensitivity to intra-class distortions, while low frequencies are attenuated to improve inter-class discrimination. The difference of Gaussian filter is alluded to as a means of implementing such a bandpass filter; it is further discussed and applied later in the thesis. Finally in Chapter 1, the Wiener filter is described since this provides local spatial frequency control dependent on the relative power of the object and background spectra. Again careful scrutiny of the statistical properties of the background scene, treated as the noise source, has to be made to avoid incorrect assumptions when dealing with data derived from images.

Chapter 2 examines through simulation the implementation of bandpass filters in an up-dateable correlator based on a non-degenerate four wave mixing (NDFWM) interaction in the photorefractive material bismuth silicon oxide ($\text{Bi}_{12}\text{SiO}_{20}$). The filter is shown to be tuneable by adjusting the strength of the plane wave reference beam used in the formation of the Fourier transform hologram recorded in the photorefractive. A filter

response similar to that of the matched spatial filter can be realised by employing a relatively strong reference beam. Attenuating its amplitude by 32 times generates a filter biased towards high spatial frequencies which is shown to produce a localised correlation response highly sensitive to variations of the input image, the full width at half maximum (FWHM) of the peak being over six times reduced from that of the matched spatial filter response. Examination of the filter impulse response demonstrates that it effects a second order differential operation on the input image intensity distribution.

The synthetic discriminant function (SDF) method of filter grey-level multiplexing is reviewed, the minimum variance (MV)SDF being identified with the matched spatial filter and the minimum average correlation energy (MACE) filter with the inverse filter; they are thus seen to be at the extremes of the continuum described in Chapter 1. The difficulties in the practical implementation of these filters is discussed. A bandpass SDF is suggested that strikes a compromise between the broad correlation peak produced by MVSDf and the ultra-high sensitivity of the MACE filter. Appropriate shifting of the training set images to maximise the correlation peak localisation is discussed and a sequential design procedure given for the filter's implementation in the photorefractive based correlator, although this is also readily adaptable to other correlator configurations. A multiplexed filter is produced of the seven stable state orientations of a test industrial component (TC3) and its performance assessed through simulation. The FWHM of the correlation response of the filter was only marginally broader (1.38 times) than the auto-correlation response, suggesting that peak localisation can still be maintained with filters biased to considerably lower frequencies than the MACE filter. The correlation peak height, however, was 38% of the lowest auto-correlation response of the TC3 views contained in the filter and only 17% of that of the maximum. This, together with the unaligned cross-correlation terms, resulted in a deterioration of the peak-to-sidelobe ratio of the SDF correlation plane in comparison to the response of the unmultiplexed filter. This fall in correlation response results from the sharing of the available dynamic range of the input transducer. The discrimination of the filter against a similarly shaped test component (TC6) was found to be good in one orientation, the cross-correlation response being 11% of the SDF auto-correlation peak. When rotated by 90° the cross-correlation rose to 23% due to the large coincidence in edge locations between TC3 and TC6 in this orientation. Discrimination, however, was achieved without including TC6 in the filter training set, thereby placing no further demands on the dynamic range of the input spatial light modulator. The conclusion of this work is that, if multiplexed filters are to be used, the number of input images comprising the filter should be relatively small, of the order of ten or less. This ensures that the filter may be tuned to a mid-band range of spatial frequencies to give maximum intra-class tolerance while

maintaining inter-class discrimination with a reasonable peak localisation and peak-to-sidelobe ratio.

Chapter 3 examines the effect of severe amplitude disruptions in the frequency plane on the correlation peak localisation. In two (or higher) dimensions, simulations show that this localisation is very robust against severe frequency plane amplitude perturbation. This effect is analysed by analogy with results obtained from x-ray crystallography and also by employing the central slice theorem. The robustness is exploited by the implementation of an algorithm that locally measures the relative phase matching between the spectrum of the reference image and those of the intra-class and inter-class images. The discrete frequency locations are zeroed in descending order of mismatch until the inter-class cross-correlation response is forced to zero, the intra-class response being simultaneously increased. The resulting highly discontinuous frequency domain spectrum is shown to yield localised correlation peaks with very minimal sidelobe disruption. The technique is applied to example orientations of the car images, simulation results demonstrating its effectiveness at suppressing inter-class cross-correlation while enhancing the intra-class response. Quantitative measures are provided for these examples but are, of course, data dependent.

In Chapter 4 bandpass filters are experimentally implemented in a joint transform correlator based on a NDFWM interaction in $\text{Bi}_{12}\text{SiO}_{20}$. The joint transform correlator is described and its full bandwidth performance examined using the car test input images. Sharp, well defined auto-correlations are obtained and inter-class discrimination is excellent. However, the intra-class response is shown to collapse, resulting in the full bandwidth filter being of little use for the orientation independent recognition of the vehicle. The response is shown to be similar to that of an inverse filter.

The difference of Gaussian approximation of a Laplacian of a Gaussian filter is described with emphasis on its spectral bandpass filtering effect. Experimental means of implementing this filter in the joint transform correlator by modulation of the helium neon read-out beam are described. Experimental results are presented for the orientation independent recognition of one car while maintaining discrimination against another. An intra-class to inter-class correlation ratio of 7.5 dB was obtained as a best case and 3.6 dB as a worst case. The results are object specific but should be extendible to similar object types such as other vehicles. $11\frac{1}{4}^\circ$ increments of orientation at zero elevation angle were used to generate the intra-class responses. It is estimated by extrapolation of these results that approximately 80 filters would be required for a 2π steradian coverage, allowing for some correlation degradation due to mismatches of in-plane rotation and

scale. Multiplexing is suggested to accommodate in-plane rotation rather than orientation changes, separate multiple filters being preferable for the latter. In those applications in which scale is not pre-constrained, scale matching should be accomplished by a supplementary technique such as laser ranging and compensation with a zoom lens system. In conclusion to this Chapter the possibility of implementation of the frequency removal technique and the Wiener filter in the joint transform correlator by appropriate modulation of the read-out beam is briefly discussed.

The work of the previous chapters has indicated that of the order of 100 filters must be searched to acquire a correlation on an unconstrained object. To maintain the correlation response in situations in which the orientation of the object is liable to abrupt change it is thus necessary to be able to perform a template search at least two orders of magnitude faster than the input frame rate. Thus, finally, in Chapter 5 a hybrid optical/digital configuration to implement such a high speed correlator is proposed. The input data is digitally Fourier transformed at video rates utilising specialised digital signal processing chip sets (GEC/Plessey PDSP16510) and the computed phase data passed to an electrically addressed phase modulating spatial light modulator. Successive templates, which are stored as angle multiplexed volume holograms in a photorefractive crystal, are reconstructed and optically correlated with the current input scene during an interval in which the spatial light modulator display is held stable. In this way, shift-invariance, one of the major advantages of the correlation technique, is preserved. Of critical importance to the speed of operation are the energy losses of the coherent wavefront through the system and the rate at which the correlation plane data can be converted to an electronic signal; both of these issues are considered in order to estimate that at least a 3 kHz search rate is achievable from a volume holographic memory containing 1000 templates with system components currently available.

It should be noted that the research has not been presented in chronological order. The work described in Chapter 4 preceded the simulation studies presented in Chapter 3. The design and simulation of the synthetic discriminant function was performed subsequently to the work described in Chapters 3 and 4. Chapter 5 details the most recent system design considerations.

References to the Introduction

- [1] D. North, " An analysis of the factors which determine signal/noise discrimination in pulsed-carrier systems", *RCA Tech. Rept. PTR-6C*, June 25, 1943 (*ATI 14009*). Reprinted in *Proc. IEEE*, Vol. 51, p.1016, (1963).
- [2] A. Vander Lugt, "Signal detection by complex spatial filtering", *IEEE Trans.*, IT-10, p.139, (1964).

Chapter 1

Development of the bandpass filter from the matched filter

1.1 Historical Development

The design of the frequency domain filter used in a 4-f coherent optical correlator system is of paramount importance to its performance in a given pattern recognition task. It is, therefore, valuable to briefly trace the historical development of the ideas that have been used in the filter design for linear invariant systems since much of the current research is based (often inappropriately) on these developments.

The original work was due to North [1] and concerned the optimal detection of a radar return from a target. In a pulsed radar system a waveform of a specified shape is emitted, propagates towards a target and is reflected; the scattered return signal is then detected at the receiver. The time delay between transmission and detection gives a measure of the target's range. In addition, the Doppler shift can be used to estimate the radial velocity of the target with respect to the receiver. The scattered signal to be detected is weak and contaminated with noise. The filter devised by North, known as the matched filter, optimises the signal to noise power ratio at the filter output. The signal output at time t_0 is given by :

$$y(t_0) = \int_{-\infty}^{\infty} G(k)H(k)\exp(j2\pi kt_0)dk \quad \{1.1\}$$

(t_0 is chosen so that the output signal is causal with respect to the input)
and the mean square value of the noise by :

$$\langle (n(t))^2 \rangle = \overline{|n(t)|^2} = \int_{-\infty}^{\infty} |G(k)|^2 |N(k)|^2 dk \quad \{1.2\}$$

(Note: $|N(k)|^2 = S_n(k)$)

The important assumption of Equation {1.2} is that the noise is uncorrelated with the input signal ie. the Fourier transforms of $S_{hn}(k)$ and $S_{nh}(k)$ are taken to be zero. This is only the case if the noise (and/or the signal) is of zero mean. This is a valid assumption for the scattered return signal, due to its random phase modulation, and also for system noise encountered in the receiving electronics. However, it does not apply to the image clutter present at the input to an optical correlator; this point will be discussed in detail below. No assumption is made about the probability density function of the noise ie. it need not be Gaussian.

The output signal to noise power ratio is defined as :

$$SNR = \frac{|y(t_0)|^2}{|n(t)|^2} = \frac{|E[y(t_0)]|^2}{\text{var}[n(t)]} = \frac{|c(t_0)|^2}{\sigma^2} \quad \{1.3\}$$

or equivalently in the (temporal) frequency domain:

$$SNR = \frac{\left| \int_{-\infty}^{\infty} G(k)H(k) \exp(j2\pi kt_0) dk \right|^2}{\int_{-\infty}^{\infty} |G(k)|^2 |N(k)|^2 dk} \quad \{1.4\}$$

(Note: $\langle n^2(t) \rangle = \int_{-\infty}^{\infty} |N(k)|^2 dk$ by Parseval's theorem)

Note also that the SNR is not dependent of the signal shape but only on the relative energies of signal and noise.

North derived the form of the transfer function $G(k)$ to give the maximum SNR as defined above using the calculus of variations, a useful outline proof being given by Horn [2]. A simpler proof, employed in most textbooks [3,4 and 5] is to use the Schwarz inequality :

$$\int |P(k)|^2 dk * \int |Q(k)|^2 dk \geq \left| \int P(k)Q(k)dk \right|^2 \quad \{1.5\}$$

to derive the same result.

Applying the Schwarz inequality to Equation {1.4} yields:

$$\text{SNR}_{\max} = \int_{-\infty}^{\infty} \frac{|H(k)|^2}{|N(k)|^2} dk \quad \{1.6\}$$

To obtain this maximum $G(k)$ is, therefore, of the form [3,4 and 5] :

$$G(k) = \exp(-j2\pi t_0) \frac{H^*(k)}{|N(k)|^2} \quad \{1.7\}$$

The result is intuitively appealing since the maximum attenuation occurs in the filter transfer function at those frequencies where the signal magnitude is much less than that of the noise. Conversely, where the signal magnitude is high compared to that of the noise, the filter transmission is at a maximum. The filter will provide the best improvement in the SNR at the output when the spectra of the signal and noise are very dissimilar. Further, the filter transfer function may be written :

$$G(k) = \frac{1}{N(k)} \left[\frac{H(k)}{N(k)} \right]^* \quad \{1.8\}$$

It can, therefore, be considered as a cascade of a pre-whitening filter, that acts to flatten the input noise spectrum, and a filter that selectively passes the signal spectrum dependent on the signal to noise ratio.

The matched filter as written in Equation {1.7} assumes $|N(k)|^2$ is known or at least that a recorded sample realisation gives an adequate approximation of the noise at any given time. Very often the noise can be assumed to be white in which case Equation {1.7} reduces to the simple form :

$$G(k) = \exp(-j2\pi t_0) H^*(k) \quad \{1.9\}$$

(In practice the system will be bandlimited and so the noise need only be uniform within this bandpass ie. so called pink noise.)

Thus when the radar return signal is contaminated with stationary, additive white noise of zero mean, Equation {1.9} provides the optimum detection filter (in the sense of maximising the SNR defined in Equation {1.3}).

Now since,

$$H^*(k) = F\{h^*(-t)\} \quad \{1.10\}$$

the frequency domain filter is equivalent to the space domain operation:

$$s(t) = \int_{-\infty}^{\infty} f(\tau) h^*(\tau - t) d\tau \quad \{1.11\}$$

(normally $h(t)$ is real so $h^*(t) = h(t)$)

ie. the correlation between the transmitted and return waveform which may be written :

$$s(t) = \int_{-\infty}^{\infty} (h(\tau) + n(\tau)) h^*(\tau - t) d\tau \quad \{1.12\}$$

or

$$R_{\text{th}} = R_{\text{hh}} + R_{\text{nh}} \quad \{1.13\}$$

and if the noise and signal are uncorrelated the matched filter output is the autocorrelation of the pulse waveform. If the return waveform is delayed by t_0 , the correlation function will peak at t_0 with the SNR being a maximum at this instant of time.

In addition, the matched filter has other useful statistical properties. It provides a maximum likelihood ratio when statistical decision theory is applied to deciding if a target is present or not at a given location. The correlation peak provides the maximum probability of detecting the target for a fixed probability of false alarm when it is not present. These properties are contingent, however, on the contaminated noise having the statistical properties assumed in deriving the filter ie. the noise is zero mean, stationary, additive and white, as emphasised earlier.

When the target has a radial velocity with respect to the receiver the return signal will be Doppler shifted ie. the signal spectrum will be displaced along the frequency axis, and the correlation signal will be reduced in magnitude or lost completely. An acceptable correlation peak degradation must be determined and a bank of matched filters employed,

each of which is capable of detecting the target position within a range of Doppler shifts, the entire bank covering the full bandwidth of expected Doppler shifts. The maximally responding filter then indicates the Doppler shift and hence radial velocity of any particular target within the uncertainty determined by its bandwidth tolerance. Where very large Doppler shifts are encountered, for instance in CO₂ laser radar systems operating at 10.6 μm wavelength in which the return signal may be frequency shifted by tens of MHz, the multiple filter approach becomes impracticable. Evaluation of the frequency/time delay ambiguity function as originally suggested by Woodward becomes an attractive alternative for such systems although its real-time evaluation is computationally intensive. This function may be written :

$$A(t, k) = \int_{-\infty}^{\infty} h(\tau) h^*(t + \tau) \exp(j2\pi k\tau) d\tau \quad \{1.14\}$$

which evaluates the auto-correlation of $h(t)$ along the time axis and the spectrum of $h(t)$ along the frequency axis. It has the property that $|A(t, k)|^2$ peaks at $t=t_0$ and $k=k_0$ to give an estimate of target range and radial velocity for a return signal delayed by t_0 and Doppler shifted by k_0 . Sadler *et al.* [6] have recently designed and demonstrated an acousto-optical system for ambiguity function evaluation capable of real time operation.

Vander Lugt [7] devised and implemented a two dimensional spatial version of the matched filter. This well known solution involved holographically recording the complex transfer function, $H^*(k)$, on a carrier frequency generated by the interference of the spectrum of an input scene and a mutually coherent plane wave. The input scene was originally a photographic transparency containing the 2-D spatial reference signal $h(r)$. An expanded laser beam is used to coherently illuminate the transparency and is amplitude modulated by it. A converging lens forms the spectrum of this amplitude distribution in its back focal plane which is interfered with the reference wave, the pattern being recorded by a high resolution holographic emulsion. If such a filter is linearly recorded by the emulsion, that is the resulting transmission of the developed plate is linear with respect to the exposure, the function $H^*(k)$ will be faithfully recorded [8]. The 4f Vander Lugt optical system will then act as a correlator, the output plane light distribution representing the 2-D correlation between $h(r)$ and the current input distribution. The latter is typically $h(r)$ together with clutter. The correlator has the important property of shift invariance, meaning $h(r)$ may be located at any position within the clutter without affecting the correlation strength. Further, the peak location corresponds to the position of $h(r)$ within $f(r)$. The property of extracting a signal and its location from a cluttered background, independent of a pixel by pixel analysis of the entire input image, is the most attractive feature of the correlation method when used for

2-D pattern recognition. However, such a filter will not work particularly well in practice. The reason for this, mentioned briefly earlier, is that the background clutter does not have the statistical properties used in deriving the optimality conditions of the matched filter.

The clutter is not white noise and, since it and the reference signal result from an amplitude modulation of the illuminating coherent wavefront, neither are of zero mean. This results in the reference and clutter being strongly correlated in many instances. In Vander Lugt's original paper [7] he suggests including an additional filter behind the MSF with a transfer function which is the reciprocal of the background clutter spectral density, ie. $1/N(\mathbf{k})$, as in the general form of the matched filter of Equation {1.7}. However, his experimental realisation of the filter (and the overwhelming majority of subsequent descriptions, for instance, Goodman [8], take $|N(\mathbf{k})|^2$ to be constant ie. the spectral density of the clutter to be flat or "white". The main reason for this appears to have been the difficulty in implementing the reciprocal noise spectral density filter.

The inverse filter may be realised photographically by recording the power spectrum of the input scene noise and developing as a negative with a γ_n of unity. The intensity transmittance of the film is then inversely related to the incident intensity (ie. the noise power spectrum) over the linear region of the Hurter-Driffield curve for the emulsion [8]. If this filter is placed directly behind the complex matched filter, an approximation of the desired transfer function will be realised. However, since the filter is a passive device, the noise power spectrum can only be whitened by attenuation of its non-uniformity. Thus a level, K , must be set to which all spectral components are attenuated. Regions where the spectral magnitude falls below K cannot be accentuated by the passive filter and thus lead to a non-uniformity of the transmitted spectrum.

There are several practical difficulties which resulted in the inverse filter never being generally implemented in combination with the conjugate phase filter in optical correlators. The holographic portion of the filter will diffract at most 6.25% of the incident light intensity into the 1st diffraction order since it is recorded as a thin amplitude hologram. The inverse filter will then further attenuate the spectrum resulting in a very low overall light transmission. However, there are other, more serious, difficulties.

The purpose of the correlator is to identify a relatively localised object in background clutter. The object is considered the signal, to which the filter is matched, and the background scene is considered as noise. However, for any given unknown input it is not possible to separate the object and background power spectra *a priori*. Further, as

just described, the inverse filter must be fabricated by a photographic process which would preclude the real-time operation of the correlator with an input scene imaged on to a spatial light modulator. It is for these reasons that the inverse filter has not been implemented and which has led, ultimately, to false assumptions about the spectral characteristics of the input "noise" to the correlator. It must be pointed out, however, that optical correlators based on holographic recording of the matched spatial filter in silver halide worked well, in the sense that their discrimination against non-target objects was generally good. The reason for this was that the filter recorded was not a matched filter in the strict sense as defined by Equation {1.9}. The fine holographic emulsion, necessary to record the high carrier frequencies associated with holography, result in a high photographic γ . Thus the dynamic range over which linear recording can be accomplished is quite limited. In contrast, the spectrum of typical real world input imagery has high values at low spatial frequencies but falls sharply with frequency. (To linearly record an input scene spectrum out to the spatial frequency achievable by even a modest Fourier transform lens would require over three orders of magnitude dynamic range.) Within the dynamic range limitations of the plate an approximation to the matched spatial filter transfer function may be holographically recorded by adjusting the amplitude of the plane wave reference beam to be equal to, or greater than, the zero frequency component of the reference function spectrum. However, as pointed out earlier, this filter will not perform well at discriminating the reference object in clutter environments.

A much more effective filter for the target recognition task, which was originally suggested by Vander Lugt himself and has been described by several authors since, eg. Binns *et al.* [9], is to employ a reference beam much weaker than the zero frequency peak of the object spectrum. Thus the reference beam amplitude is set to achieve maximum fringe modulation within a bandpass of spatial frequencies and the exposure set to achieve linear recording of the fringe pattern within this bandpass. Both at lower and higher spatial frequencies the modulation will be very low due to the unbalanced amplitudes of the interfering beams and so these regions of the spectrum will not diffract significantly into the 1st order. In addition, the high intensity low spatial frequency regions of the spectrum will be strongly over-exposed. Since the hologram is recorded on a carrier frequency, the photographic negative can, and always is, used to reconstruct $H^*(\mathbf{k})$ (this results in an irrelevant overall π phase shift of the illuminating wavefront). Thus the low frequency regions of the input scene are essentially "dc" blocked by the over-exposed holographic plate which acts to whiten the spectrum. The overall result of these effects is that a correlation peak is obtained that is narrower than that which would result from a faithful realisation of $H^*(\mathbf{k})$, leading to higher discrimination due to

the low frequency removal. In addition, removal of the high frequencies allows the filter to maintain desirable tolerance to minor changes of the input object spectrum, giving some degree of noise tolerance. This is achieved similarly in the matched filter by attenuation of the reference spectrum high frequency components, $|H(\mathbf{k})|$ being recorded as well as $\arg H^*(\mathbf{k})$. These desirable properties of the holographically recorded filter also allow good results to be achieved when multiple images are multiplexed within a single filter, the resulting correlation response remaining sufficiently localised to be useful. Some of the best multiplexing results to date were achieved by Guenther *et al.* [10] using silver halide amplitude matched filters. The bandpass characteristics of the filter are of paramount importance to its pattern recognition performance and will be discussed in some detail later.

One further point that is of importance to the operation of the correlator is that if the input plane image is recorded on photographic film, a phase modulation coupled to the intensity distribution is known to occur [8]. This will reduce the value of the mean of the input amplitude distribution such that the correlation between unrelated backgrounds will be reduced and thus more closely match the assumption made in deriving the matched filter (see Equation {1.2} and the following discussion). Similarly, a spatial light modulator based on the hybrid field effect in a twisted nematic liquid crystal generates a phase shift of the coherent read-out beam in addition to the polarisation rotation (which is converted to an amplitude modulation by an analyser). Again the "dc" bias is removed that would otherwise be introduced by a pure amplitude modulation of the read-out wave coincident with the input scene intensity distribution. Contrary to this, however, when modelling the performance of a coherent optical correlator, the assumption is usually made that the device acts simply as a square law detector.

Whatever the chosen bandpass, the frequency plane filter should have a phase transfer function equal to the conjugate of the reference signal (as in the matched filter) in order to maximise the correlation signal. The reason is that this maximizes the projection onto the frequency plane real axis of the phase of the processing coherent wavefront as it emerges from the filter plane. Now since for a real input function:

$$s(\mathbf{r}) = \iint_{-\infty}^{\infty} \text{Re}[F(\mathbf{k})H^*(\mathbf{k})] \exp(j2\pi\mathbf{k} \cdot \mathbf{r}) d\mathbf{k}_x d\mathbf{k}_y \quad \{1.15\}$$

(since then $F(\mathbf{k})H^*(\mathbf{k})$ is diametric Hermitian and hence $s(\mathbf{o})$ is entirely real) the correlation peak, $s(\mathbf{o})$, is maximized for a filter of given amplitude modulation. (If the input scene contains a shifted version of $h(\mathbf{r})$, eg. $h(x_1, y_1)$ the Fourier transform will be

multiplied by an appropriate linear phase factor, $\exp(j2\pi(k_x x_1 + k_y y_1))$, which will result in a corresponding shift of the correlation maximum to location (x_1, y_1) in the output plane.)

The amplitude attenuation of the filter may then be used to minimize the variation of $s(\mathbf{o})$ for disruptions of the reference object from that used to form the filter, while simultaneously ensuring its value for $\forall x, y$ is minimized in response to scene clutter. As defined above the MSF achieves an optimum solution to this criterion when the clutter has specific statistical characteristics but these are often not appropriate to pattern recognition problems. Thus the $|H(\mathbf{k})|$ amplitude fall off with frequency of the MSF must be modified to achieve a more acceptable performance, as alluded to above. A range of filters can be generated by altering the amplitude modulation characteristics in the frequency domain to vary their sensitivity (ie. correlation peak fall-off from the maximum) and noise robustness. For the types of non-random clutter encountered in pattern recognition these characteristics are generally conflicting and a compromise in filter performance must be arrived at. To achieve this, a full range of filter modulus transfer functions are considered in some detail below, from the extreme of the inverse filter (very selective and noise sensitive) through to the MSF. This process may be considered as the variation of a single thresholding parameter, K , and will be discussed further in the next section.

Further refinements are possible by the use of the Wiener filter transfer function; the approximations and difficulties with this approach are also discussed below. An alternative, although similar, method which is easier to implement is to record the filter in a bandpass of spatial frequencies using a difference of Gaussian amplitude envelope. The theory behind this technique is discussed, and experimental results given, in Chapter 4. Improvement in intra-class tolerance, while maintaining inter-class discrimination, is demonstrated which is superior to that obtainable with a filter of unmodified modulus transfer function.

1.2 Variations from the matched filter

1.2.1 The inverse filter

An image degraded by a linear, space invariant process may be written :

$$s(\mathbf{r}) = h(\mathbf{r}) * g(\mathbf{r}) \quad \{1.16\}$$

where $g(\mathbf{r})$ is the point spread response of the process. $h(\mathbf{r})$ may be recovered by application of the inverse filter :

$$G_{\text{INV}}(\mathbf{k}) = \frac{1}{G(\mathbf{k})} \quad \{1.17\}$$

since in the Fourier domain Equation {1.16} becomes:

$$S(\mathbf{k}) = H(\mathbf{k}) G(\mathbf{k}) \quad \{1.18\}$$

which when multiplied by $G_{\text{INV}}(\mathbf{k})$ recovers $H(\mathbf{k})$ and hence $h(\mathbf{r})$. For pattern recognition purposes $g(\mathbf{r})$ is given the value $h^*(-\mathbf{r})$ and the optical system becomes a correlator matched for the recognition of $h(\mathbf{r})$ since the output is then:

$$s(\mathbf{r}) = h(\mathbf{r}) * h^*(-\mathbf{r}) = h^*(\mathbf{r}) \otimes h(\mathbf{r}) \quad \{1.19\}$$

In the frequency domain this becomes:

$$S(\mathbf{k}) = H(\mathbf{k}) H^*(\mathbf{k}) \quad \{1.20\}$$

ie. the field emerging directly behind the filter plane is proportional to $|H(\mathbf{k})|^2$, the transfer function of the filter being that of the MSF, $H^*(\mathbf{k})$.

If, instead, the filter is given the inverse filter transfer function, ie. $G_{\text{INV}}(\mathbf{k}) = \frac{1}{H(\mathbf{k})}$,

which may be equivalently written $G_{\text{INV}}(\mathbf{k}) = \frac{H^*(\mathbf{k})}{|H(\mathbf{k})|^2}$, the field emerging from behind

the filter when $h(\mathbf{r})$ is present at the input is:

$$S(\mathbf{k}) = \frac{H(\mathbf{k})H^*(\mathbf{k})}{|H(\mathbf{k})|^2} \quad \{1.21\}$$

ie. a uniform distribution, which Fourier transforms to the psf response of the (finite aperture) optical system. In practice this is a very localised impulse-like correlation response, much more localised than the true auto-correlation function of Equation {1.19} as generated by an MSF. Due to the absence of any high frequency attenuation the filter response is ultra-sensitive to minor variations of the reference function, the correlation signal being lost if $h(\mathbf{r})$ alters slightly, which may be brought about, for example, by a minor change in perspective of a solid object. For this reason the filter is of little practical value in correlator systems. Recently, synthetic discriminant function (SDF) filters have incorporated this transfer function characteristic into their design but have suffered from the expected drawbacks. These will be discussed further in the consideration of SDF filter design given in Chapter 2.

When the input signal contains additive noise the output signal may be written, in the frequency plane, as:

$$S(\mathbf{k}) = [H(\mathbf{k}) + N(\mathbf{k})] \frac{1}{H(\mathbf{k})} = \frac{N(\mathbf{k})}{H(\mathbf{k})} + 1 \quad \{1.22\}$$

Thus the field uniformity is disrupted in amplitude and phase by the noise contamination. This will be particularly severe at high frequencies if the noise is white since in these regions $|N(\mathbf{k})| \gg |H(\mathbf{k})|$ and the signal will be dominated by noise.

In addition, the effect of the inverse filter of boosting regions where $|H(\mathbf{k})|$ is small, especially where it passes through zero, leads to noise amplification in these regions (or no attenuation in the case of an optically implemented passive filter). Bates [11] suggests a modification to the inverse filter to improve its noise robustness.

The field emerging behind the filter may be divided into three distinct regions :

1. R_S where $|S(\mathbf{k})|$ has significant value.
2. R_H where $|H(\mathbf{k})| \geq |N(\mathbf{k})|$.
3. R_N where that region of R_S is not intersected by R_H ie. where $|N(\mathbf{k})| > |H(\mathbf{k})|$.

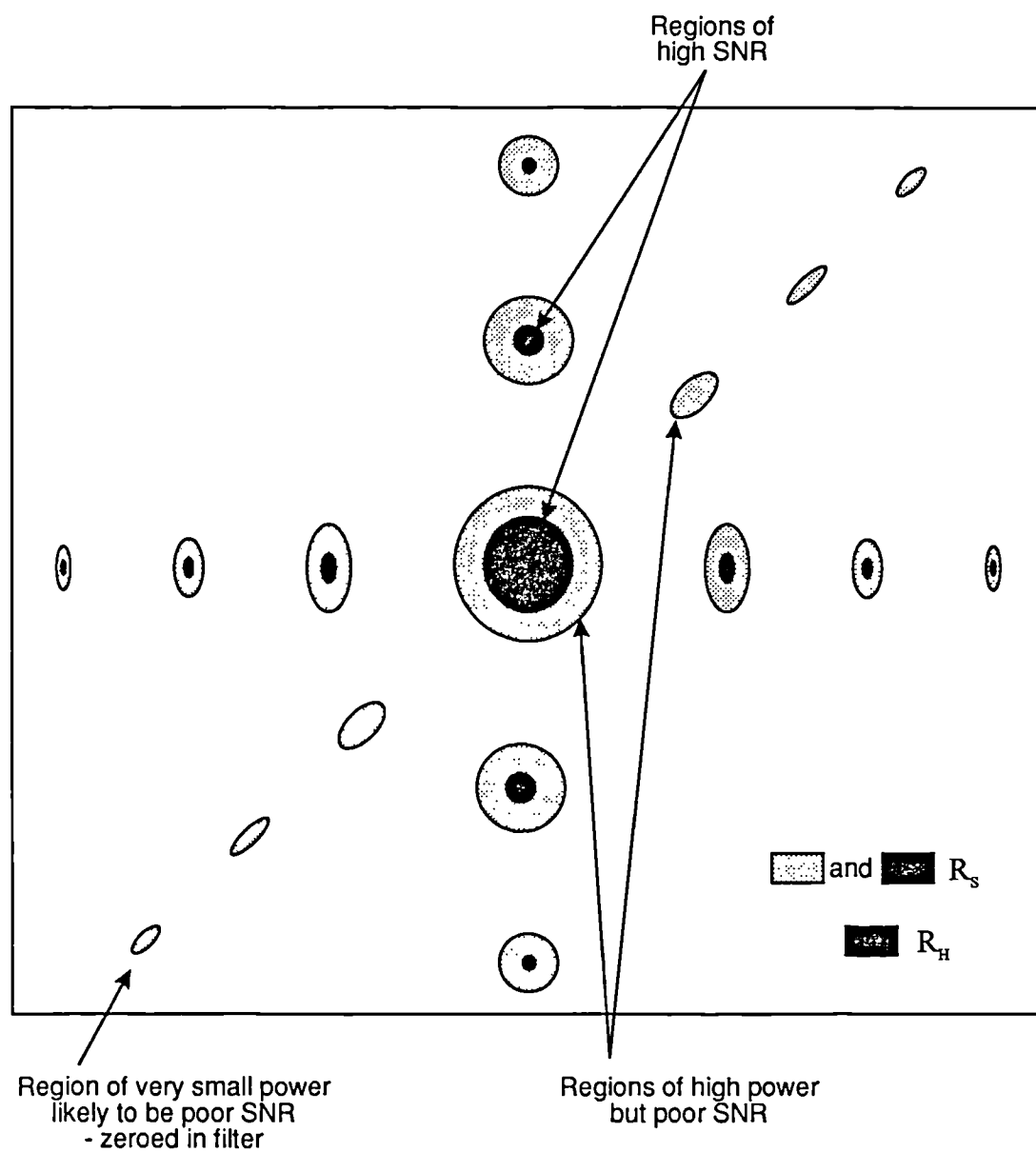


Figure 1.1 Schematic of frequency plane regions used in defining the modified inverse filter

These regions are shown schematically in Figure 1.1 in a distribution that would be typical of that resulting from an image and clutter spectrum. These regions may be used to define a modified inverse filter with transfer function :

$$\begin{aligned} M(\mathbf{k}) &= \frac{1}{H(\mathbf{k})} & \text{for } \mathbf{k} \in R_H \\ &= 0 & \text{for } \mathbf{k} \in R_N \end{aligned} \quad \{1.23\}$$

This filter provides better noise immunity but when used for image restoration introduces artefacts into the restoration due to the sharp discontinuities of the frequency domain function. The impulse response has an oscillatory form due to these discontinuities which when convolved with the image result in edge "ringing" (Gibbs phenomenon). However, used as a correlation filter the sidelobe structure of the resulting isolated correlation peak is of less importance, especially since the modulus squared at the output is measured, effectively attenuating the sidelobe structure and sharpening the peak. An exception is when the filter is heavily multiplexed since then the sidelobe structure from many cross-correlations sum in amplitude before detection, leading to a considerable broadening of the correlation response. This difficulty is discussed more fully in Section 2.3.2.

1.2.2 Reduction of filter sensitivity

A further refinement of the filter transfer function can be made by including a frequency dependent measure of the signal to noise ratio. This leads to a Wiener filter which will be described shortly. However, first it is useful to consider the modification of the inverse filter, via a single parameter K , through a series of intermediate states. These intermediates increase progressively in noise robustness and decrease in sensitivity as K is increased. Where $|H(\mathbf{k})|$ is less than K , the spectrum is conjugated and completely transmitted by the filter. However, in regions where $|H(\mathbf{k})| > K$, the transmittance function becomes $K(H^*(\mathbf{k}))^{-1}$, so that high intensity regions are attenuated to an extent determined by the magnitude of K . For any given application a value of K must be chosen which strikes a compromise between these two conflicting properties of the filter.

If the noise is not known *a priori* and cannot be reliably estimated, it is often simply assumed to have a uniform power spectrum. In such a case, the inverse filter transmittance will be reduced where $|H(\mathbf{k})|$ is small, ie. near the zero crossings of $H(\mathbf{k})$,

and at high spatial frequencies, where the SNR is low. As explained above in Section 1.1, this is necessary for a passive filter that cannot boost $H(\mathbf{k})$ and is the closest approximation of an inverse filter physically realisable with a passive device. Note that the efficiency will be very low since $H(\mathbf{k})$ must be attenuated to a low value of K in order to achieve a reasonable degree of uniformity in the transmitted spectrum, $S(\mathbf{k})$.

The field emerging from the filter will be approximately uniform in amplitude and phase when $H(\mathbf{k})$ is the input to the filter, Fourier transforming to an impulse-like auto-correlation function. The high sensitivity of the filter to deviations of the input spectrum from $H(\mathbf{k})$ can be further appreciated by remembering that differentiating a space domain function is equivalent to multiplying its spectrum by $j\omega$ ie.

$$F \{ \nabla h(\mathbf{r}) \} \equiv j 2\pi |\mathbf{k}| H(\mathbf{k}) \quad \{1.24\}$$

Thus amplification of the high frequency components of an image spectrum leads to an accentuation of the fine detail in the image or to an edge enhancement which is similar, in effect, to a differentiation of the image. The spectra of natural scenes tend to have a fractal, $1/|\mathbf{k}|$ magnitude fall-off with frequency. Thus since the field emerging from the inverse filter may be written :

$$S(\mathbf{k}) = (G(\mathbf{k}) + N(\mathbf{k})) \frac{1}{|G(\mathbf{k})|} \left[\frac{1}{|G(\mathbf{k})|} G^*(\mathbf{k}) \right] \quad \{1.25\}$$

it can be seen that the filtering operation is equivalent to correlating a differentiated reference with an approximately differentiated input scene. Thus an MSF could be used in place of the inverse filter if made with a differentiated reference function and illuminated with an image differentiated prior to input to the filter; the correlation response would be similar to that obtained with an inverse filter.

1.2.3 The phase only filter

Increasing K to be equal to $|H(\mathbf{k})|_{\max}$ leads to the phase only filter, about which there has been considerable interest in the recent literature since attention was drawn to it by Horner and Gianino [12]. The main motivation of these authors was the observation that the information content of an image (in the sense of intelligibility of features and structures within the image) seems to be almost entirely contained within the phase of the image spectrum [13]. (However, rather paradoxically, it is also a fact that a signal of two

or higher dimensions can, at least in theory, be uniquely recovered from knowledge of its power spectrum alone [11].) Thus if the phase of the reference signal spectrum is recorded by this filter, it should contain adequate information about the reference function to generate a strong recognition signal in the form of a correlation peak.

It has been found that the correlation decreases quite sharply when the input object changes from that of the reference ie. alters in scale, in-plane rotation or aspect. The decrease is not as severe, however, as that for an inverse filter (which may be considered as the final member of the filter family of which the MSF corresponds to the opposite extreme). The reason can again be clearly seen by examining the transfer function of the POF, which is simply $\exp(-j\phi_H(\mathbf{k}))$ or, equivalently, $H^*(\mathbf{k})/|H(\mathbf{k})|$. For a $1/|\mathbf{k}|$ frequency fall-off of $H(\mathbf{k})$, it will correspond to the storage of a differentiated, ie. edge enhanced, reference function. In contrast to the inverse filter, this is then addressed by the unmodified spectrum of the input distribution. Thus the filter output is equivalent to the correlation of an edge enhanced reference function with the unmodified input scene. This will, therefore, result in less sensitivity to alterations of the target object as compared to the inverse filter, which, as shown above, is equivalent to the correlation between an edge enhanced reference *and* input. On the other hand, cross-correlations with non-target objects and scene clutter are likely to increase. This could also be anticipated by noting that all the energy of the input scene spectrum reaches the correlation plane. Thus high frequency components will be transmitted to the output plane and are likely to be of little value for pattern recognition since they correspond to very fine detail in the image and will change with minor alterations of the target object from that of the reference. Here they will only contribute to a decrease in the SNR (in the sense of Equation {1.3}). Although of low intensity, the high frequency regions occur over a large area and so their degrading effect can be significant. This is, of course, why the MSF transfer function attenuates this contribution of the input spectrum.

In addition, the POF allows the low frequency components of the input spectrum to pass unattenuated. For an image, these components are very large and correspond to gross image features which will contribute to high false cross-correlation responses. The MSF transfer function does not attenuate these components, but it should be remembered that in its design the background scene has been treated as noise uncorrelated with the signal. In pattern recognition, as repeatedly emphasised, the input image is not of zero mean (the zero frequency spectral component being proportioned to the integral of the image intensity distribution) and the "noise" may well consist of a similar shaped object to the reference which is highly correlated with it, particularly at low spatial frequencies. Thus the low spatial frequency response of the filter must be attenuated. These points again

lead back to the idea of a filter operating in a bandpass of spatial frequencies which is chosen so as to maximize intra-class tolerance versus inter-class discrimination for the target and non-target objects being considered. Since the POF filter transmits all of the input spectrum to the output plane, it obviously has the maximum optical efficiency possible. Horner and Gianino [12] originally cited the ratio:

$$\eta_H = \eta_m \frac{\int \int_{-\infty}^{\infty} |s(\mathbf{r})|^2 dx dy}{\int \int_{-\infty}^{\infty} |f(\mathbf{r})|^2 dx dy} \quad \{1.26\}$$

(where η_m is a parameter used to account for losses in the physical implementation of the POF) as a measure of the light efficiency of an optical correlator. This ratio does not take into account the degree to which the available energy is concentrated into the correlation signal as opposed to noise in the output plane. The definition has recently been modified to [14]:

$$\eta_H = \eta_m \frac{s^2(\mathbf{o})}{\int \int_{-\infty}^{\infty} |f(\mathbf{r})|^2 dx dy} \quad \{1.27\}$$

in order to give a measure of how much of the input energy is concentrated into the output correlation signal. However, as can be appreciated from the preceding discussion, the filter maximising η_H in Equation {1.27} does not by any means maximize the output SNR (in the sense of the definition of Equation {1.3}) or maximize the probability of true target identification simultaneously with false target rejection. A reduced η_H may have to be accepted as a necessary consequence of the increased pattern recognition performance of such a filter.

The POF has attracted considerable attention mainly because it is physically realisable by non-interferometric means on phase-modulating spatial light modulators. Further, the correlation response has proved to be robust against gross quantisation of the phase values stored as the filter transfer function. Binary phase quantisation still yields usable correlation performance although the correlation peak height can be expected to be decreased considerably from that resulting with a finer phase quantisation. This can be seen to result from a reduction of the projection of the spectrum phasors onto the frequency plane real axis as shown in Equation {1.15}. However, since spatial light modulators such as the SIGHT-MOD [15] exist that are capable of being operated in a binary phase only mode, a considerable amount of simulation work has been done of

binary POFs [16]. Hardware realisations with a binary amplitude SLM at the input and a binary phase SLM in the filter plane have been constructed [17].

Methods of improving the performance of the POF have been suggested by incorporating an additional binary amplitude modulation. Vijaya Kumar and Bahri [18] use the MSF SNR criterion to optimise phase only filters by reducing the area of support of the filter which leads to a low pass filter with a response similar to that of an MSF. Flannery *et al.* [19] have modified a binary POF by the addition of binary amplitude modulation. The filter, known as a transform ratio ternary phase-amplitude filter, attempts to maximize the ratio of true-class to false-class correlation signal by comparing the ratio of the magnitudes of the power spectra of each class at each discrete spatial frequency. If this is below a certain value the pixel is set to zero and this spectral frequency component is removed. This method thus explicitly attempts to improve discrimination based on training data and so produces better performance than those filters optimised for inappropriate statistical SNR criteria. The technique is, in fact, a binary approximation to the MSF of Equation {1.7} with $|N(\mathbf{k})|^2$ being set equal to the average of the false-class power spectra. Practical implementation of the ternary filter is possible with a SIGHT-MOD as described by Kast *et al.* [20]. Although more difficult to practically implement, the Wiener filter used with $|N(\mathbf{k})|^2$ generated in a similar manner, yields an improved performance. For this reason it is discussed in the following section.

1.2.4 The Wiener filter

Given a system input described by a linear, space invariant process and degraded by additive stationary noise, the system output may be written as :

$$s(\mathbf{r}) = g(\mathbf{r}) * h(\mathbf{r}) + n_g(\mathbf{r}) \quad \{1.28\}$$

The Wiener filter finds an estimate:

$$\hat{s}(\mathbf{r}) = w(\mathbf{r}) * s(\mathbf{r}) \quad \{1.29\}$$

such that the mean square error :

$$\varepsilon = \langle [s(\mathbf{r}) - \hat{s}(\mathbf{r})]^2 \rangle \quad \{1.30\}$$

is minimised. Thus :

$$\varepsilon = \left\langle [h(\mathbf{r}) - w(\mathbf{r}) * s(\mathbf{r})]^2 \right\rangle \quad \{1.31\}$$

By the orthogonality principle [5] this is a minimum when the error signal, $\varepsilon(\mathbf{r}) = h(\mathbf{r}) - \hat{s}(\mathbf{r})$, is orthogonal to the data which leads to the expression :

$$[h(\mathbf{r}) - w(\mathbf{r}) * s(\mathbf{r})] \otimes s^*(\mathbf{r}) = 0 \quad \{1.32\}$$

which in the frequency domain becomes :

$$H(\mathbf{k}) S^*(\mathbf{k}) - W(\mathbf{k}) S(\mathbf{k}) S^*(\mathbf{k}) = 0 \quad \{1.33\}$$

Hence :

$$W(\mathbf{k}) = \frac{S^*(\mathbf{k})}{|S(\mathbf{k})|^2} H(\mathbf{k}) \quad \{1.34\}$$

or since $S(\mathbf{k}) = G(\mathbf{k}) H(\mathbf{k}) + N(\mathbf{k})$, written in full Equation {1.34} becomes :

$$W(\mathbf{k}) = \frac{G^*(\mathbf{k}) |H(\mathbf{k})|^2 + N^*(\mathbf{k}) H(\mathbf{k})}{|G(\mathbf{k})|^2 |H(\mathbf{k})|^2 + |N(\mathbf{k})|^2 + G(\mathbf{k}) H(\mathbf{k}) N^*(\mathbf{k}) + G^*(\mathbf{k}) H^*(\mathbf{k}) N(\mathbf{k})} \quad \{1.35\}$$

Generally, the assumption is made that the signal and noise are uncorrelated and so Equation {1.35} reduces to :

$$W(\mathbf{k}) = \frac{G^*(\mathbf{k}) |H(\mathbf{k})|^2}{|G(\mathbf{k})|^2 |H(\mathbf{k})|^2 + |N(\mathbf{k})|^2} \quad \{1.36\}$$

which is conveniently written :

$$W(\mathbf{k}) = \frac{G^*(\mathbf{k})}{|G(\mathbf{k})|^2 + \frac{|N(\mathbf{k})|^2}{|H(\mathbf{k})|^2}} \quad \{1.37\}$$

from which it can be seen that in regions of the signal spectrum where the SNR is high, the filter $W(\mathbf{k})$ approaches the inverse filter transfer function, $G_{\text{inv}}(\mathbf{k})$, and so has unity transmittance. However, where $|H(\mathbf{k})|^2 \rightarrow 0$, and in other regions of poor SNR, the Wiener filter has a low transmittance.

The Wiener filter has been widely applied to image restoration, its use for this purpose being introduced by Hellstrom [22]. However, since $|N(\mathbf{k})|^2 / |H(\mathbf{k})|^2$ varies smoothly, the transfer function of the filter does not contain the abrupt discontinuities of the modified inverse filter, $M(\mathbf{k})$, (see Equation {2.22}) and so alleviates the "ringing" artefacts associated with the frequency domain discontinuities introduced by this filter. In this application a problem arises in that the image to be restored, $h(\mathbf{r})$, and hence its spectrum, $H(\mathbf{k})$, is not known *a priori*. Further, explicit knowledge of $|N(\mathbf{k})|^2$ is also usually unavailable and must be estimated. Often the ratio $|N(\mathbf{k})|^2 / |H(\mathbf{k})|^2$ is simply taken as a constant and varied in value until the best subjective restoration of the image is obtained.

As with the inverse filter, the Wiener filter may be used for pattern recognition if $G(\mathbf{k})$ is set equal to $H(\mathbf{k})$ ie. the spectrum of the object to be recognised. When this object is input to the system contained in background clutter with spectrum $N(\mathbf{k})$, the field directly behind the filter may, therefore, be written :

$$S(\mathbf{k}) = [H(\mathbf{k}) + N(\mathbf{k})] \frac{H^*(\mathbf{k})}{|H(\mathbf{k})|^2 + \frac{|N(\mathbf{k})|^2}{|H(\mathbf{k})|^2}}$$

or

$$S(\mathbf{k}) = \frac{H^*(\mathbf{k})N(\mathbf{k}) + |H(\mathbf{k})|^2}{|H(\mathbf{k})|^2 + \frac{|N(\mathbf{k})|^2}{|H(\mathbf{k})|^2}} \quad \{1.38\}$$

If $N(\mathbf{k})$ is low, this reduces to unity (except in the zero crossing regions of $H(\mathbf{k})$) and the filter becomes an inverse filter maximally sensitive to the target signal. However, under poor signal to clutter ratio conditions Equation {1.38} may be approximated by :

$$S(\mathbf{k}) = \frac{H^*(\mathbf{k})N(\mathbf{k})|H(\mathbf{k})|^2 + |H(\mathbf{k})|^4}{|N(\mathbf{k})|^2}$$

or

$$S(\mathbf{k}) = \left\{ \left(\frac{H(\mathbf{k})}{N(\mathbf{k})} \right)^* + \frac{|H(\mathbf{k})|^2}{|N(\mathbf{k})|^2} \right\} |H(\mathbf{k})|^2 \quad \{1.39\}$$

This should be compared to the field emerging from behind the MSF :

$$S(\mathbf{k}) = \left\{ \left(\frac{H(\mathbf{k})}{N(\mathbf{k})} \right)^* + \frac{|H(\mathbf{k})|^2}{|N(\mathbf{k})|^2} \right\} \quad \{1.40\}$$

Thus under very high noise conditions the Wiener filter further attenuates the high frequency components of $S(\mathbf{k})$ as compared to the MSF where the value of $|H(\mathbf{k})|$ is likely to be much lower than the noise values. The Wiener filter is thus adaptive from a maximally highpass filter to a lowpass filter depending on the SNR conditions. Importantly, it is also locally adaptive; in the regions of the input spectrum where the SNR deteriorates, the filter transmittance is reduced and this portion of the spectrum is attenuated.

Problems are encountered, however, in its use as a modified matched filter for pattern recognition purposes. The reduced form of the Wiener filter for uncorrelated signal and noise (Equation {1.37}) is suggested in the literature [22]. As has been emphasised throughout the discussion this assumption is, in general, invalid for the pattern recognition application; the "noise" may well consist of an object of similar shape to the target object and so be highly correlated with it. Even background clutter may be strongly correlated with the target object. Therefore, the full form of the Wiener filter (Equation {1.35}) should ideally be used, although this is considerably more complicated in functional form.

The further difficulty also again arises as to how to assign $|N(\mathbf{k})|^2$. The input to the correlator contains the target signal plus the background and so cannot be used to obtain $|N(\mathbf{k})|^2$. The best approximation would be to generate $|N(\mathbf{k})|^2$ from typical terrain data. The spectral characteristics of wooded land or other natural scenes could be averaged if this was appropriate to the anticipated background against which the target objects must be recognised. Such data could, possibly, be gathered in real-time from the terrain encountered as it was overflown and introduced to the Wiener filter just prior to its use. A possible means of accomplishing this dynamic filter generation will be discussed in Section 4.6.

A method of improving the discrimination performance of the simplified Wiener filter of Equation {1.37} would be to consider non-target spectra as part of $|N(\mathbf{k})|^2$ and include them as part of this term. For situations in which the background can be constrained to be relatively uniform, as in most industrial recognition/inspection scenarios, $|N(\mathbf{k})|^2$ may be constructed solely of averaged non-target spectra, weighted by the probability of their occurrence, as recently suggested by Yaroslavsky [23]. This essentially acts to decorrelate the target and non-target objects without recourse to the more involved transfer function of Equation {1.35} which would be harder to implement since it involves the complex terms $N(\mathbf{k})$, $N^*(\mathbf{k})$, $H(\mathbf{k})$ and $H^*(\mathbf{k})$. In contrast, in Equation {1.37} the only complex term is $G^*(\mathbf{k})$, which is in the numerator and may be realised holographically by the method of Vander Lugt. The denominator may be approximated with a photographically recorded inverse filter, as with the synthesis of the (non-white) matched filter, since it does not include any complex terms. Alternatively, the transfer function of a hologram recorded in a photorefractive crystal may be appropriately modified, with the advantage that this can be done dynamically and so updated with changes of $|N(\mathbf{k})|^2$. Further, it should be observed that the transfer function of Equation {1.37} is bandpass in nature, the high intensity low frequency spectral components being attenuated by $|G(\mathbf{k})|^2$ in the denominator and the high frequency terms being attenuated as $|H(\mathbf{k})|$ decreases and the ratio $|N(\mathbf{k})|^2/|H(\mathbf{k})|^2$ becomes large. In Chapter 4 a bandpass approximation to the Wiener filter is discussed which is more conveniently optically implemented, particularly when a Gaussian TEM_{00} mode laser is used to provide the processing wavefront, as is generally the case in coherent optical processing systems.

Use of the Wiener filter has also been suggested in conjunction with multiplexed filters designed by the synthetic discriminant function technique [22]. The correlation peak localisation resulting from use of a bandpass filtered SDF is of particular value and is discussed in Chapter 2 where the transfer function modifications required to improve SDF performance are investigated. Again, the bandpass filtered SDF is shown to result in very similar compromises between signal discrimination and distortion tolerance to those considered above. Correct design allows the minimum number of object views to be incorporated into the SDF for recognition of the target object over a given swathe of orientations.

References to Chapter 1

- [1] D. North, "An analysis of the factors which determine signal/noise discrimination in pulsed-carrier systems," *RCA Tech. Rept.* PTR-6C, June 25, 1943 (ATI 14009). Reprinted in *Proc. IEEE*, Vol. 51, p.1016, (1963)
- [2] B. Horn, "Robot vision," MIT Press and McGraw-Hill, (1986)
- [3] M. Skolnik, "Introduction to radar systems," McGraw-Hill International Edition, 20th Ed., (1988)
- [4] C. McMullen, "Communication theory principles," MacMillan, (1968)
- [5] J. Blackledge, "Quantitative coherent imaging," Academic Press (1989)
- [6] R. Sadler, M. Buttinger, "Acousto-optic ambiguity function processor," *IEEE Proc.*, Vol. 133, Pt.J, No.1, p.35, (1986)
- [7] A. Vander Lugt, "Signal detection by complex spatial filtering," *IEEE Trans.*, IT-10, p.139, (1964)
- [8] J. Goodman, "Introduction to Fourier optics," McGraw-Hill, (1968)
- [9] R. Binns, A. Dickinson, B. Watrasiewicz, "Methods of increasing discrimination in optical filtering," *Appl. Opt.* Vol. 7, p.1047, (1968)
- [10] B. Guenther, C. Christensen, J. Upatnieks, "Coherent optical processing: another approach," *IEEE, J. Quantum Elect.*, Vol. QE-15, No. 12, p.1348, (1979)
- [11] R. Bates, M. McDonnell, "Image restoration and reconstruction," Clarendon Press, Oxford, (1989)
- [12] J. Horner, P. Gianino, "Phase-only matched filtering," *Appl. Opt.*, Vol. 16, No. 11, p.829, (1991)
- [13] A. Oppenheim, J. Lim, "The importance of phase in signals," *Proc. IEEE*, Vol. 69, No. 5, p.529, (1981)

- [14] J. Horner, "Clarification of Horner efficiency," *Appl. Opt.*, Vol.31, No. 21, p.4629, (1992)
- [15] W. Ross, D. Psaltis, R. Anderson, "Two dimensional magneto-optic spatial light modulator for signal processing," *Opt. Eng.*, Vol. 24, p.485, (1983)
- [16] J. Horner, J. Leger, "Pattern recognition with binary phase only filters," *Appl. Opt.*, Vol.24, p.609, (1985)
- [17] D. Ben-Meir, S. Mills, "Optical correlator based ATR system implementation and evaluation," in *Optical Pattern Recognition III*, D. Casasent, T.-H. Chao, Eds, *Proc. SPIE*, Vol. 1701 p.239, (1992)
- [18] B. Vijaya Kumar, Z. Bahn, "Phase-only filters with improved signal-to-noise ratio," *Appl. Opt.*, Vol. 28, No.2, p.250, (1989)
- [19] D. Flannery, J. Loomis, M. Milkovich, "Transform-ratio ternary phase-amplitude filter formulation for improved correlation discrimination," Vol. 27, No. 19, p.4079, (1988)
- [20] B. Kast, M. Giles, S. Lindell, D. Flannery, "Implementation of ternary phase amplitude filters using a magneto-optic spatial light modulator," *Appl. Opt.*, Vol. 28, No. 6, p.1044, (1989)
- [21] C. Helstrom, "Image restoration by the method of least squares," *JOSA A*, Vol. 57, No. 3 p.297, (1967)
- [22] Ph. Réfrégier, J. Figue, "Optimal trade-off filters for pattern recognition and their comparison with the Wiener approach," *Optical computing and processing*, Vol. 1, No. 3, p.245, (1991)
- [23] L. Yaroslavsky, "Is the phase-only filter and its modifications optimal in terms of the discrimination capability in pattern recognition?" *Appl. Opt.*, Vol. 31, No. 11, p.1677, (1992)

Chapter 2

Design and simulation of a bandpass synthetic discriminant function

2.1 Introduction

It has been recognised for a number of years now, that to exploit to its full advantage the application of a coherent optical correlator to two dimensional image recognition, as originally proposed by Vander Lugt [1], a hybrid implementation is required, the reference templates being selected and passed to the optical correlator by a host digital computer [2]. The reason for this is that in typical application areas, such as in the recognition and location of a military vehicle in a generally cluttered environment, there are variations in the input image that cause loss of the correlation signal. These include changes of orientation, in-plane rotation, scale and direction of illumination. Several other optically implementable algorithms, such as various pre-processing coordinate mapping operations and filter multiplexing techniques, can be used to ameliorate these difficulties. However, the resulting multi-channel system requires control of the sequencing of the various operations and interpretation of the channel outputs by a digital processor. Even with such a system, the acquisition and maintenance of a correlation signal will still realistically require many directed, sequential template matching operations. For a system aimed solely at object recognition and only requiring infrequent re-programming, these templates can be stored optically by exploiting the high storage capacity of volume holographic media [3,4]. Rapid random access of these templates can then be achieved by control of the angle or wavelength of the readout beam from the digital computer.

The same basic optical pattern recognition algorithms may also be employed for industrial recognition and inspection tasks. In these application areas the initial object recognition task can be made more highly constrained than in the military case by, for instance, ensuring uniform lighting and constant camera to object distance. However, the subsequent inspection stages require the flexibility and decision making capabilities of a knowledge based system, implemented in an artificial intelligence language such as LISP. The optical processor is then under direct control from the digital computer, its function being to implement the otherwise CPU intensive basic pattern recognition algorithms. This was the concept behind the work carried out under a recent European

BRITE/EuRAM funded project [5] of which the work reported in this Chapter forms a part.

It can be appreciated that central to such hybrid schemes is an effective digital to optical interface. Several solutions to this problem are currently being investigated and include: techniques based on direct phase modulation of the coherent processing wavefront by a spatial light modulator [6], the use of high resolution SLMs to act as dynamic holograms [7] and the use of photorefractive materials for a similar purpose [8,9]. The third alternative was chosen for implementation of an up-dateable correlator in the cited Project. This allows both the input and reference to be transferred to the optical system as space domain images via the more readily available amplitude modulating SLMs. The use of the widely available photorefractive material, Bismuth Silicon Oxide has several important consequences for the overall system design. In the following Section a brief description of the BSO based up-dateable correlator is given. Factors are considered that must be taken into account in order to obtain optimal performance for the system in terms of spatial bandwidth product and speed of response.

However, the finite response time of the input SLMs and the required formation time of the holographic grating in the photorefractive crystal, combined with the requirement for many reference searches per input frame in the initial stages of object acquisition, still compromise the real-time capability of the system. Thus any reduction in the number of reference template searches that can be realised by the use of filter multiplexing techniques is valuable, provided the resulting deterioration in the correlation signal can be minimized. In this regard, modifications to the synthetic discriminant function filter algorithm are shown to produce a high discriminatory, localised correlation response for all inputs included in the training set. The design of this filter is described which is implementable by recording a multiplexed space domain reference function as a Fourier transform hologram written to the BSO with a relatively weak reference beam, such that high spatial frequencies are accentuated. Simulation results are presented for its performance on one of the test components, TC3, used during the BRITE/EuRAM project. Modifications to the filter design to increase the flexibility of both its application and hardware realisation are briefly discussed in the conclusion to this work.

2.2 The up-dateable correlator

2.2.1 System configuration

The use of degenerate four wave mixing (DFWM) to implement the real time correlation between two images was first proposed by Pepper *et al.* [8]. Much work on the use of the photorefractive crystal Bismuth Silicon Oxide (BSO) has since been done by Huignard and co-workers who demonstrated a joint transform correlator (JTC) configuration employing non-degenerate four wave mixing in BSO [10]. An alternative NDFWM configuration that more closely resembles a Vander Lugt correlator was proposed by Cooper *et al.* [10] and further modelled by Nicholson *et al.* [11]. This configuration has certain inherent advantages over the JTC and was adopted as the basis for the up-dateable correlator in the work carried out during the BRITE/EuRAM project. A system diagram of the NDFWM scheme used to implement the up-dateable correlator is shown in Figure 2.1. The input and reference images are displayed on the input SLMs resulting in the field modulations, $u_1(\mathbf{r})$ and $u_2(\mathbf{r})$, of the Ar^+ and HeNe beams respectively. A hologram is written to the BSO by the interference of $U_1(\mathbf{k})$, the Fourier transform of $u_1(\mathbf{r})$, and U_3 , a plane wave reference beam of variable strength. $U_2(\mathbf{k})$, the Fourier transform of $u_2(\mathbf{r})$, is diffracted from the grating formed in the BSO to yield, after a further Fourier transform, the correlation between $u_1(\mathbf{r})$ and $u_2(\mathbf{r})$ at plane P_4 .

Two factors of critical importance in the use of a photorefractive material in the up-dateable correlator are the magnitude of the induced refractive index change and the speed of the effect. It has been pointed out [12] that, since the electro-optic and permittivity tensors of electro-optic materials are related, a compromise between these two properties has to be made. Thus materials with very large e-o coefficients, such as the ferroelectric Barium Titanate (BaTiO_3), have a response time of the order of seconds at mW/cm^2 power densities. Whereas, at the other extreme, III-V compounds such as Gallium Arsenide (GaAs) exhibit a much weaker non-linear effect on microsecond time scales. Bismuth Silicon Oxide ($\text{Bi}_{12}\text{SiO}_{20}$) strikes a compromise between these two extremes with an effective Pockels coefficient of about 5 pmV^{-1} and a response time of several tens of milliseconds at mW/cm^2 power densities.

These two factors govern the configuration of the correlator shown in Figure 2.1. This would more closely resemble a Vander Lugt correlator if the reference image Fourier transform and plane wave reference beam were to form the grating in the BSO. The reference and input image positions are transposed, however, so that the grating is

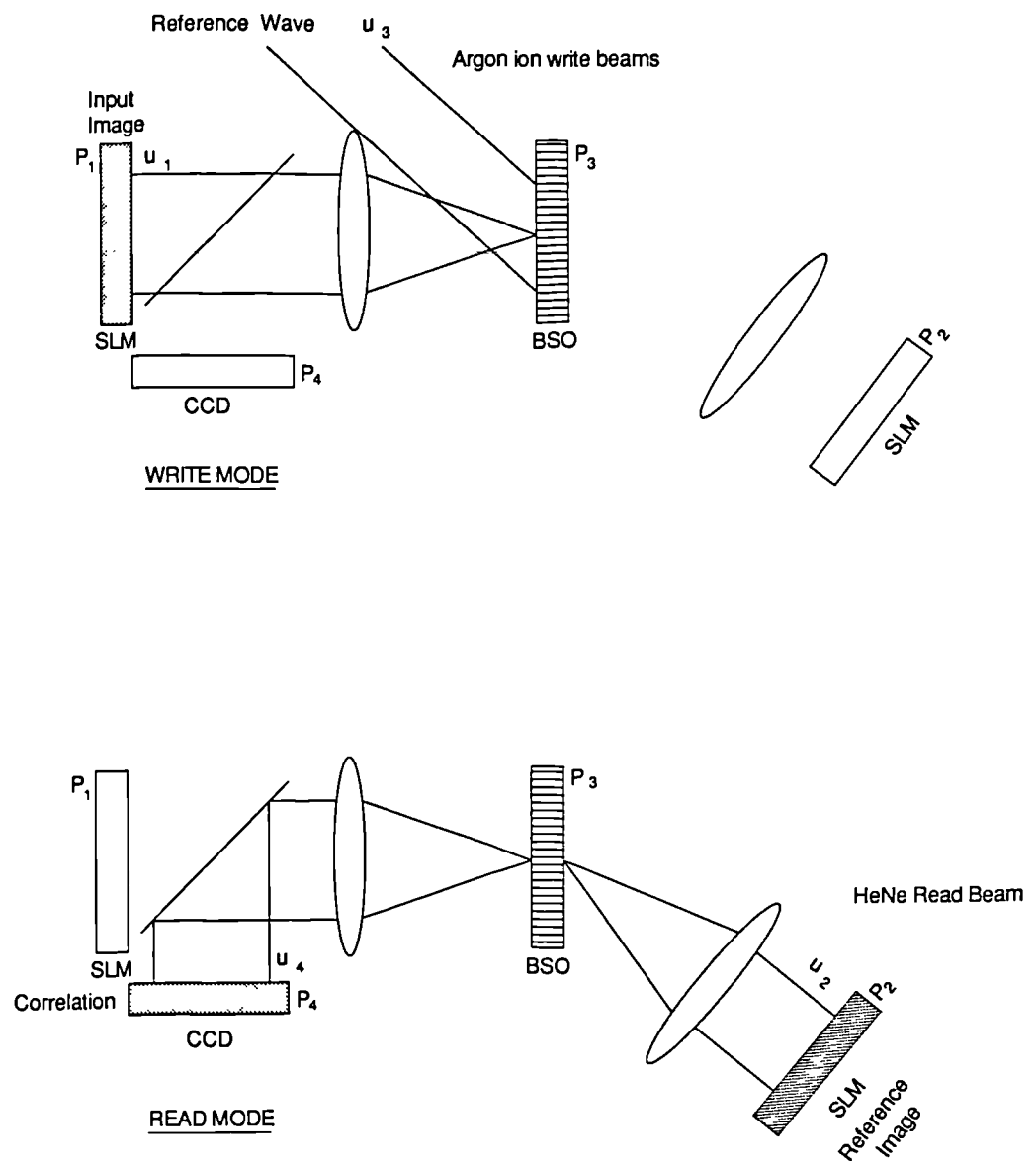


Figure 2.1 Schematic of up-dateable photorefractive correlator

formed by the input image Fourier transform. There are two reasons for this. Firstly, the Bragg phase matching constraints of the volume hologram formed in the BSO are much less severe for this beam, allowing a larger effective field of view for a given lens f number. Thus the spatial invariance property of the correlation algorithm can be exploited since the unknown image can be located anywhere within this input area. In contrast, the reference template must be exactly centred in the HeNe beam, any translation resulting in loss of the correlation signal due to the rapid dephasing that occurs in this beam. Secondly, there is an asymmetry in the speed requirements between the input and reference Fourier transforms. In general, many reference templates must be searched to identify the unknown input. In the NDFWM configuration the reference template can be updated as rapidly as the SLM and readout CCD can be operated, as the reference Fourier transform simply diffracts from an existing grating in the BSO formed once per input cycle [10]. Thus the relatively slow response time of the BSO can be prevented from degrading the system response time too greatly.

As mentioned above, Bragg matching constraints impose a minimum f number on the Fourier transform lenses that can be employed for a given input and reference function spatial bandwidth product. The required f number can be directly related to the magnitude of the e-o effect in the photorefractive material. This is because the resulting change of the refractive index induced in the material governs how thick the volume hologram must be to achieve the necessary diffraction efficiency.

The use of a HeNe read beam, in addition to having certain practical advantages, simplifies the modelling of the correlator since non-linear interaction occurs only between the Ar^+ write beams. The HeNe beam is simply diffracted from the static grating formed by these two beams, rather than writing a further grating as in the case of DFWM. Theoretical estimation of diffraction efficiency, and its loss due to Bragg mismatching, is complicated by the high degree of optical activity and electric field induced birefringence in BSO [13]. Thus only relative diffraction efficiencies are accounted for by considering the overall modulation of the grating formed by $U_1(\mathbf{k})$ and U_3 . This modulation spatially alters the diffracted amplitude of $U_2(\mathbf{k})$ giving rise to the spatial filtering effects to be exploited in the operation of the correlator.

For the steady state and ignoring any beam coupling, the grating may be written [11]:

$$m(\mathbf{k}, z) = \frac{2U_1^*(\mathbf{k})\exp\left(\frac{-a_w z}{2}\right)U_3(\mathbf{k})\exp\left(\frac{-a_w z}{2}\right)}{|U_1(\mathbf{k})|^2 \exp(-a_w z) + |U_3(\mathbf{k})|^2 \exp(-a_w z) + \eta_q |U_2(\mathbf{k})|^2 \exp(-a_r(d-z))} \quad \{2.1\}$$

where $a_{r/w}$ are the absorption coefficients of the write and read beams in BSO.

Since η_q , the quantum efficiency of the interaction of the HeNe beam with the grating is only about 0.06 of that of the Ar^+ beam, Equation {2.1} reduces to:

$$m(\mathbf{k}) = \frac{2U_1^*(\mathbf{k})U_3}{|U_1(\mathbf{k})|^2 + I_3} \quad \{2.2\}$$

where U_3 , and hence I_3 , are constants (plane wave reference).

If the assumption of linearity is made between the grating modulation and the locally generated space charge field in the material, the amplitude of the diffracted field directly behind the grating, $U_4(\mathbf{k})$ will be proportional to $m(\mathbf{k})$ and so can be written:

$$U_4(\mathbf{k}) = m(\mathbf{k})U_2(\mathbf{k}) \quad \{2.3\}$$

The Fourier transform of $U_4(\mathbf{k})$ gives $u_4(\mathbf{r})$, the weighted correlation between $u_1(\mathbf{r})$ and $u_2(\mathbf{r})$.

2.2.2 Variation of response with beam ratio

As mentioned in the Introduction, the spatial frequency weighting of the correlation operation can be altered by varying the amplitude of the plane wave reference beam U_3 . This is accomplished with an electrically controlled electro-optic beam attenuator developed during the BRITE/EuRAM project. The beam ratio between $U_1(\mathbf{k})$, the Fourier transform of the input image, and U_3 is defined as:

$$BR = \frac{\text{Amplitude of } U_1(\mathbf{k}) \text{ peak (at zero spatial frequency)}}{\text{Amplitude of plane wave reference beam } U_3}$$

(Note that since any experimentally measured values will be an intensity ratio they must be square rooted to give the corresponding BR).

The 2-D Fourier transforms of images have large zero and low frequency components. Thus if the beam ratio is set to unity, values of $|U_1(\mathbf{k})|^2$ for all but the lowest frequencies will be small compared to I_3 . Thus a good approximation to Equation {2.2} is:

$$m(\mathbf{k}) \approx \text{const.} * U_1^*(\mathbf{k}) \quad \{2.4\}$$

When $U_2(\mathbf{k})$ is diffracted from the grating the field *just* behind the filter will be:

$$U_4(\mathbf{k}) = \text{const.} * U_1^*(\mathbf{k}) U_2(\mathbf{k}) \quad \{2.5\}$$

Thus with BR set to unity the BSO hologram will act as a classical matched filter [10]. The correlation peak obtained will be broad and discrimination between non zero mean input imagery will be poor.

If the beam ratio is set to a much higher value by attenuating the plane wave reference, for low frequencies where $|U_1(\mathbf{k})|^2 \gg |U_3|^2$, Equation {2.2} becomes:

$$m(\mathbf{k}) = \frac{2U_3}{U_1(\mathbf{k})} \quad \{2.6\}$$

(BR=32 was the highest ratio examined ie. an intensity ratio of 1000). Under these circumstances there is a small modulation of the grating in the central region of the hologram and thus the low frequency components of $U_2(\mathbf{k})$ are diffracted by the grating with very low efficiencies. At higher frequencies, where $|U_1(\mathbf{k})| \approx |U_3|$, the modulation is nearer unity leading to the maximum diffraction efficiency consistent with the parameters of the volume hologram. As will be shown below, these modulation conditions lead to an effect that is very similar to a 2nd differential pre-processing operation on the input images prior to an all-pass correlation. This gives rise to a high discriminatory correlation response but with the possibility of varying the frequency response by direct control from the beam attenuator. However, since the amplitude of the Ar⁺ write beams are considerably lower for the high frequency modulation conditions, a penalty has to be paid in that the photorefractive response time will be slower and the filter will take longer to write.

Initially, the impulse response of the Fourier transform hologram written to the BSO was calculated for different values of the beam ratio to demonstrate the edge enhancement possible with high values of BR. A simple square of size 65*65 pixels in a 256*256 array was used as the test input object to beam 1. The array for input plane P_2 was set to zero apart from an on-axis pixel set to 255 to represent an impulse function. After Fourier transformation this gives rise to a constant array representing a plane wave HeNe beam addressing the hologram. Figure 2.2 shows cross-sections through the field values of the reconstruction of the square at P_4 for the values of BR indicated. The cross-sections demonstrate that a zero-crossing occurs at the location of a step change in the input function, the change becoming progressively more localised the higher the value of BR. The cross-section resembles the negative of the 2nd differential of the step change. An explanation of this effect can be given as follows. An equivalent of the 2nd differential operation in the space domain is multiplication of the Fourier transform by $-\omega^2$. Further, make the assumption that the amplitude of the spectrum of the input image $u_1(\mathbf{r})$ has a fractal, $1/\omega$, frequency dependence (this is a good approximation for natural images but less so for a more regularly shaped input such as TC3). When the hologram is illuminated by a plane wave beam, U_2 , the field emerging directly behind the filter, $U_4(\mathbf{k})$, may then be written:

$$U_4(\mathbf{k}) = \frac{2U_1^*(\mathbf{k})U_2U_3}{|U_1(\mathbf{k})|^2 + |U_3|^2} = \frac{2U_2U_3 \frac{1}{\omega} \exp(-j2\pi\phi(\mathbf{k}))}{\frac{1}{\omega^2} + |U_3|^2} = \left[\frac{2U_2U_3}{1 + |U_3|^2\omega^2} \right] \omega \exp(-j2\pi\phi(\mathbf{k})) \quad \{2.7\}$$

If the bracketed term can be treated as a constant, $U_4(\mathbf{k})$ will have a linear dependence on ω which is equivalent to $\omega^2 * 1/\omega$ and so its Fourier transform will be the negative of the 2nd differential of $u_1(\mathbf{r})$. This approximation obviously depends on the second term in the denominator of Equation {2.7} being small compared to unity, which will be only true when $|U_3|^2 \ll \omega^2$ ie. for high values of BR. The higher the frequency content of the image, the lower $|U_3|$ must be set to approximate the 2nd differential operation. In practice $|U_3|$ can not be set too low since the areas of the grating in which there is then significant modulation will have a long formation time. In addition, it has been found that diminishing returns in terms of correlation discrimination are achieved by increasing the value of BR too far.

Low pass correlation will give tolerance to in-plane rotation for initial identification purposes whereas a high-pass correlation will yield a sharp, localised response for high discrimination recognition and accurate determination of rotation angle. Thus a valuable feature of the up-dateable channel is the ability to vary the bandpass by adjusting the

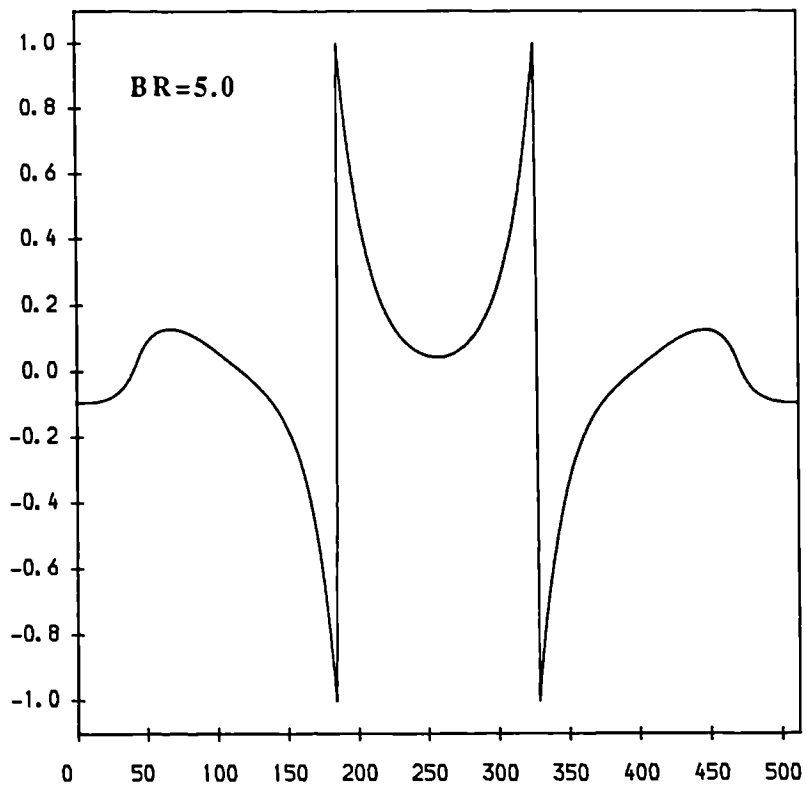
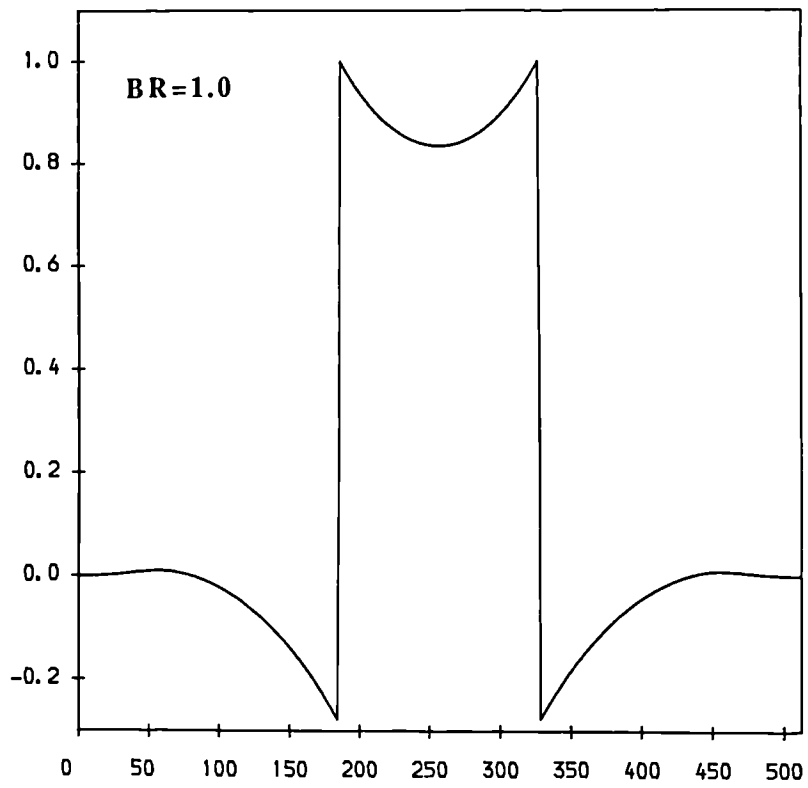


Figure 2.2 Edge reconstructions for the indicated beam ratios

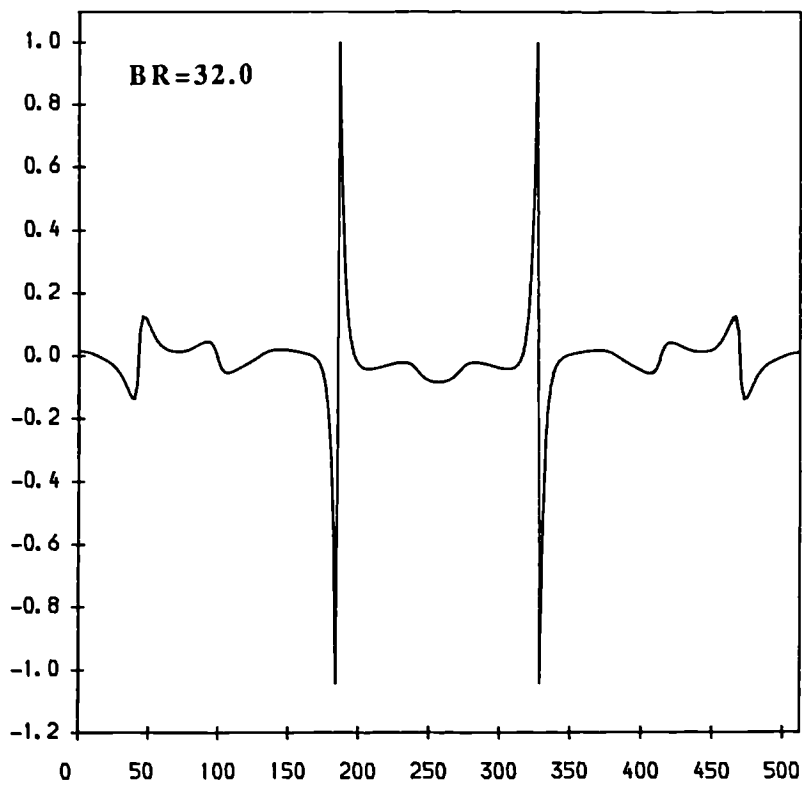
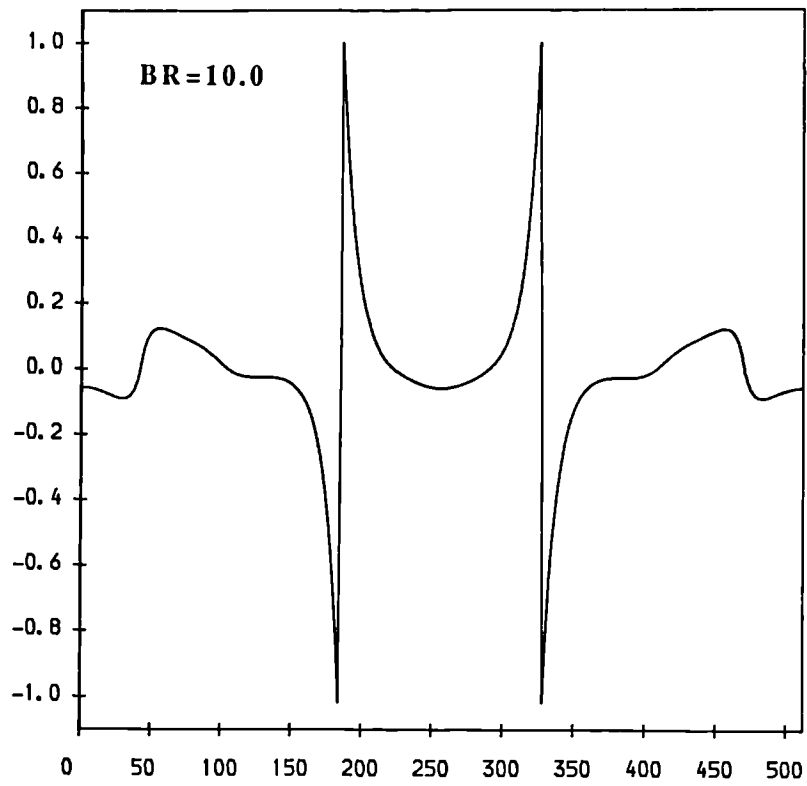


Figure 2.2(cond.) Edge reconstructions for the indicated beam ratios

amplitude of the plane-wave reference beam writing the Fourier transform hologram to the BSO. A range of correlation responses can be generated in this way to increase the flexibility of the technique.

The correlation response of the NDFWM scheme was simulated by taking the Fourier transform of the reference at $u_2(\mathbf{r})$ to yield the frequency domain field distribution $U_2(\mathbf{k})$. The complex multiple of this with the transfer function of the BSO filter, formed by the non-linear interaction of $U_1(\mathbf{k})$ and U_3 in the BSO as previously described, yields an approximation to the field emerging directly behind the filter. Subsequent Fourier transformation of this distribution produces the field amplitude at plane P_4 of Figure 2.1 which is the cross-correlation between the input images $u_1(\mathbf{r})$ and $u_2(\mathbf{r})$. The simulation was run for four values of BR ie. 1.0, 5.0, 10.0 and 32.0 in order to determine the correlation response over a range of frequency weighting conditions. This range encompasses that similar to a matched spatial filter to, at the other extreme, that of a highly discriminatory filter which is equivalent to performing an all-pass correlation on pre-edge enhanced imagery.

The auto-correlation response of the device was initially simulated with TC3 view 3 (see Figure 2.8) being used to modulate both beams $u_1(\mathbf{r})$ and $u_2(\mathbf{r})$. The intensity at the P_4 output plane was computed since this is the quantity measured by a square law detector such as a photodiode or CCD camera. Figure 2.3 shows cross-sections through the auto-correlation intensity peaks of TC3 view 3 for the values of BR indicated. The results are normalised to the maximum ACF peak which occurs for BR=32.0. This is also shown in isometric form in Figure 2.4 for direct comparison with the SDF filter responses shown in Figure 2.12. The ACF peak height is plotted against beam ratio in Figure 2.5. The total correlation plane output energy (defined as the sum of squared pixel values over the plane) is plotted against beam ratio in Figure 2.6. It can be seen from these results that as the hologram is biased towards progressively higher frequencies the total correlation plane energy falls. However, the energy available is concentrated into a progressively more localised correlation response which results in the peak height increasing with BR. The rate of decrease of correlation plane energy appears to fall approximately exponentially with BR, the correlation peak height rising at a roughly one minus exponential rate. When biased to high frequencies a larger area of the HeNe addressing beam, $U_2(\mathbf{k})$, will be strongly diffracted from the grating which is near full modulation in the high frequency regions of the spectrum. However, in contrast to the image reconstruction arrangement, in which a plane wave reads out the frequency plane hologram, the addressing spectrum has an approximately $1/\omega$ amplitude dependence. The larger area of diffraction appears to be more than compensated for by the fall-off in intensity with spatial frequency. Thus an overall reduction in the total

correlation plane energy occurs with increasing values of BR as indicated by Figure 2.6. By Parseval's theorem, this is equal to the energy of the diffracted field directly behind the filter, indicating an overall decrease in diffraction efficiency with BR for this particular optical configuration. Table 2.1 gives a summary of the auto-correlation function properties for the four different settings of BR examined. An important result, as regards a system implementation, is that the auto-correlation peak intensity varies by almost an order of magnitude between BR settings of 1.0 and 32.0. Hence, with alterations of BR setting a means of varying the read beam power to avoid saturation of the CCD sensor is required.

The in-plane rotation sensitivity of the dynamic correlator response was determined for the four values of BR previously examined by computing the cross-correlations for rotated views of TC3. However, small changes in the reflection highlights altered the energy of the inputs for different rotation angles resulting in a change in energy of the inputs between frames. Thus to obtain valid comparisons of the cross-correlation responses between the variously rotated images, each frame was normalised in energy. Figure 2.7 shows the fall-off with rotation angle for the different values of BR. All the results have been normalised to their individual ACF response in order to allow for direct comparison of the fall-off rate between BR settings on the same graph, but it is worth reiterating that the ACF peak for BR=32.0 is almost nine times that for BR=1.0. Table 2.2 summarises the rotation angles for the 90% and 50% fall-off points of the correlation response. For BR=32.0 the fall-off is rapid, the 90% point is within $\pm 1^\circ$, and the 50% fall-off point less than $\pm 3^\circ$. This degree of sensitivity is useful for final object alignment. At the other extreme with the BR=1.0 setting, the 50% point is not reached within the 15° of rotation examined. This allows considerable tolerance to object misalignment for initial recognition purposes. Since the object is to be recognised from a fixed overhead CCD camera, changes in scale of the whole object, in contrast to sub-features, do not have to be accommodated in the initial acquisition phase. Thus the correlator response to scale changes were not modelled.

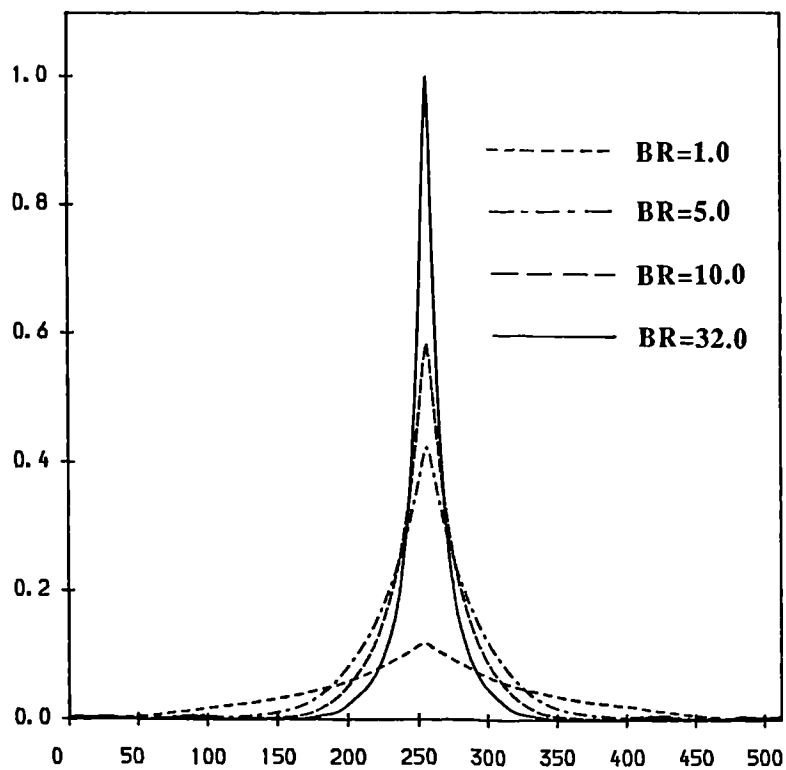


Figure 2.3 Auto-correlation functions of TC3 view 3 for the indicated beam ratios

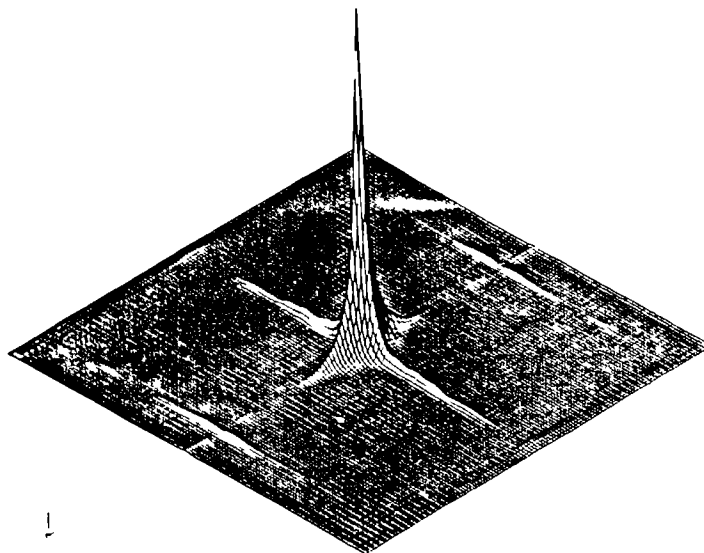


Figure 2.4 Isometric of the auto-correlation of TC3 view 3 with BR=32.0

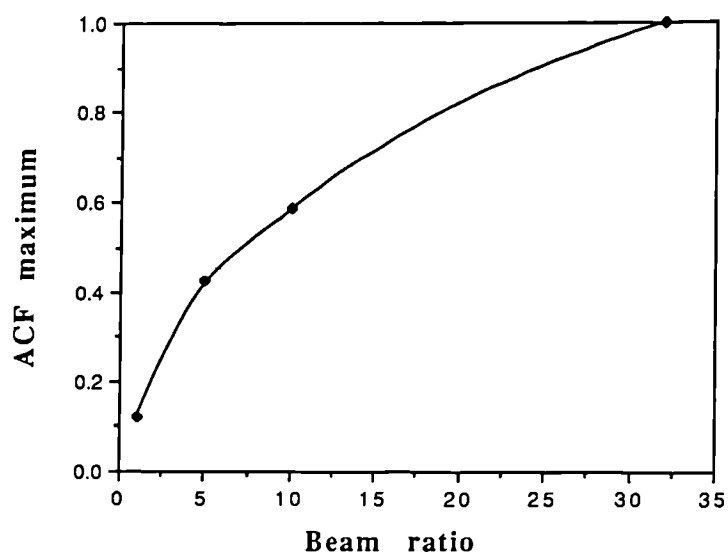


Figure 2.5 Auto-correlation peak intensity variation with beam ratio

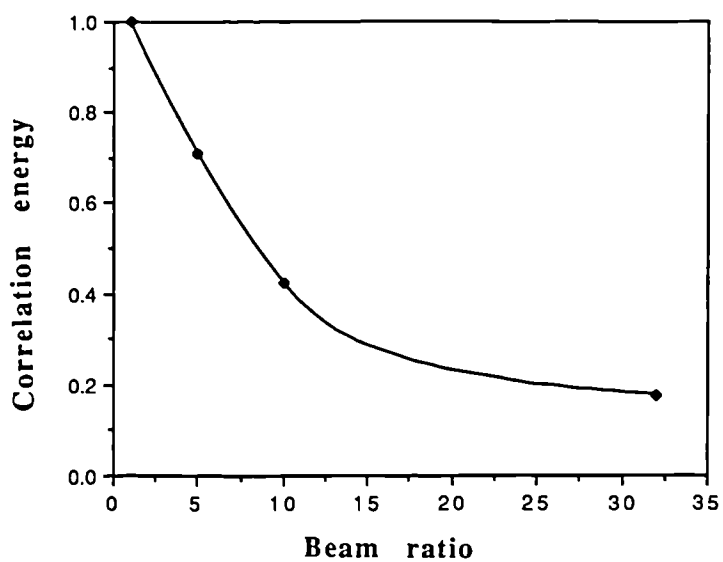


Figure 2.6 Correlation plane energy variation with beam ratio

Amplitude Beam Ratio	1.0	5.0	10.0	32.0
ACF Peak Intensity $\times 10^5$	0.826	2.932	4.041	6.905
ACF Ratio	1.0	3.543	4.883	8.344
FWHM (Pixels)	51	24	17	8
Correlation Plane Energy $\times 10^8$	5.653	3.973	2.433	0.999
Correlation Plane Ratio	1.0	0.703	0.430	0.177

Table 2.1 Summary of auto-correlation properties for the indicated beam ratios

Amplitude Beam Ratio	1.0	5.0	10.0	32.0
90 % Fall-off	3.8°	1.9°	1.2°	0.7°
50 % Fall-off	-	8.1°	5.2°	2.9°

Table 2.2 Summary of correlation peak fall-off with rotation angle for the indicated beam ratios

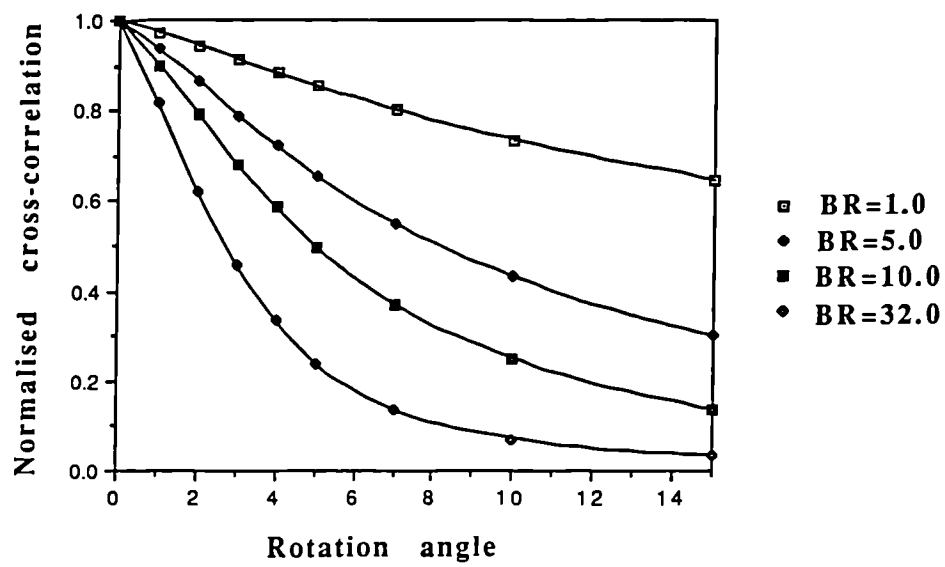


Figure 2.7 Fall in correlation with rotation of TC3 view 3 for the indicated beam ratios

2.3 The synthetic discriminant function filter

2.3.1 Background

The initial task of the inspection system was the identification of a test industrial component from a known set. This requires the determination of its orientation, in-plane rotation and position within the field of view. It is not until the object is constrained in this way that a detailed inspection process can be initiated since this involves moving a robot mounted camera around the component.

In-plane rotation is initially determined by the use of a log. r - θ coordinate transformation of the input image power spectrum followed by correlation. This gives the rotation to one of a number of possibilities governed by the symmetry of the component shape. Orientation determination is simplified by the *a priori* knowledge that the component is resting in one of a number of stable states determined in advance from its CAD model. Thus, for each component, the view presented to an overhead camera is limited to one of these stable state orientations. The seven stable states of TC3 are shown in Figure 2.8 together with a view of TC6 used to demonstrate the discrimination ability of the filter to be described.

In order to expedite the initial recognition phase, multiplexed filter techniques are employed in the up-dateable correlator to reduce the number of sequential searches required. The design was made using the synthetic discriminant function technique [14] so that predictable correlation peak heights were obtained allowing the orientation independent recognition of a particular component. This Section gives a summary of the basic theory behind this technique and a detailed description of how it was applied in the photorefractive based correlator. Simulation results of the performance of the SDF generated for the orientation independent recognition of TC3 are given in the following Section.

A refinement of the matched spatial filter multiplexing technique was originally suggested by Caufield and Maloney [15] and extended and named as the synthetic discriminant function filter by Casasent and co-workers [14]. The main advantage of this formulation over the original multiplexed filter is that it allows the on-axis correlation peak height to be controlled. Discrimination against non-target objects can then be accomplished by including them in the filter training set and assigning their cross-correlation response to zero. The theory of the SDF filter is based on the fact that an input to the correlator, that is a member of the training set used to create the filter, will correlate not only with itself but to some greater or lesser degree with all other members

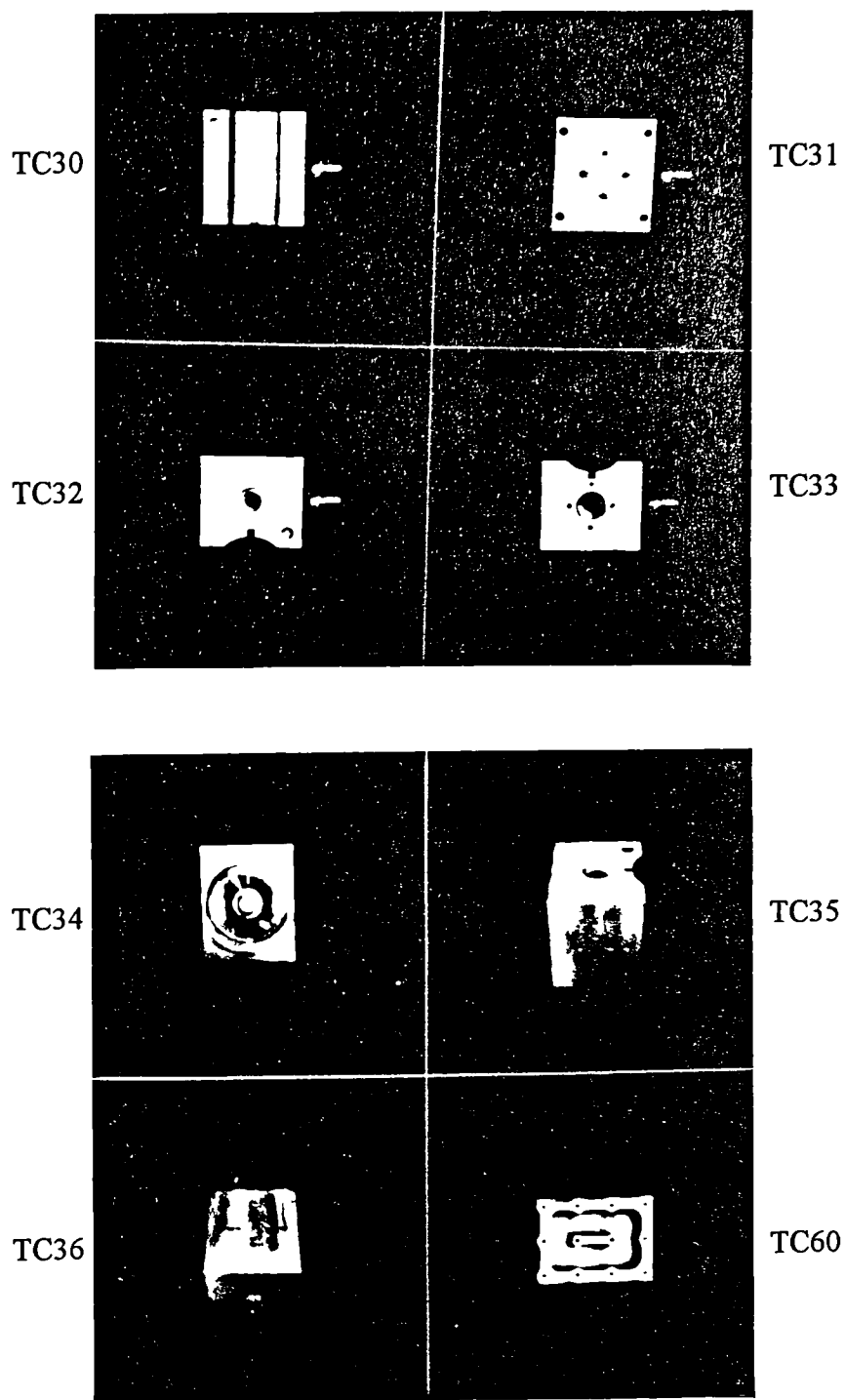


Figure 2.8 Stable state views of TC3 and one view of TC6

of the training set stored in the filter. Since the SDF seeks only to control the on-axis correlation height, the filter response can be specified as the vector inner product (VIP) between the space domain representation of the filter function, g , and the input training set image, h_i . Each pixel can then be regarded as a component of the $d \times d$ dimensional vector generated from the image by lexicographical reordering. An SDF satisfying the condition $h(\mathbf{r}) \otimes g(\mathbf{r}) = (\text{user specified on axis value})$ can be determined for a training set of N images and hence an associated symmetric $N \times N$ correlation or VIP matrix, r , with the equation:

$$\mathbf{a} = r^{-1} \mathbf{u} \quad \{2.8\}$$

where the column vector \mathbf{u} has N elements each specifying the on-axis correlation heights for input image $h_i(\mathbf{r})$. \mathbf{a} is also a column vector, the elements of which are the weighting coefficients in the SDF:

$$g(\mathbf{r}) = \sum_{i=1}^N a_i h_i(\mathbf{r}) \quad \{2.9\}$$

The $h_i(\mathbf{r})$ training set images are the different distorted versions of the object to be recognised. Thus the SDF differs from the original multiplexed MSFs in that the training images are weighted by the a_i coefficients during filter synthesis. This ideally results in an equal on-axis correlation response for all members of the training set. The appropriate weightings are achieved by multiplying the user specified correlation values by the inverse of the cross-correlation matrix as shown by Equation {2.8}. This accounts for the contributions to the correlation response of the filter from the cross-correlations of a particular h_i with all other members of the training set.

The various SDFs can be unified with a notation developed by Vijaya Kumar [16]. Images are treated as d^2 dimensional column vectors $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_N$ obtained by the concatenation of the pixel elements; this allows the use of methods from vector and matrix algebra. Rewriting the SDF formulation of Equation {2.9} as a vector sum:

$$\mathbf{g} = \sum_{i=1}^N a_i \mathbf{h}_i \quad \{2.10\}$$

The composite image vector, \mathbf{g} , is thus a linear combination of the training vectors weighted such that :

$$\mathbf{g}^T \mathbf{h}_j = u_j \quad j=1,2,\dots,N \quad \{2.11\}$$

is satisfied.

The a_i values can be determined by substitution of {2.10} into {2.11} to obtain:

$$\sum_{i=1}^N a_i (\mathbf{h}_i^T \mathbf{h}_j) = u_j \quad j=1, 2, \dots, N \quad \{2.12\}$$

where $\mathbf{h}_i^T \mathbf{h}_j$ is the vector inner product between the i^{th} and j^{th} training images. Thus Equation {2.12} can be written as:

$$\mathbf{r} \mathbf{a} = \mathbf{u} \quad \{2.13\}$$

as in {2.8} above where $\mathbf{a}=(a_1, a_2, \dots, a_N)^T$ and $\mathbf{u}=(u_1, u_2, \dots, u_N)^T$. If it is assumed \mathbf{r} is invertible the coefficients of \mathbf{a} for the SDF can be obtained directly.

Vijaya Kumar further compresses the notation by defining the matrix \mathbf{h} with $d \times d$ rows and N columns, the i^{th} column being given by \mathbf{h}_i . The SDF composite image can then be written:

$$\mathbf{g} = \mathbf{h} \mathbf{a} \quad \{2.14\}$$

where the coefficients of vector \mathbf{a} satisfies:

$$(\mathbf{h}^T \mathbf{h}) \mathbf{a} = \mathbf{u} \quad \{2.15\}$$

$(\mathbf{h}^T \mathbf{h})$ being equivalent to \mathbf{r} .

An extension of the original SDF result, with which it can be unified using the vector-matrix notation, is the minimum variance SDF due to Vijaya Kumar. The original SDF is not guaranteed to give the specified on-axis correlation value $u_i(\mathbf{o})$ in the presence of noise in the input. The minimum variance SDF can be expressed as:

$$\mathbf{g}_{\text{mv sdf}} = \mathbf{c}^{-1} \mathbf{h} (\mathbf{h}^T \mathbf{c}^{-1} \mathbf{h})^{-1} \mathbf{u} \quad \{2.16\}$$

where the additive noise in the input is taken as zero mean with co-variance matrix c . This solution can be shown to yield the smallest variance in the output for any given constraint vector \mathbf{u} [16]. Comparing {2.16} to the expression for the conventional SDF:

$$\mathbf{g}_{\text{conv}} = h(h^T h)^{-1} \mathbf{u} \quad \{2.17\}$$

it can be seen that this is a special case of the MVSDF when the noise in the input is white and the covariance matrix reduces to the identity matrix [16]. The filter will thus give minimal variation in u_i from the specified values under these input conditions. The more general case of coloured noise requires the inclusion of its covariance matrix in the filter synthesis procedure. The resulting filter then becomes a linear combination of appropriately pre-processed training set images.

These SDF filter solutions, however, suffer from a number of drawbacks for the general pattern recognition problem. Due to the omission of any high frequency weighting in the filter synthesis procedure, the correlation response is similar to that of a conventional MSF. This leads to very broad multiple peaked correlation plane responses from the SDF filter. The reason for this is that, of the total $N(N-1)$ cross-correlations possible for an N training set image filter, only $2(N-1)$ of these can be centred by appropriately shifting the training images to produce an on-axis correlation response. This results in the remaining cross-correlation peaks being located off the central axis. Since the SDF formulation only controls the on-axis pixel value and leaves the rest of the correlation plane unconstrained, these sidelobe values become significant compared to the specified on-axis correlation value and confound correct classification of the input. This problem is particularly severe when the central pixel value has been constrained to be zero for the purposes of out-of-class object rejection.

Within the same general framework as the SDF, but using a different optimisation technique, a filter formulation known as the minimum average correlation energy filter (MACE) has been proposed by Mahalanobis *et al.* [17]. These authors use the same notation as previously but apply the analysis in the frequency domain, considering the field directly behind the filter plane. Thus if the correlation of the i^{th} image sequence, $h_i(\mathbf{r})$, with the filter is denoted by $s_i(\mathbf{r})$, ie. $s_i(\mathbf{r}) = h_i(\mathbf{r}) \otimes g(\mathbf{r})$, the discrete Fourier transform (DFT) of this, $S_i(\mathbf{k})$, describes the field distribution in the filter plane. The energy of the i^{th} correlation plane is equal to this and hence given by [17] :

$$E_i = \frac{1}{d^2} \sum_{\mathbf{k}} |G(\mathbf{k})|^2 |H(\mathbf{k})|^2 \quad \{2.18\}$$

Equation {2.18} may be more compactly written:

$$\mathbf{E}_i = \mathbf{G}^+ D_i^{-1} \mathbf{G} \quad \{2.19\}$$

where the superscripted cross indicates the conjugate transpose of the complex vector \mathbf{G} representing the filter transfer function. D_i is a $d^2 \times d^2$ diagonal matrix where the elements are:

$$D_i(k, k) = |\mathbf{H}(k)|^2 \quad \{2.20\}$$

This energy function is averaged over the individual images comprising the training set by replacing the elements of D_i by the sum of their power spectra at each pixel location, ie. $D(k, k) = \sum |\mathbf{H}_i(\mathbf{k})|^2$. From Parseval's theorem this will then represent the average correlation plane energy produced by the filter for the training set images. By the method of Lagrange multipliers the average energy is minimised subject to the SDF constraint:

$$\mathbf{H}^+ \mathbf{G} = \mathbf{u} \quad \{2.21\}$$

from which the frequency plane solution vector is determined to be [17]:

$$\mathbf{G} = D^{-1} \mathbf{H} (\mathbf{H} D^{-1} \mathbf{H})^{-1} \mathbf{u} \quad \{2.22\}$$

The Fourier transform of this result yields the space domain formulation of the filter and can be seen to be of the same general form as that shown in Equation {2.16}. Thus the MACE filter can be considered as a conventional SDF operating on an image pre-processed by multiplying by the inverse of D . This acts to flatten the input spectrum. Mahalanobis in fact shows that the MACE filter can be considered as a cascade of a pre-whitening filter and a conventional SDF based on a weighted linear superposition of the training set data. Vijaya Kumar's result shows that the minimum variance filter, that most accurately maintains the assigned correlation peak heights for a particular type of input noise, is a cascade of the whitening filter for that noise and the conventional SDF based on the transformed data. The MACE filter response will only be undegraded in the presence of input noise if this is included in the calculation of D , so allowing D^{-1} to act as a pre-whitening filter for the noise corrupted training set input image.

The MACE formulation attempts to produce a uniform field distribution behind the filter which will Fourier transform to an on-axis impulse function in the correlation plane, all other points being zero. Thus the correlation plane energy has been minimized for the specified correlation peak height of u . This is accomplished by flattening of the input power spectrum by the MACE filter. When more than one training image is included in the filter, the results presented by Mahalanobis *et al.* [17] show that the localised nature of the correlation response remains good. Only a slight broadening of the peak occurs, with the background noise in the rest of the correlation plane remaining low. Discrimination against non-targets is achieved by including them in the filter training set with an associated constraint of zero in the response vector \mathbf{u} . This forces the on-axis point of the correlation plane to be zero as in the conventional SDF. However, in the MACE filter the sidelobe disturbance is very much reduced, as compared to the SDF, due to the very much more localised nature of the correlation response.

2.3.2 SDF filter design considerations

The MACE filter formulation does, however, have its own associated difficulties. It will be recalled from the discussion of the conventional SDF above, that the sidelobe problem arose because the input image cross-correlated with the similar training set images stored in the filter. These cross-correlations can not all be centred and are broad, so leading to a multiple peaked, very broad correlation response from the SDF. The MACE filter overcomes this problem by maximally high-pass filtering the images prior to correlation (ie. by completely flattening the object spectrum). This achieves a very localised correlation response. However, part of the reason the correlation response can be so localised is that, due to the high-pass filtering, the input object will only correlate strongly with the exactly corresponding view stored in the filter. All other cross-correlations with other members of the training set are negligible. This is undesirable for the orientation independent recognition of a 3-D object such as a military vehicle. If the vehicle orientation does not closely match one included in the training set, correlation will be lost. Thus many training images are required with only small increments of viewpoint between each. This leads to problems of storage on a single holographic plate due to the limited dynamic range of this medium. The filter will also be very sensitive to in-plane rotation and scale changes of the input object.

A further difficulty of the MACE filter lies in its actual design which involves the inversion of the diagonal matrix D . Each element of this matrix is the sum of the pixels at the k^{th} location of the power spectra of the training set images

(ie. $D(k, k) = \sum |H_1(\mathbf{k})|^2$). For this matrix to be invertable no elements can be zero. However, in the high frequency regions of the spectra there will be typically very many near zero pixels and the resulting sum in this region may well be very small. This will result in problems in forming the inverse of D . The physical reason for this mathematical difficulty is clear. The filter is attempting to flatten the average power spectrum of the training set data. The filter is, however, a passive device and so can only do this by absorbing spectral energy. Thus if there are zero or low value regions in the averaged spectrum it is not possible to produce a uniform field directly behind the filter plane. In all cases the filtering will involve considerable attenuation of the low frequency components of the spectrum. This reduces these high intensity regions to the same value as the relatively weak high spatial frequency regions of the spectrum. This inevitably means that the light throughput efficiency of an optical correlator based on the MACE filter will be low. However, since the available energy in the correlation plane will be concentrated into a very localised correlation response this may not be serious problem. It will be remembered that a similar situation occurs as the beam ratio is set to higher values when writing the BSO filter. Although the overall correlation plane energy decreases with increasing values of BR, the progressively more localised correlation peak increases in intensity by almost an order of magnitude.

Thus the MACE filter algorithm attempts to produce as uniform a field distribution as possible behind the filter plane. Subsequent Fourier transformation to the correlation plane then results in a maximal localisation of the correlation response. This should approach the diffraction limit of the point spread response for the optical system. However, this extreme leads to practical difficulties in the filter's implementation and, in fact, is not required for the production of a sufficiently localised correlation response to avoid high correlation plane energies for non-target training set members. Réfrégier has performed an analysis which seeks a trade-off in the optimisation of output variance and correlation plane energy [18,19]. The MVSDF and MACE filters are seen as the two extremes in the independent optimisation of these two parameters. An important point is that a very much more localised correlation response can be achieved, as compared to a conventional SDF, even though the spectrum emerging from behind the filter is by no means flat. This is also seen clearly by comparing the auto-correlation response of the phase-only filter of Horner and Gianino [20] in which the field behind the filter falls off at $|H(\mathbf{k})|$ rather than $|H(\mathbf{k})|^2$ (as in the conventional matched spatial filter). A very much more localised auto-correlation response is produced in this way.

Similarly, an SDF used in combination with the BSO filter at a high BR setting can be used to create an SDF with a localised peak response. The field emerging from behind

the BSO filter is certainly not flat. Figure 2.9 shows this field distribution for BR settings of 1.0 and 32.0. These cross-sections should be compared to their corresponding Fourier transforms, ie. the ACFs for the same BR values. This indicates the very marked increase in peak localisation that occurs due to the increased energy in the high frequency regions of the spectrum. For the higher beam ratio the ACF is sufficiently localised to give minimal sidelobe disruption when used as part of an SDF.

It is important to mention a difference between the POF and the BSO filter. The POF does not attenuate the spectrum of the input object. Since the impulse response of the POF is an edge enhanced version of the stored reference image, the action of a correlator employing such a filter will, therefore, be the correlation of an unprocessed input with an edge enhanced reference. This should be compared to the BSO filter (or the extreme case of the MACE filter) in which the transfer function of the filter attenuates low frequency regions of the addressing spectrum for high BR settings (see Equation {2.6} and the following discussion). This has the important effect of inhibiting false cross-correlations, since the matching low frequency components from non-target objects are prevented from passing through the filter to the correlation plane. Here they can only degrade the SNR.

Since the discriminatory ability of the high-pass BSO filter is good, except in the case of very similar inputs with many common edges, the only training images that need be included in the SDF design are those for which positive identification is required. This is in contrast to the inclusion of non-target objects in the training set and setting the SDF response to be zero for these views. This is valuable since it reduces the storage requirement from that of the conventional SDF and so allows the greyscale of each training image to be more effectively represented. This is particularly important when the device used for the SDF storage has a limited greyscale capability, as is the case in a photorefractive based up-dateable correlator where the SDFs are input to the system via an space domain SLM. In addition, it avoids the negative a_i coefficients which result when correlation responses are suppressed by the SDF and which are difficult to realise when the SDF is written to an SLM. For these reasons the SDF design is required to contain only the component views for which positive identification is required, discrimination against different components being sought by the high-pass nature of the filter when it is designed using high BR values.

As mentioned previously, component recognition is only required for a limited number of stable state views. This facilitates the use of a high-pass filter since all these views can be explicitly included in the training set. There is no requirement for full 4π steradian orientation independent recognition. Thus an SDF can be designed for each component

to be inspected by the system. Included in the training set for each SDF are all possible views of the component from an overhead camera position when it is resting in any of its stable states. Initially the elements of the 7*7 cross-correlation matrix, R , for the seven stable states of TC3 were calculated. The on-axis cross-correlation value for a direct correlation is simply the vector inner product of the two input images, as detailed above. However, in the present case, the full cross-correlation must be calculated using the model outlined in Section 2.2.1 since the transfer function of the BSO filter for a particular BR setting must be accounted for. The R matrix as initially calculated is shown in Table 2.3. The entries in bold on the leading diagonal are the ACFs of the 7 views of TC3, the maxima of which occur on-axis, ie. the (128,128) data point of the unsupplemented 256*256 correlation plane array. Ideally, for the high-pass filtered SDF, matrix elements off the leading diagonal ie. the cross-correlations between the component views, should be zero. However, due to the large commonality of edges between the different faces of TC3 some cross-correlation heights are significant. This increases the computational requirement in designing the filter. If the cross-correlation terms were low enough to be ignored only 7 ACFs would need to be calculated and the resulting diagonal matrix could be readily inverted. However, as can be seen from Table 2.3, high cross-correlation values occur due to the similarity of some of the TC3 views. Also, in contrast to a conventional SDF matrix which is symmetric about the leading diagonal, the SDF matrix created using the BSO filter is asymmetric, requiring all 49 entries to be calculated explicitly. $R_{hg} \neq R_{gh}$ because the energy between any pair of views is not the same, requiring the amplitude of the reference beam to be reset when each member of the pair rewrites the BSO grating. In this way the BR value is kept constant at 32.0. This results in a different filter modulation in each case and hence an altered correlation response.

The high cross-correlation values compromise the performance of the filter with regard to the localisation of the correlation response that can be achieved. These can occur off the central axis and hence lead to a broadening of the SDF correlation peak. To minimize this effect the training images comprising the SDF must be shifted so as to centrally locate the cross-correlation responses. However, this can only be done with six pairs of cross-correlations since there are only six training images to align with any arbitrarily chosen member of the training set. The most effective result can be achieved if this is done with the six highest cross-correlation pairs. These values are shown italicised in Table 2.3. Those which cannot be aligned will degrade the peak quality and increase the background noise levels away from the peak. This is particularly so for cross-correlations that result in a multiple peak structure such as that between TC33 and TC31, which is shown in Figure 2.10. Even though this problem cannot be avoided the

SDF peak quality is, nevertheless, much better than that of the conventional SDF. This is due to the increased localisation of the ACFs and maximal reduction of the unwanted cross-correlation terms by way of the high-pass filtering effect of the filter. The simulation results demonstrate that the technique is capable of giving a sufficiently good quality correlation response to allow the recognition and accurate location of the component in the input camera field of view.

The bracketed numbers in Table 2.4 show the location of the cross-correlation peak heights for the corresponding matrix elements in Table 2.3, the six highest pairs being aligned on axis. However, the matrix shown in Table 2.3 is not that required for inversion since, as emphasised earlier, the technique only controls the on-axis (128,128) peak height. Thus the contributions to *this* point of the cross-correlation terms are the required entries for the SDF matrix. These values are shown in Table 2.5.

In summary the steps required for calculating the SDF are:

1. Calculate the $N \times N$ cross-correlation values between the N training images of the SDF using the filter transfer function given by Equation {2.2} with a high BR setting.
2. Align the training images such that the largest $(N-1)$ cross-correlation pairs occur on-axis.
3. Re-compute the on-axis contributions of the remaining $N(N-3)+2$ unaligned cross-correlation values.
4. Invert the resulting matrix and multiply by the assigned correlation values, u_i , to determine the SDF weights, a_i .

This procedure results in the SDF weights, shown in Table 2.6 for TC3, such that all views give an equal correlation peak height. It can be seen that all the a_i coefficients are positive, allowing direct implementation of a space domain SDF with an intensity modulating device such as an SLM. However, it should be noted that the smallest coefficient, a_2 , is over eight times lower than the largest, a_3 . Since the constancy of the correlation peak obtained from the SDF will depend on how accurately the individual training images are weighted, a high demand will be placed on the grey-level capability of the SLM if the optical implementation is to work well.

The space domain SDF containing the seven stable state views of TC3 is shown in Figure 2.11. The weights of Table 2.6 have been applied to the individual views and each has been appropriately shifted in order to align the highest six pairs of cross-correlations on-axis. (Unfortunately, the photograph of the IBM 5080 graphics terminal display does not have sufficient dynamic range to show the SDF well.)

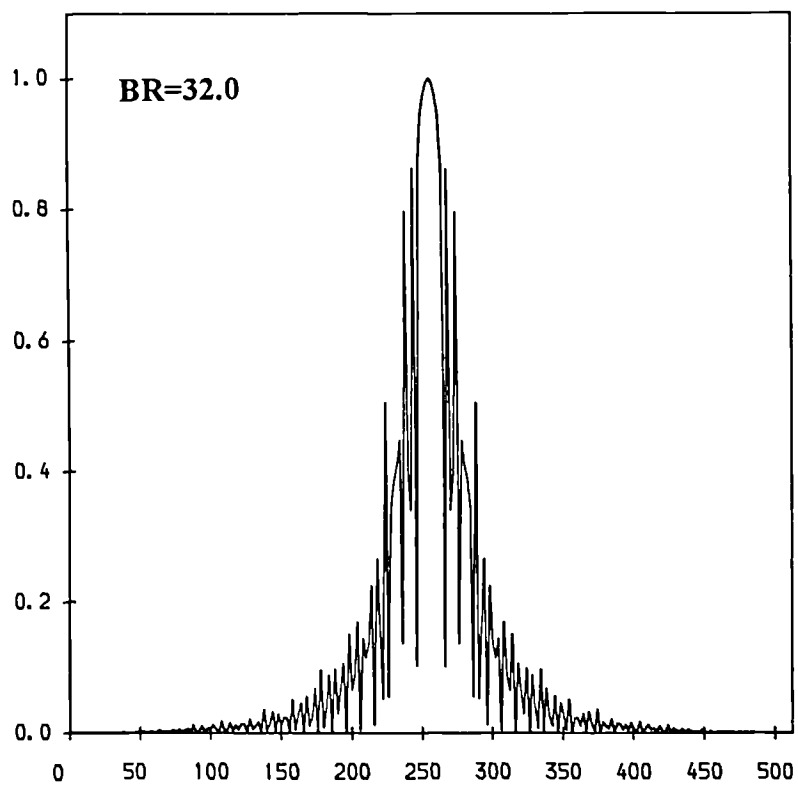
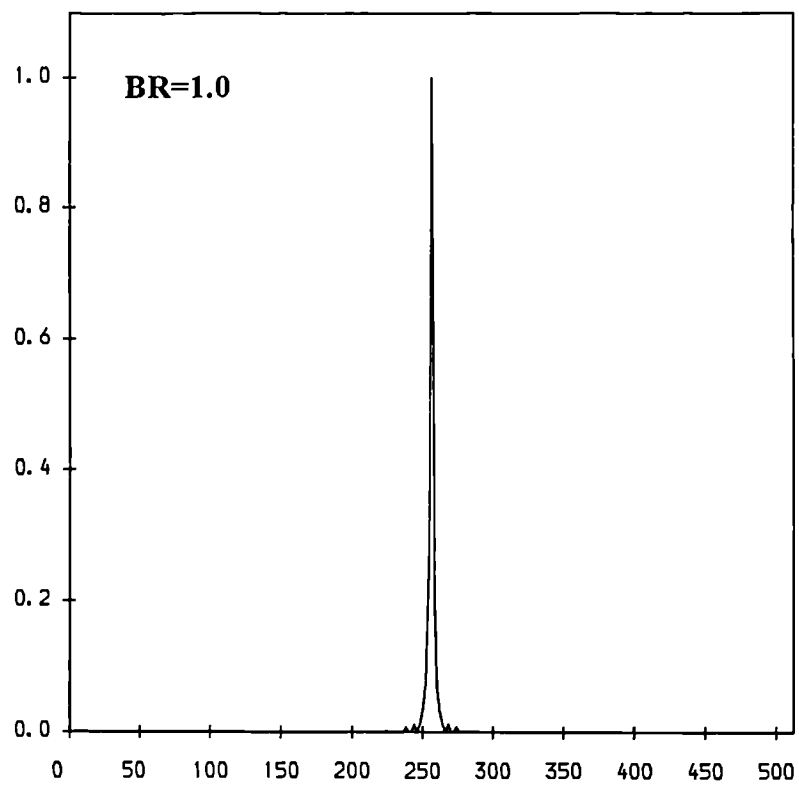


Figure 2.9 Field amplitudes, $U_4(k)$, directly behind filter plane

View	TC30	TC31	TC32	TC33	TC34	TC35	TC36
TC30	725	645	334	296	361	226	226
TC31	599	803	341	297	387	233	233
TC32	346	395	789	542	336	201	210
TC33	286	335	523	831	237	150	173
TC34	382	442	384	255	738	302	287
TC35	257	309	285	194	342	680	347
TC36	306	354	300	248	380	366	559

Table 2.3 Cross-correlation maxima (amplitude) for TC3 stable state views

View	TC30	TC31	TC32	TC33	TC34	TC35	TC36
TC30	(128, 128)	(128, 128)	(128, 128)	(128, 129)	(128, 127)	(136, 126)	(131, 127)
TC31	(128, 128)	(128, 128)	(128, 128)	(128, 98)	(128, 128)	(135, 127)	(134, 127)
TC32	(128, 128)	(128, 128)	(128, 128)	(128, 128)	(142, 127)	(136, 127)	(123, 127)
TC33	(128, 137)	(128, 138)	(128, 128)	(128, 128)	(122, 160)	(137, 165)	(123, 116)
TC34	(128, 129)	(128, 128)	(134, 129)	(134, 130)	(128, 128)	(126, 128)	(128, 128)
TC35	(118, 129)	(120, 129)	(120, 129)	(120, 90)	(130, 128)	(128, 128)	(128, 128)
TC36	(127, 129)	(125, 129)	(132, 129)	(133, 140)	(128, 128)	(128, 128)	(128, 128)

Table 2.4 Cross-correlation peak locations for TC3 stable state views

View	TC30	TC31	TC32	TC33	TC34	TC35	TC36
TC30			334	295	352	178	200
TC31				283		190	211
TC32	346				231	167	164
TC33	255	277			106	56	91
TC34	369		282	154		292	
TC35	195	256	247	137	333		
TC36	290	339	254	154			

Table 2.5 On-axis contributions of un-aligned cross-correlations

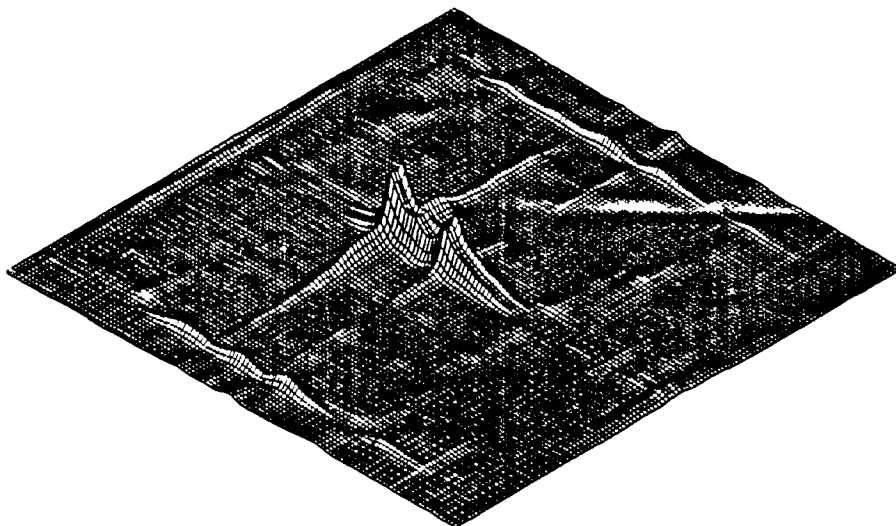
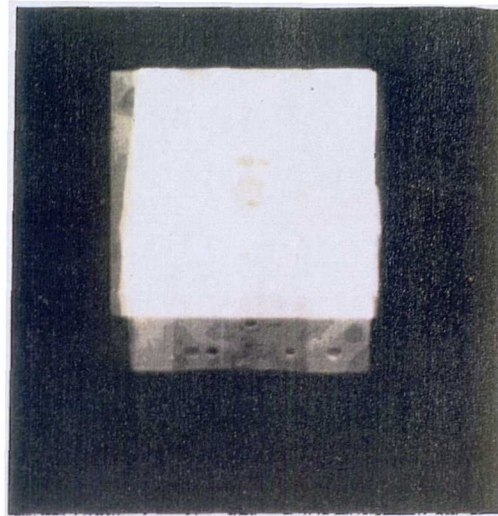
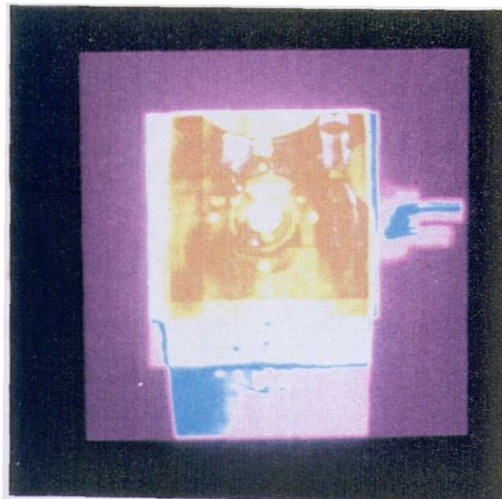


Figure 2.10 Cross-correlation of TC3 view 3 and TC3 view 1 normalised to the ACF of TC3 view 1



Grey scale image



Pseudo-colour coded image

Figure 2.11 Space domain synthetic discriminant function (SDF)

2.4 Simulation results of bandpass SDF

Simulations were performed to determine the correlation performance of the SDF with each individual training image in turn. The input image was used to modulate the Argon ion beam with BR set to the value used in the filter design, ie. 32.0, while the SDF modulated the HeNe beam. This is the configuration that would be used in practice since it allows maximum shift invariance in the correlator, the input image modulating the maximum field of view beam. Also, importantly, in this configuration the SDF will produce constant correlation peak heights for each training set image. This is because in its design a constant BR setting of 32.0 was used to calculate the elements of the cross-correlation matrix. Thus BR must be reset to this ratio when each new image is input to the correlator. If, instead, the SDF is used to modulate the Argon ion beam, with the reference beam set to a constant value, the variable energies of the unknown inputs modulating the HeNe beam will disrupt the constancy of the correlation plane peak heights and so cause the SDF to fail.

The simulation results are summarised in Table 2.7. It can be seen that the correlation peak intensities are on-axis and within a fraction of a percent of each other. However, it should be noted the the peak intensity is 0.38 of the minimum ACF of the training set, which is that of TC36, and only 0.17 that of the maximum ACF which is that for TC33. Further, it must be emphasised that the space domain SDF employed in the simulation used eight bit pixels (ie. 256 grey level), to represent the weighted training images. Since the SLM used in the experimental implementation of the SDF has an equivalent greyscale capability considerably less than this, the constancy of the SDF correlation response is expected to be degraded from the simulation results which must be regarded only as an ideal response. The grey level realisable in practice can be readily incorporated into the simulation and its likely effect on the SDF performance quantified.

Figure 2.12 (a) and (b) show isometric plots of the SDF cross-correlation response with two typical TC3 stable state views. These should be compared to the isometric of the ACF of TC33 for BR=32.0 shown in Figure 2.4. This comparison reveals that the correlation response of the SDF has been broadened somewhat from the ideal of the single ACF. The FWHM is 15 data points for the correlation of TC35 and 11 data points for that of TC33 as compared to that of the ACF which is 8 data points. In addition, the sidelobe structure of the correlation peak is more evident in the output from the SDF, particularly that of TC35. The peak broadening and increased sidelobe structure are due to the cross-correlation terms that remain unaligned. Examination of Table 2.4 reveals that the cross-correlation peak of TC33 and TC35 occurs at location (120,90) and contributes to the high sidelobe value evident in the isometric of the SDF

View	SDF Coefficient a_i
TC30	0.2890
TC31	0.1347
TC32	0.0823
TC33	0.6865
TC34	0.3438
TC35	0.5198
TC36	0.4545

Table 2.6 Calculated SDF coefficients for TC3

View	TC30	TC31	TC32	TC33	TC34	TC35	TC36
CCF Peak Intensity $\times 10^{-5}$	1.199	1.198	1.197	1.198	1.200	1.198	1.199
Peak Location	(128, 127)	(128, 128)	(128, 128)	(128, 128)	(128, 128)	(128, 128)	(128, 128)

Table 2.7 SDF responses to the training set images (TC3 stable state views)

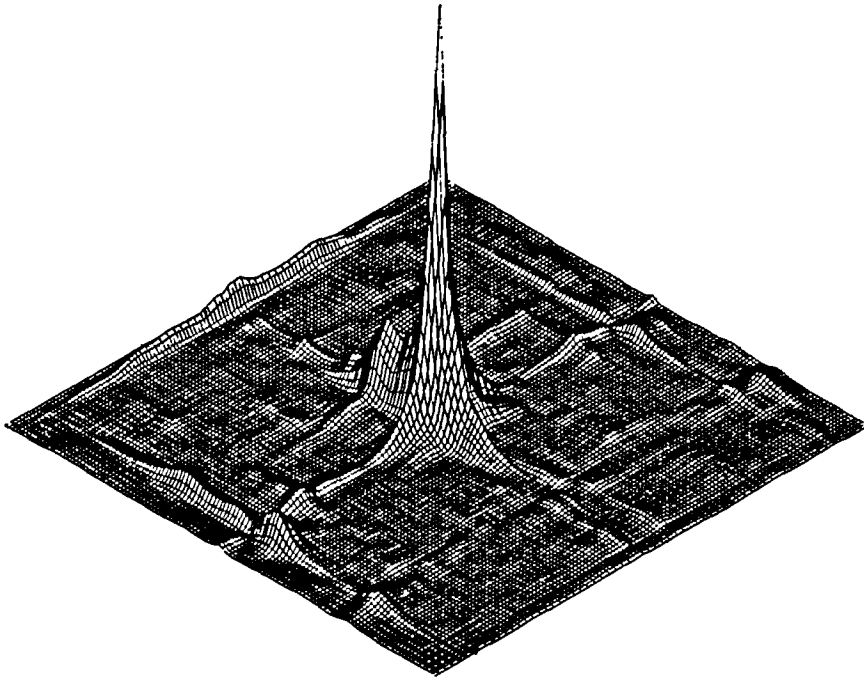


Figure 2.12(a) Isometric of SDF response to TC3 view 3

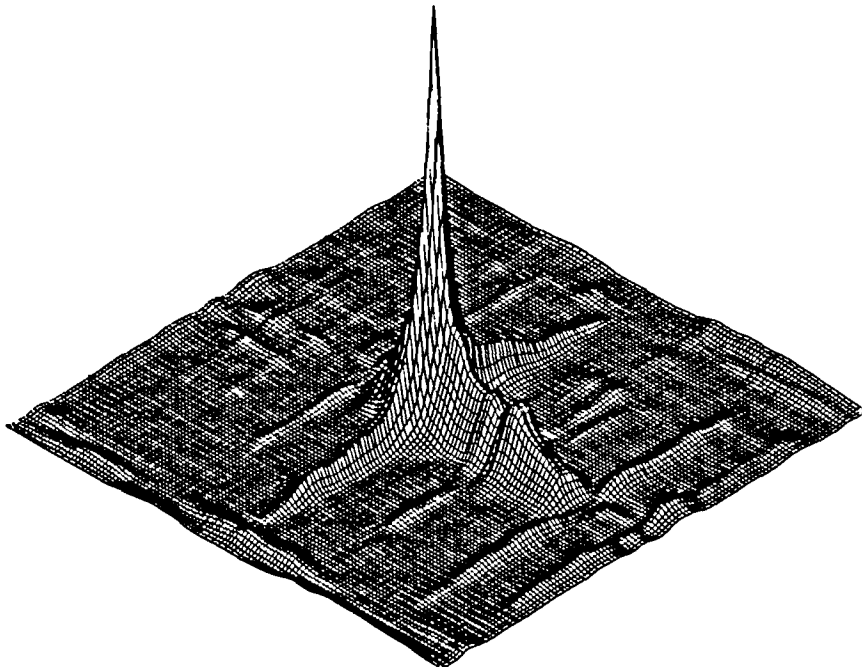


Figure 2.12(b) Isometric of SDF response to TC3 view 5

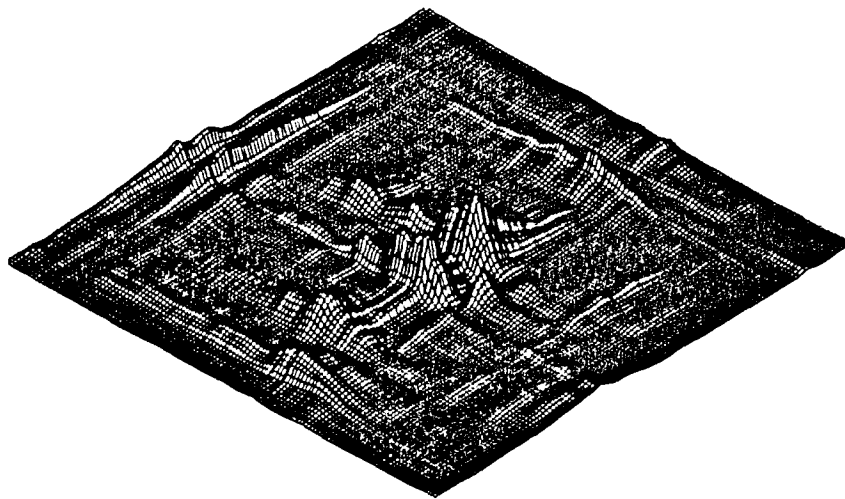


Figure 2.13(a) Isometric of SDF response to TC6 view 0

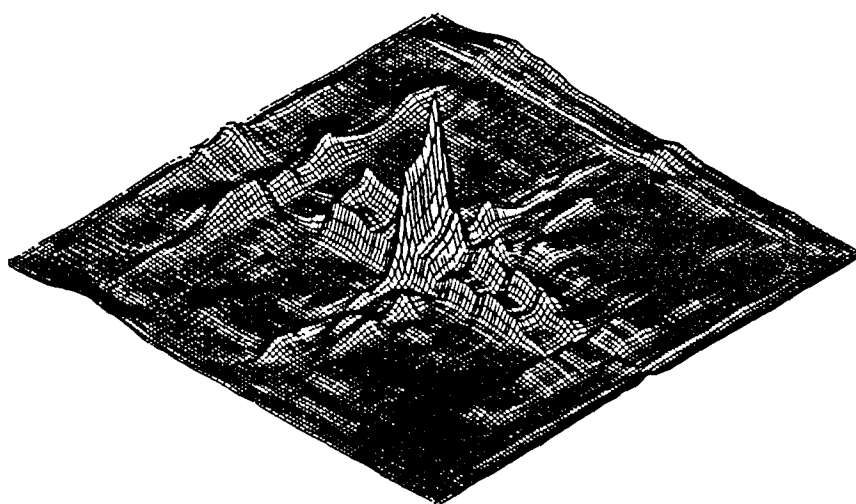


Figure 2.13(b) Isometric of SDF response to TC6 view 0, rotated 90°

response for the TC35 input. The noise structure over the correlation plane away from the central peak can be seen to be greater in Figure 2.12 than that in the auto-correlation plane of Figure 2.4. This is at least partly due to the fact that the SDF correlation peak intensity is only a sixth of that of the ACF. Thus the correlation peak has fallen towards the noise floor of the correlation plane leading to an overall degradation in the SNR.

An isometric of the cross-correlation response of the SDF filter to TC6 is shown in Figure 2.13(a). This has been normalised with respect to the SDF recognition correlation peak height to indicate the discrimination ability of the filter without inclusion of TC6 in the training set. The maximum peak height is 0.11 of the SDF peak. However, the discrimination is poorer when TC6 is rotated by 90°. The isometric of the resulting cross-correlation is shown in Figure 2.13(b). Discrimination is still possible but the peak height has risen to 0.23 of the SDF recognition peak. This is due to the fact that TC6 is very similar in size and shape to some views of TC3, particularly view 4. Thus the resulting large overlap in edge location makes even a high-pass correlation give a relatively strong cross-correlation signal. Since this peak occurs at location (129,128), it could be effectively suppressed by including TC6 in the filter training set and setting its response to zero. However, this was not done since adequate discrimination is possible without this measure and the negative a_i coefficients that this would introduce are thus avoided. Negative a_i coefficients would result in an increased grey-level requirement on the input SLM since a bias would have to be added to make the space domain SDF image all-positive (and subsequently removed in the frequency plane to maintain the assigned SDF responses).

2.5 Conclusion

Thus, in summary, an SDF constructed using the high frequency biased BSO filter is capable of view independent recognition of the BRITE/EuRAM project test components. Its design requires moderate amounts of CPU time but is straightforward without any analytical difficulties. As mentioned previously, its characteristics are in many ways appropriate to the needs of the task since its discriminatory ability is sufficiently high to avoid the difficulty of having to include non-target objects in the training set data. This reduces the computational requirements in synthesizing the filter and avoids negative a_i coefficients. For the more general pattern recognition task of the detection of a 3-D object, when totally unconstrained in orientation, the filter will suffer from lack of tolerance to viewpoint misalignments due to its high-pass characteristic. For the present application, however, the test components are constrained to rest in known stable states, so presenting a constant viewpoint angle to an overhead camera.

The only unconstrained variable is in-plane rotation, to which the SDF will have the same sensitivity as that given in Table 2.2 for the single high-pass filter with BR set to 32.0. This indicates that a 10% fall-off in correlation peak height should be expected for a $+1^\circ$ rotation misalignment. Thus the SDF is capable of providing accurate component alignment information.

Further work on the SDF implementation in a photorefractive based correlator should concentrate on increasing the filter's tolerance to unexpected component orientations. These occur due to faults on the inspected components causing stable state orientations that are different from those anticipated from the CAD model of the ideal component. Thus the filter must be able to accommodate some degree of orientation independent recognition whilst maintaining inter-class discrimination. For this to be possible the high-pass characteristic of the filter must be reduced, since this leads to over sensitivity to view point mis-alignment. The filter weights could be calculated with BR set to a lower value than 32.0 since, as can be seen from Figure 2.3, peak localisation remains good for BR values down to about 10.0. However, some deterioration in SDF peak quality is inevitable due to the increased correlation of the filter views that cannot be aligned on axis. Also the inter-class discrimination will be adversely affected and non-target objects will probably need to be included in the filter training set.

For such a filter, finer control over the modulation function recorded in the photorefractive hologram may be required. This could be achieved by adaptively amplitude modulating the reference beam U_3 with an SLM. This would permit closer control of the bandpass nature of the recorded hologram. A function particularly suitable for implementing this is the difference of Gaussian (DOG) filter [21] which has been used previously for the recording of MSFs [22]. Thus the SDF can be tuned to a band of spatial frequencies; this should be chosen such that the high frequencies, which make the filter over sensitive to intra-class variations, are attenuated - as are the low frequencies which lead to high false alarm cross-correlations with inter-class objects. Such an SDF filter will inevitably generate a correlation signal of larger width than the high-pass varieties but this is a necessary trade-off.

Although designed with a photorefractive material in mind the filter synthesis procedure described in this Chapter can be applied, with minor modifications, to other up-dateable correlator realisations. For instance, the phase portion of the SDF design could be written to a phase modulating SLM and this addressed by an appropriate bandpass spectrum generated by DOG filtering the input scene.

References to Chapter 2

- [1] A. VanderLugt, "Signal detection by complex spatial filtering". *IEEE Trans.*, IT-10, p. 139, (1964).
- [2] D. Casasent, "Hybrid processors", Chapter 5 in: *Optical information processing*, Springer-Verlag, Topics in Applied Physics, Vol. 48, Lee, S., (Ed), New York, (1981).
- [3] F. Mok, M. Tackitt, H. Stoll, "Storage of 500 high resolution holograms in a LiNbO_3 crystal". *Opt. Lett.*, Vol. 16, No. 8, p. 605, (1991).
- [4] F. Yu, S. Wu, A. Mayers, S. Rajan, "Wavelength multiplexed reflection matched spatial filters using LiNbO_3 ". *Opt. Comm.*, Vol. 81, No.6, p. 343, (1991).
- [5] Project funded under the CEC BRITE/EURAM initiative (Reference RI1B/0294/C), "Hybrid optical/electronic industrial inspection system".
- [6] D. Psaltis, E. Paek, S. Venkatesh, "Optical image correlation with a binary spatial light modulator". *Opt. Eng.*, Vol. 23, p. 698, (1984).
- [7] T. Iwaki, Y. Mitsuoka, "Optical pattern recognition of letters by a joint transform correlator using a ferroelectric liquid-crystal spatial light modulator". *Opt. Lett.*, Vol. 15, No. 21, p. 1218, (1990).
- [8] D. Pepper, J. AuYeung, D. Fekete, A. Yariv, "Spatial convolution of optical fields via degenerate four wave mixing". *Opt. Lett.*, Vol. 3, No. 1, p. 7, (1978).
- [9] J-P. Huignard, H. Rajbenbach, Ph. Réfrégier, L. Solymar, "Wave mixing in photorefractive Bismuth Silicon Oxide crystals and its applications". *Opt. Eng.*, Vol. 24, No. 4, p. 586, (1985).
- [10] I. Cooper, M. Nicholson, C. Petts, "Dynamic frequency plane correlator". *Proc. IEE*, Vol. 133, Pt. J, No. 1, p. 70, (1986).
- [11] M. Nicholson, I. Cooper, M. McCall, C. Petts, "Simple computational model of image correlation by four wave mixing in photorefractive media". *Appl. Opt.*, Vol. 26, No. 2, p. 278, (1987).

- [12] P. Foote, T. Hall, N. Aldridge, A. Levenston, "Photorefractive materials and their applications to optical processing". *IEE Proc.*, Vol. 133, Pt. J., No. 1, p. 83, (1986).
- [13] S. Mallick, D. Roude, A. Apostolidis, "Efficiency and polarisation characteristics of photorefractive diffraction in a $\text{Bi}_{12}\text{SiO}_{20}$ crystal". *JOSA*, Vol. 4, No. 8, p. 1247, (1987).
- [14] D. Casasent, "Unified synthetic discriminant function computational formulation". *Appl. Opt.*, Vol. 23, No. 10, p. 1620, (1984).
- [15] H. Caulfield, W. Maloney, "Improved discrimination in optical character recognition", *Appl. Opt.* Vol. 8, p. 2354, (1968).
- [16] B. Vijaya Kumar, "Minimum-variance synthetic discriminant functions". *JOSA*, Vol. 3, No. 10, p. 1579, (1986).
- [17] A. Mahalanobis, B. Vijaya Kumar, D. Casasent, "Minimum average correlation energy filters". *Appl. Opt.*, Vol. 26, No. 17, p. 3633, (1987).
- [18] Ph. Réfrégier, "Filter design for optical pattern recognition: multicriteria optimisation approach". *Opt. Lett.*, Vol. 15, No. 15, p. 854, (1990).
- [19] Ph. Réfrégier, "Optimal trade-off filters for noise robustness, sharpness of the peak and Horner efficiency". *Opt. Lett.*, Vol. 16, No. 11, p. 829, (1991).
- [20] J. Horner, P. Gianino, "Phase-only matched filtering". *Appl. Opt.*, Vol. 23, No. 6, p. 812, (1984).
- [21] D. Marr, E. Hildreth, "Theory of edge detection". *Proc. R. Soc. Lond. B*, Vol. 207, p. 187, (1980).
- [22] D. Gibson, N. Aldridge, M. Brown, A. Levenstone, H. White, "Automatic recognition and tracking of targets from visible or thermal imagery using optical processing". *SPIE*, Vol. 492, p.165, (1984).

Chapter 3

Discrete frequency removal technique and its application to filter performance enhancement

3.1 Introduction

Previous sections have emphasised the importance of correlation peak localisation in producing a well defined correlation response with minimal sidelobe disruption, particularly when a heavily multiplexed filter is employed. To increase intra-class distortion tolerance a reduced bandwidth must be employed with a consequent increase in correlation width. However, since the width of the correlation peak is directly related to the frequency bandpass of the filter, a degradation has to be made from the ideal of an impulse-like correlation peak generated by the inverse filter (or MACE filter in the multiplexed case). In addition, the cross-correlation terms will increase in strength; since all these cannot be aligned on-axis they will cause an increase in the sidelobe structure of the correlation response of a multiplexed filter.

However, the observation has been made that a very marked increase in correlation peak localisation occurs with a moderate increase in power in the high frequency terms of the spectrum. Thus a filter recorded in BSO with a beam ratio of 1.0, which closely approximates a classical matched filter frequency response, produces the broad correlation signal shown in Figure 3.1(a). When BR is increased to 32.0, a much more localised correlation signal of high intensity is produced as shown in Figure 3.1(b). Considering the field directly behind the filter, that from the MSF falls with frequency as $|G(\mathbf{k})|^2$, whereas that behind the POF falls only as $|G(\mathbf{k})|$. Nevertheless, this variation makes a very considerable difference in the width of the correlation signal. To go from the POF correlation response to an impulse, however, requires the field values at high frequencies to be boosted to a very much greater degree, the field emerging from the filter being nominally uniform to achieve this. Thus there is a diminishing return, in terms of peak localisation, with too great an increase in amplitude of the high frequency terms.

The purpose of the next section is to examine more closely the relation between the amplitude structure of the field emerging directly behind the filter plane and the localisation and peak-to-sidelobe ratio of the corresponding +correlation (ie. the Fourier transform of this distribution). This was done by computer simulation using the fast

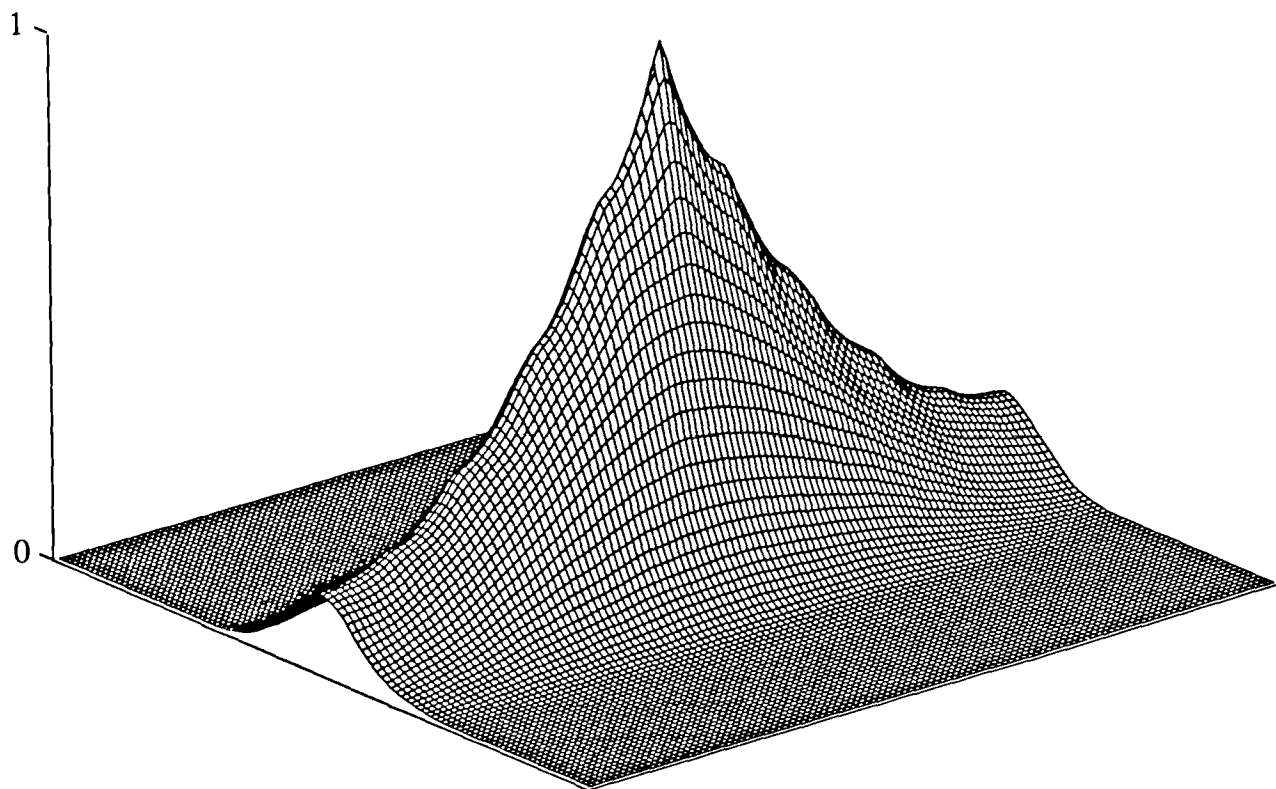


Figure 3.1(a) Matched spatial filter auto-correlation response

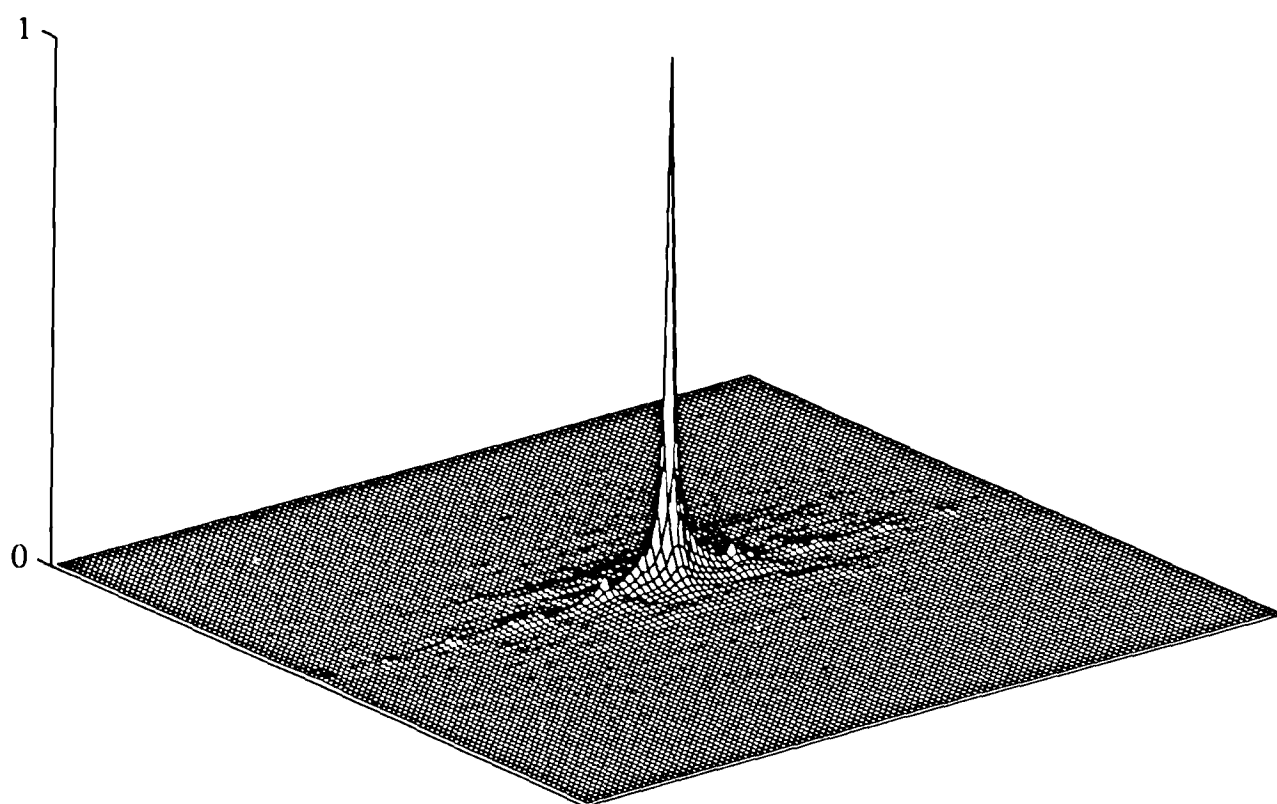


Figure 3.1(b) Phase only filter auto-correlation response

Fourier transform algorithm to compute the 2-D discrete Fourier transform of the field distributions. These simulations demonstrated that the correlation peak structure is remarkably robust against perturbation of the amplitude structure of the frequency plane field distribution. Two separate derivations are employed to suggest a reason for this observation. This property is used to generate a maximum discrimination filter by discrete frequency removal. The noise robustness of the frequency plane signal also suggests a possible alternative multiplexing method to that of the SDF technique which is briefly described, although not simulated, in the final sub-section.

3.2 Robustness of correlation peak localisation to frequency plane amplitude discontinuities

3.2.1 Demonstration by simulation

Consider a frequency plane filter in a 4-f optical correlator, matched to a particular signal. When this signal alone is present at the input to the system the filter has the effect of rotating the phasors, representing the complex field distribution of the Fourier transform of the signal, completely onto the real axis. Thus the field directly behind the filter plane is entirely real. In addition to altering the phase of the Fourier transform of the input, the modulus may also be attenuated in various ways depending on the exact design of the filter as has been thoroughly discussed in previous sections. In the case of a matched spatial filter the signal, $G(\mathbf{k})$, will be attenuated with frequency by $|G(\mathbf{k})|$, leading to the field emerging from the filter falling as $|G(\mathbf{k})|^2$ with frequency which is proportional to the power spectrum of the input. Thus its 2-D Fourier transform, effected by the second lens of the correlator will, according to the Wiener-Khinchine theorem, be the auto-correlation function of the signal. This is necessarily measured as the modulus squared of the output plane field distribution by an intensity sensitive device such as a CCD camera.

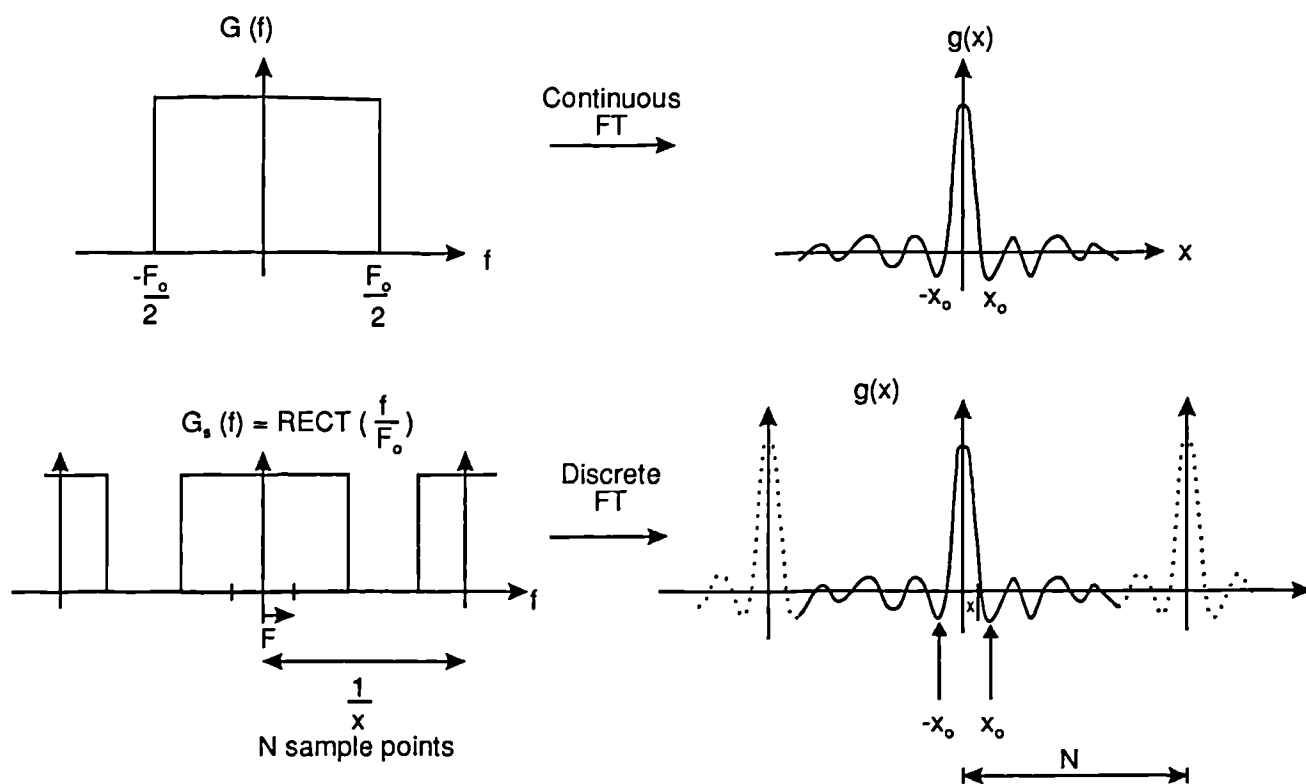
This section will concentrate on the situation in which there is no systematic attenuation of the frequency plane field with frequency. For a passive filter this can only be approximated since the input object spectrum cannot be boosted in those regions near its zero crossings. However, photographically implemented inverse filters, such as the MACE multiplexed filter, in practice yield very localised correlation responses. This is expected to closely resemble the impulse response of the optical system without the filter in place, since the field emerging from the filter is of nominally uniform modulus but apertured by the finite extent of the Fourier transform lenses. For the purposes of the

simulations which follow this aperturing effect is ignored. Thus a uniform frequency plane field distribution will Fourier transform to an on-axis impulse function. The simulations are performed on discrete data samples over a finite window area. However, if the window extends over the whole width of the uniform frequency plane data array the impulse response in reciprocal space will be a single data point in the centre of the array, all other points in the array being zero to within the precision of the computation; this is the discrete approximation to an impulse function. The sampling effects of the discrete Fourier transform are summarised diagrammatically in Figure 3.2 but are more thoroughly discussed by Brigham [1].

The field emerging from behind an ideal inverse filter matched exactly to the input signal will be uniform in modulus and of zero phase. Thus it is identical to the spectrum of a white noise signal since this is defined as having a uniform power spectrum. Equivalently white noise is defined as having an auto-correlation function which is an impulse function (any physical realisation will be bandlimited and so give rise to a sinc function approximation to the delta function). This equivalence is made readily apparent by invoking the Wiener-Khinchine theorem ie. a Fourier transform relationship exists between a signal's power spectrum and its auto-correlation function. Thus the ideal inverse filter will give an impulse-like auto-correlation function.

However, the simulation results which follow reveal that the frequency plane definition for a white noise signal in two (or higher) dimensions is a much stronger condition than demanding that its auto-correlation function should be impulse-like. The 2-D signal may have a power spectrum that is produced by a random noise generator, further thinned by thresholding, and still Fourier transform to a delta-like function. It should be pointed out that this result is only of direct interest in the present context of correlator filter design because the field emerging directly behind such a filter is separated from the correlation plane by a Fourier transform. It may thus be directly compared to the spectral density function of a noise process and its corresponding correlation function.

To investigate the degree to which the amplitude of the frequency domain waveform can be disrupted from a uniform distribution while still maintaining an impulse-like Fourier transform, the pseudo-random number generator G05 DDF from the NAG sub-routine library was employed to generate numbers with properties very close to those of a time random process. The numbers generated by the DDF sub-routine have a normal (Gaussian) probability density function. In addition, they are stationary and have a white noise spectral characteristic. The standard deviation of the distribution was set to various values which will be indicated for each result; the mean was set to zero in all



As $F_0 \rightarrow \frac{1}{X}$ the discrete SINC function tends to a δ functional approximation

(X, F) sampling interval

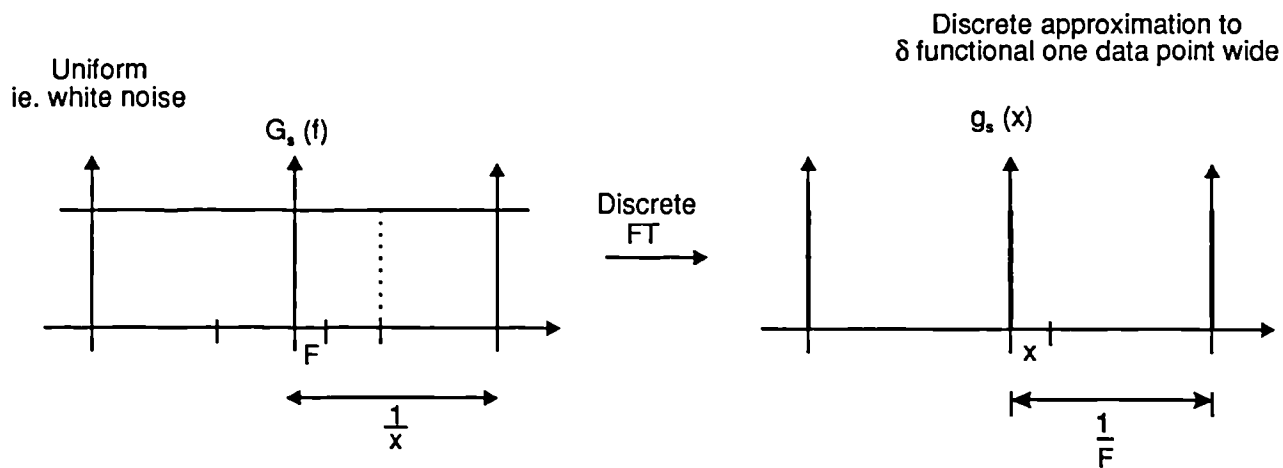


Figure 3.2 The discrete approximation to a δ -functional and its discrete Fourier transform

cases. The routine was used to fill half a 2-D data array with random numbers, the remaining half being filled with these numbers transposed so as to give a diametrically symmetric data array. This symmetry results from the filtering of a real input function and also leads to a totally real correlation function. The regions of the area that were below a certain number of standard deviations from the zero mean of the process were set to zero. The remaining array locations were then replaced by data generated from a further call to G05 DDF over half the array and the remaining half filled from this to produce a diametric symmetry. The absolute value of the zero numbers obtained was initially taken to generate the positive real field components as required. In this way random zero locations are generated but the data remaining have a full statistical spread. (If the original data above the zeroed cut-off were left in place their statistic would be modified due to the higher standard deviation contribution having been removed.) Figures 3.3(a) and (b) show x and y cross-sections through the centre of the array when the signal is generated in the manner first described with the cut-off threshold being set at one standard deviation. Thus just over 84% of the array elements are zeroed. The remaining points have a normal distribution (with a standard deviation of 10.0 in the units of the ordinate of the graph).

Figures 3.4(a) and (b) show radial cross-sections in the x and y direction of the real component of the field values of the 2-D Fourier transform of the field distribution shown in Figure 3.3. It can be clearly seen from these cross-sections that an impulse-like Fourier transform is obtained with the rest of the simulated correlation plane containing an approximately zero mean random noise signal of very low magnitude. It should be noted that since the Fourier plane field values have a null on-axis zero spatial frequency component for this particular noise realisation, the total integral of the distribution shown in Figure 3.4 will be zero. This is why the noise signal surrounding the impulse is not of exactly zero mean - it will have a small negative bias to balance the positive component making up the impulse spike. Notice also that the signal has radial symmetry due to the original frequency plane signal being totally real. Since it was also symmetric, the simulated correlation plane signal has zero imaginary field component. Figure 3.5(a) and (b) show radial x and y cross-sections through the modulus squared of the field values ie. the signal that would be measured as an intensity at the output of an optical correlator. To the limits of the accuracy shown by the plot this signal is a zero array with a single on-axis impulse function. In fact the highest noise spike away from the centre of the array was 250 times smaller than the impulse height. This ratio was found to be roughly constant between runs with different realisations of the noise process used to simulate the frequency domain signal, as might be expected. Figure 3.6 is a cross-section through the

correlation plane data array to show this noise spike. It can be seen that the noise in general is of lower magnitude due to its normal distribution.

The simulation was repeated with the frequency domain cut-off increased to 2σ ie. 97.7% of the data array being set to zero. The data array is shown in isometric form in Figure 3.7. Figure 3.8(a) and (b) show the corresponding real field values of the Fourier transform of this data. It can be seen that the signal to noise ratio has decreased from that of the previous example but it is important to notice in what way. The impulse in Figure 3.8 has a magnitude of 47 as compared to 325 for that in Figure 3.4. The excursion of the noise, however, rather than increasing, has fallen to about half of its previous value. These values are due to the much fewer number of non-zero frequency domain components of the input data array; the impulse height is equal to the integral of their values. Thus the reduced signal-to-noise ratio is brought about by the signal sinking towards the noise floor, which itself has also decreased, although by a smaller margin. The modulus squared of the simulated correlation plane, however, still maintains its near perfect impulse like appearance as can be seen from the cross-sections shown in Figure 3.8(c) and (d). In the next section it will be shown that the peak intensity is N times that of the background (where N is the number of frequency domain impulse pairs).

The remarkable degree to which the simulated correlation peak remains impulse-like, despite very large deviations of the frequency plane signal from the uniform distribution associated with a white noise spectral characteristic, may be analysed in two distinct ways. Each method gives a different insight into the property and so both will be described in the following two sub-sections. The first method is derived from the theory developed to analyse diffraction patterns generated by x-ray crystallography. The other method approaches the problem in a different way by using the central-slice theorem to reduce the 2-D Fourier transform relationship between the two signals to a projection followed by 1-D Fourier transformation.

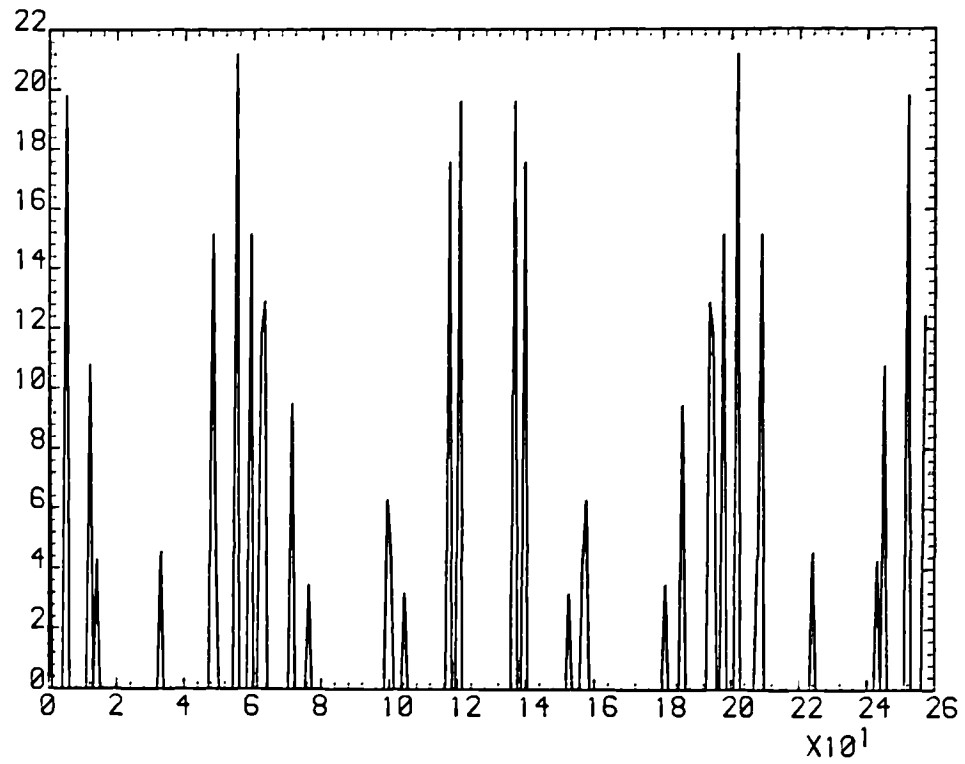


Figure 3.3(a) Cross-section at $y = 0$ through random data array with threshold set at one standard deviation

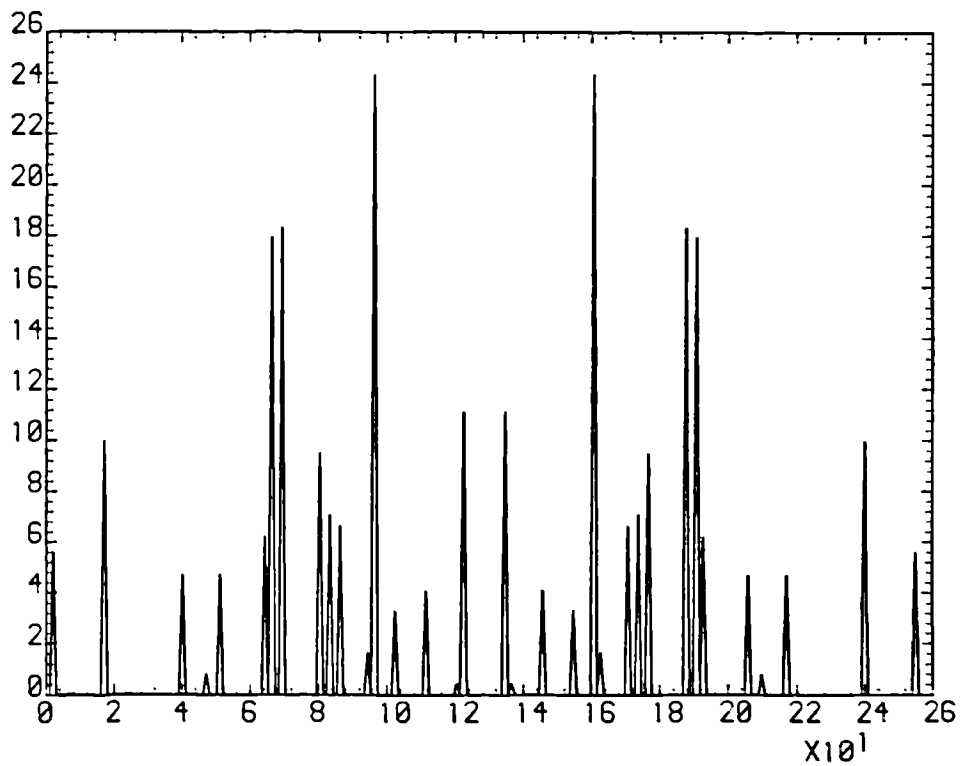


Figure 3.3(b) Cross-section at $x = 0$ through random data array with threshold set at one standard deviation

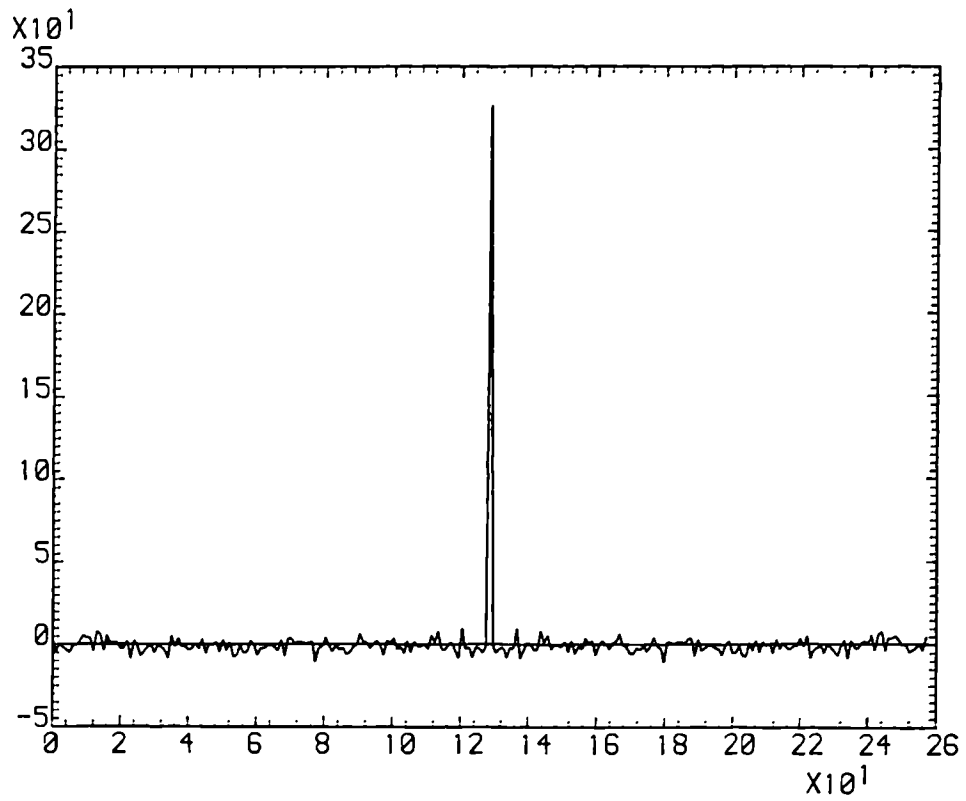


Figure 3.4(a) Cross-section at $y = 0$ through Fourier transform of random data array shown in Figure 3.3

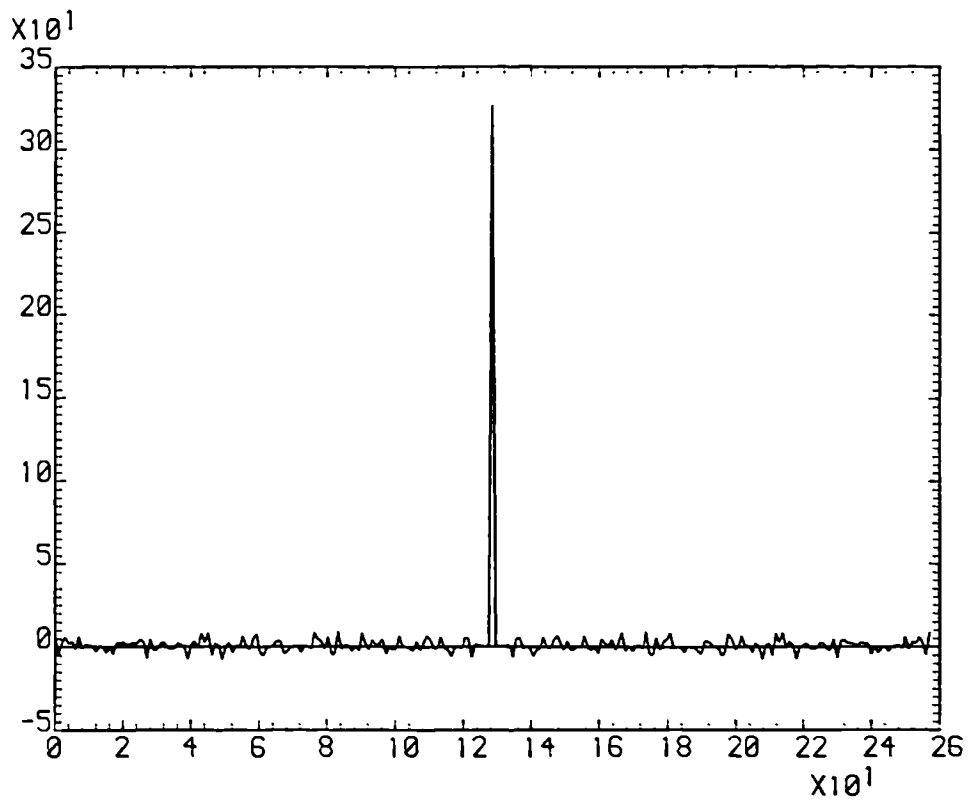


Figure 3.4(b) Cross-section at $x = 0$ through Fourier transform of random data array shown in Figure 3.3

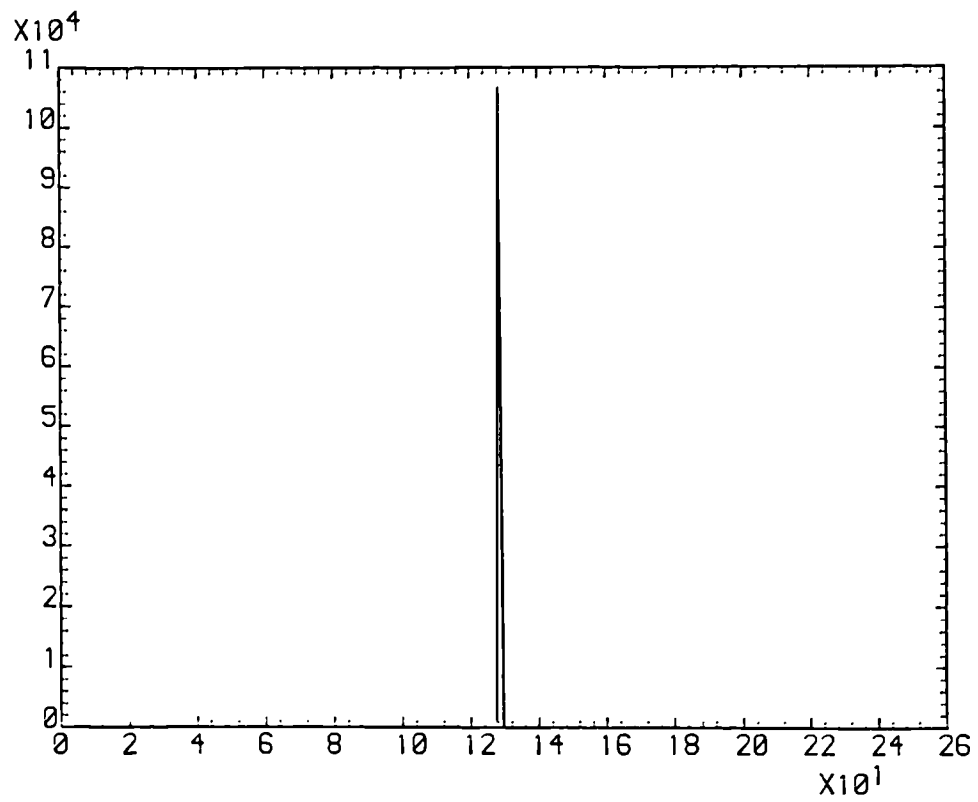


Figure 3.5(a) Cross-section at $y = 0$ through intensity values of the simulated correlation plane

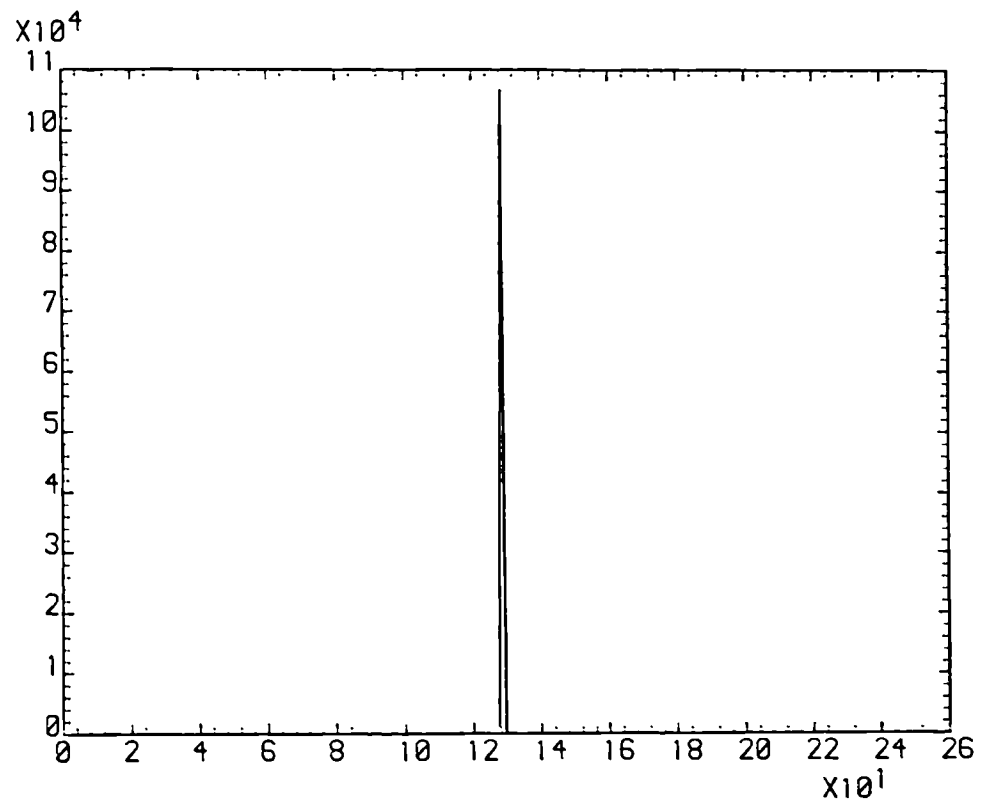


Figure 3.5(b) Cross-section at $x = 0$ through intensity values of the simulated correlation plane

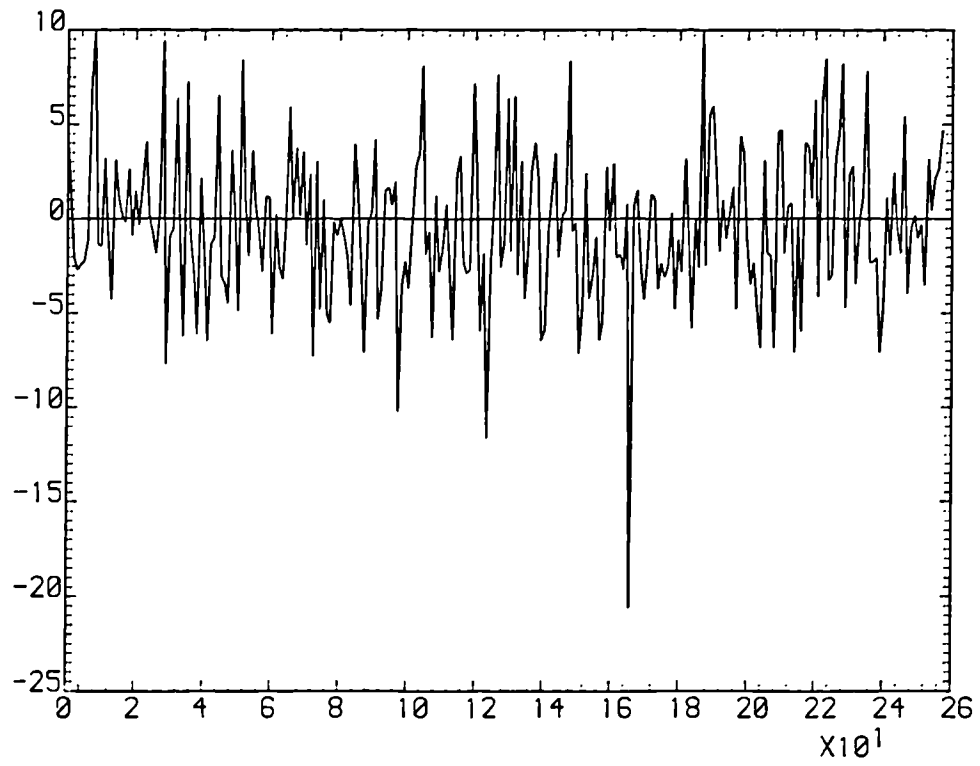


Figure 3.6 Cross-sections through maximum noise spike of the simulated correlation plane real field values

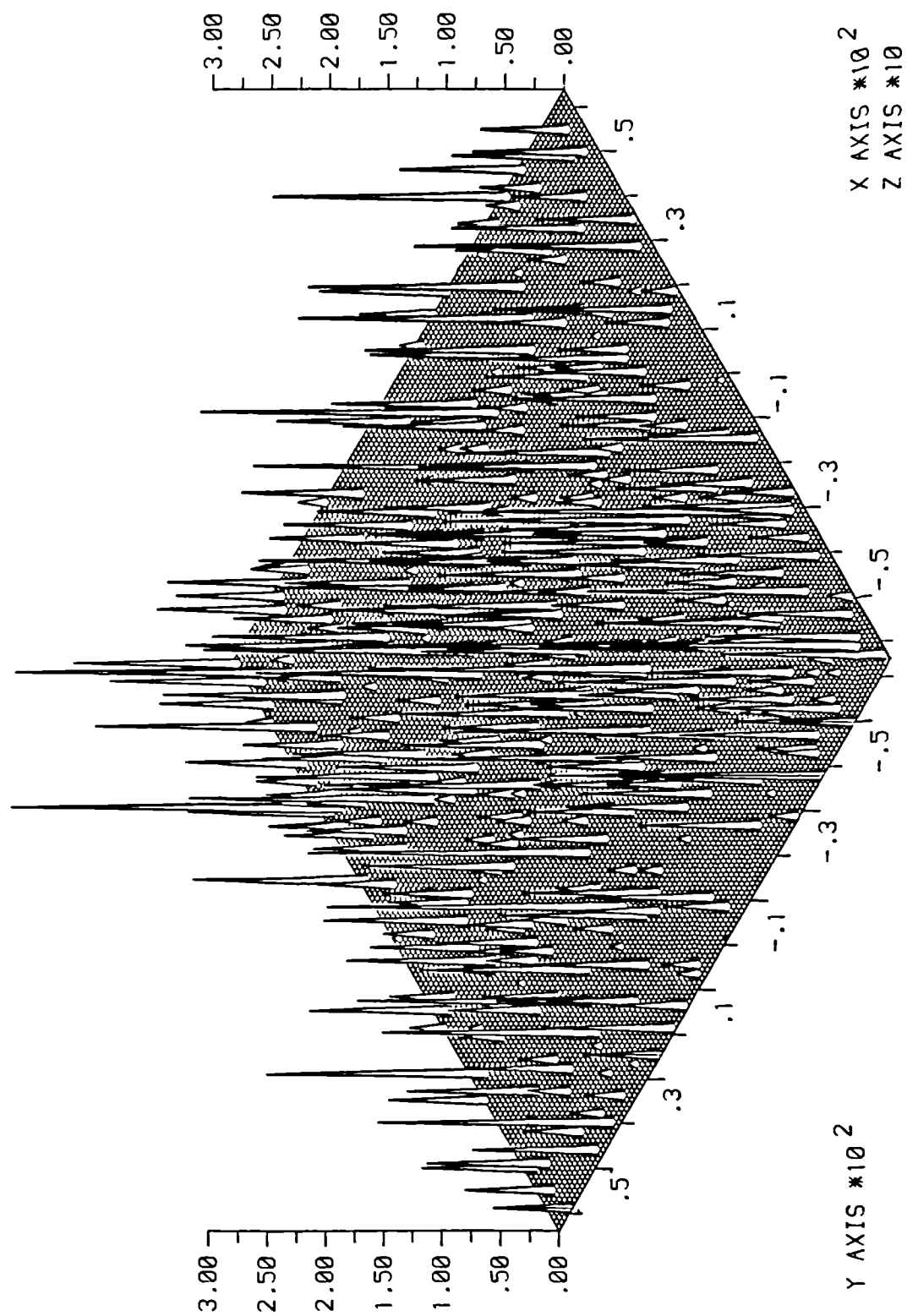


Figure 3.7 Isometric of random data array with threshold set at two standard deviations

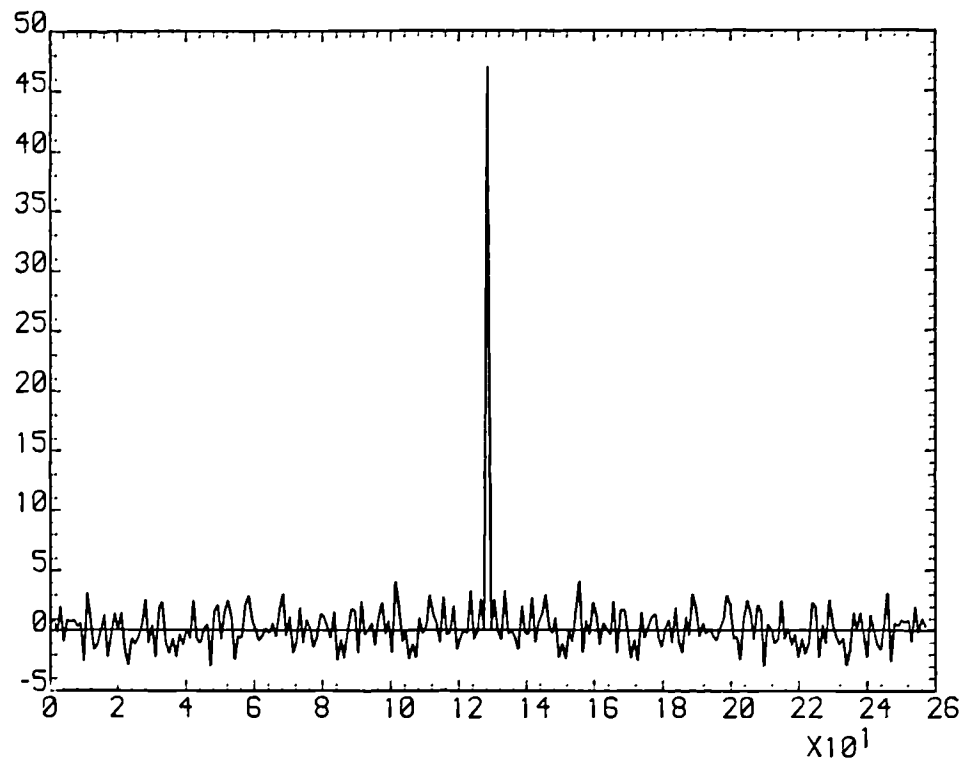


Figure 3.8(a) Cross-section at $y = 0$ through Fourier transform of random data array shown in Figure 3.7

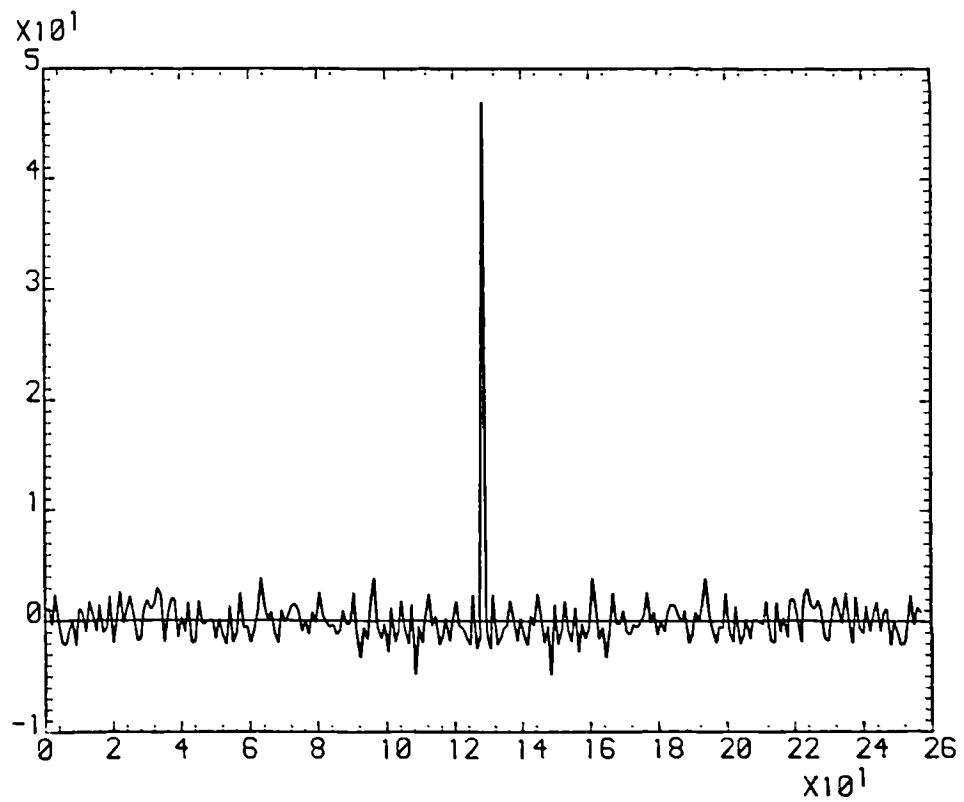


Figure 3.8(b) Cross-section at $x = 0$ through Fourier transform of random data array shown in Figure 3.7

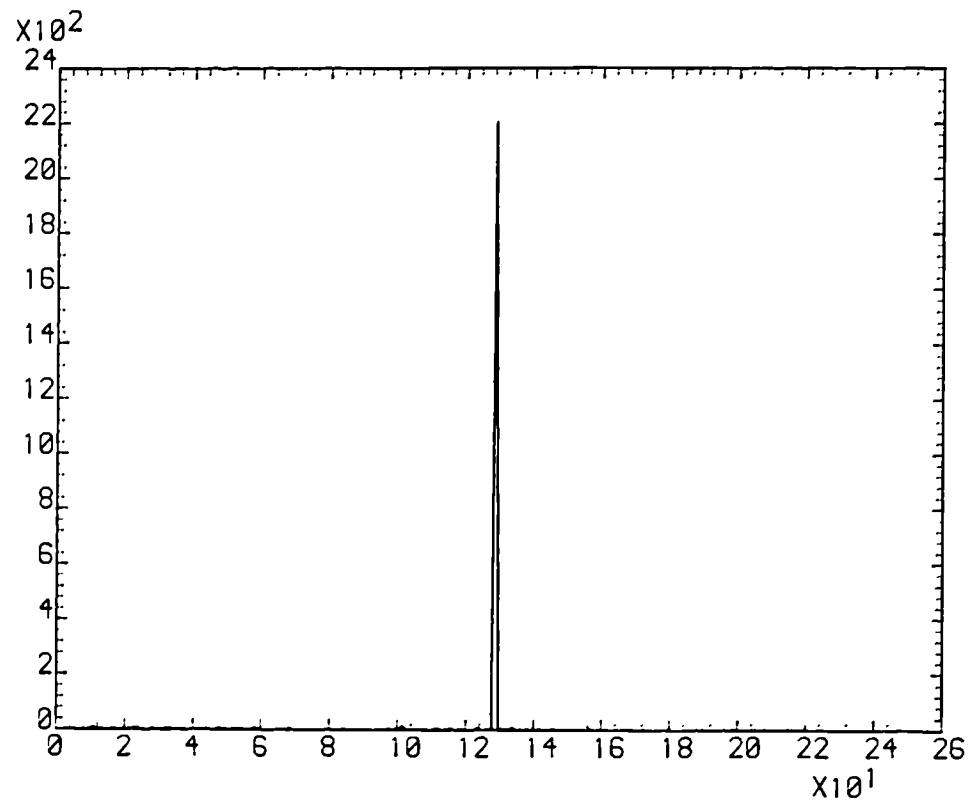


Figure 3.8(c) Cross section at $y = 0$ through intensity values of the simulated correlation plane

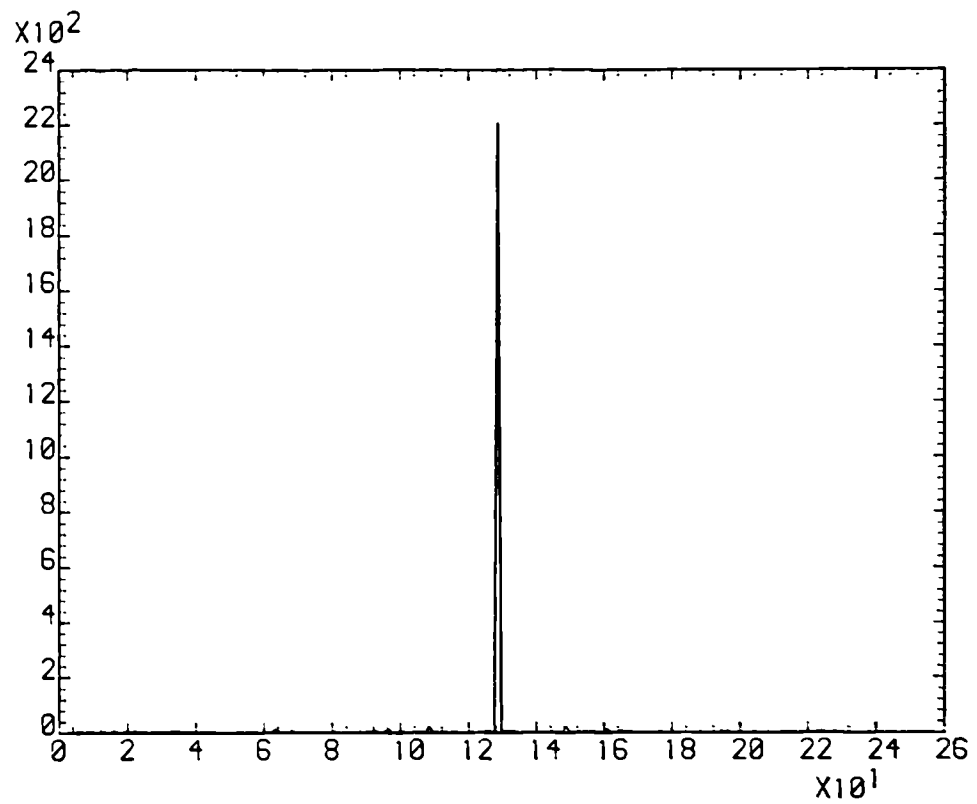


Figure 3.8(d) Cross section at $x = 0$ through intensity values of the simulated correlation plane

3.2.2 Analysis of correlation peak localisation by consideration of diffraction by a random set of delta functionals

A problem that arises in x-ray crystallography is the calculation of the diffraction pattern arising due to the diffraction of an x-ray beam from a large number of small absorbing molecules located at random in a plane [2]. By Babinet's principle this will yield a diffraction pattern identical to that generated from the complementary transmittance ie. a set of randomly located apertures in an opaque screen. The far field Fraunhofer diffraction pattern of the x-rays will be closely approximated by the 2-D Fourier transform of the aperture field distribution [3]. It will be detected as the power spectrum (modulus squared of the Fourier transform) by an intensity sensitive detector such as a photographic plate. The problem may be simplified by the use of the sifting property of the δ -functionals to appreciate that the aperture may be considered as the convolution of a single aperture with a series of δ -functionals at the locations of the apertures [3]. The problem is then almost identical to that of the simulated correlation response considered above, with the difference that the digital approximation to the δ -functionals have an imposed random amplitude. The 2-D Fourier transform of the set of randomly located δ -functionals may be written :

$$s(x, y) = \sum_{n=1}^N a_n \delta(k_x - k_{xn}, k_y - k_{yn}) \exp \left\{ -j2\pi (k_x x + k_y y) \right\} \quad \{3.1\}$$

where a_n is the random amplitude of a pair of diametrically opposed δ -functionals about the origin of the frequency plane. This sum cannot be evaluated analytically but some properties of $|s(x, y)|^2$ can be seen if it is written as a double sum:

$$\begin{aligned} |s(x, y)|^2 &= \left| \sum_{n=1}^N a_n \exp \left\{ -j2\pi (k_{xn} x + k_{yn} y) \right\} \right|^2 \\ &= \sum_{n=1}^N \sum_{m=1}^N a_n a_m \exp \left\{ -j2\pi [x(k_{xn} - k_{xm}) + y(k_{yn} - k_{ym})] \right\} \end{aligned} \quad \{3.2\}$$

Since k_{xn} , k_{xm} , k_{yn} and k_{ym} are random variables, so are $(k_{xn} - k_{xm})$ and $(k_{yn} - k_{ym})$. Thus the exponential terms generally make randomly positive and negative contributions to the overall sum. However, there are two exceptions. When $n = m$ the exponential is unity and the sum is equal to a_k^2 ($k=1,2,\dots,N$). This occurs N times which, since the a_k terms have a Gaussian probability density of standard deviation σ_a , results in a mean background intensity in the correlation plane equal to :

$$I_{\text{mean}} = \frac{1}{\sqrt{2\pi}\sigma_f} \sum_{k=1}^N a_k^2 \exp\left(-\frac{a_k^2}{2\sigma_f^2}\right) \quad \{3.3\}$$

In the limit as $N \rightarrow \infty$ this becomes :

$$I_{\text{mean}} = \frac{1}{\sqrt{2\pi}\sigma_f} \int_0^\infty a^2 \exp\left(-\frac{a^2}{2\sigma_f^2}\right) da = \frac{\sigma_f^2}{2} \quad \{3.4\}$$

The second contribution occurs when x and y are zero since then the exponential is unity and Equation {3.2} becomes :

$$|s(0,0)|^2 = \sum_{n=1}^N \sum_{m=1}^N a_n a_m \quad \{3.5\}$$

The statistics of the contributions to the on-axis term in Equation {3.3} is identical and so its value will, on average, be N times larger than the background intensity since N times as many a_n terms contribute to it. Thus the Fourier transform of the random array of δ -functionals has a large on-axis term surrounded by lower valued random terms, as observed in the simulations.

The statistics of the "noise" terms surrounding the central peak may be analysed further by exploiting the ergodic properties of the simulation. Thus rather than calculating the intensity for many points in the simulated correlation plane, the intensity fluctuations of a fixed point, P , are considered over many realisations of the frequency domain random function. For each trial a new field distribution at P results, which is nevertheless constrained to be along the real axis since the random frequency domain points have a diametric symmetry about the zero frequency point imposed upon them for the reason discussed previously. The probability that a particular realisation produces a modulus at P between m and $m + \delta m$ on the real axis is:

$$\text{Prob}(m) = \left(\frac{1}{\sqrt{2\pi}\sigma_f} \right) \exp\left(-\frac{m^2}{2\sigma_f^2}\right) \delta m \quad \{3.6\}$$

where σ_f is again the standard deviation of the frequency domain impulse pairs.

The intensity at the point P is the mean square value of m ie. $\langle m^2 \rangle$ which may be evaluated from :

$$\begin{aligned} I &= \frac{1}{\sqrt{2\pi}\sigma_f} \int_0^\infty m^2 \exp\left(-\frac{m^2}{2\sigma_f^2}\right) dm \\ &= \frac{\sigma_f^2}{2} \end{aligned} \quad \{3.7\}$$

which is identical to Equation {3.4} as expected by ergodicity. The standard deviation of the intensity fluctuations around the mean intensity at P over many realisations is :

$$\sigma_I = \sqrt{\langle m^2 - \langle m^2 \rangle \rangle^2}$$

and since

$$\langle 2m^2 \langle m^2 \rangle \rangle = 2(\langle m^2 \rangle)^2$$

this reduces to :

$$\sigma_I = \sqrt{\langle m^4 - \langle m^2 \rangle^2 \rangle} \quad \{3.8\}$$

Evaluating $\langle m^4 \rangle$ from :

$$\langle m^4 \rangle = \frac{1}{\sqrt{2\pi}\sigma_f} \int_0^\infty m^4 \exp\left(-\frac{m^2}{2\sigma_f^2}\right) dm \quad \{3.9\}$$

which after integration by parts and the use of the result $\frac{1}{\sqrt{2\pi}} \int_0^\infty \exp\left(-\frac{x^2}{2}\right) dx = \frac{1}{2}$

yields:

$$\langle m^4 \rangle = \frac{3\sigma_f^4}{2} \quad \{3.10\}$$

Substituting {3.7} and {3.10} into {3.8} gives :

$$\sigma_I^2 = \frac{\sqrt{5}}{2} \sigma_f^2 \quad \{3.11\}$$

which, again by ergodicity, also relates the intensity fluctuations observed over the simulated correlation plane to the standard deviation of the frequency plane random signal.

It should be mentioned that the mean and standard deviation of the background intensity is altered if the frequency plane does not have a conjugate symmetry since the phase at point P is then complex. This leads to an additional $2\pi m$ factor in Equations {3.7} and {3.9} and hence modified values for I and σ_I .

3.2.3 Analysis of correlation peak localisation using the central-slice theorem

The reason for the maintenance of an impulse-like correlation peak despite large deviations of the Fourier plane signal from a flat "white noise" spectrum in the two dimensional case may be further appreciated by applying the central-slice theorem to the problem as mentioned in the introductory remarks to this Chapter.

In general, the central-slice theorem relates the N-1 dimensional projections of an N dimensional signal to the N-1 dimensional cross-sections (or slices) through the N dimensional Fourier transform of the signal. In the 2-D case, the 1-D polar projections of the function form the 2-D Radon transform of that function. The 1-D Fourier transform of these projections are then cross-sections through the 2-D Fourier transform of the function at the corresponding projection angle. Figure 3.9 depicts this relationship and indicates the required notations. The rigorous mathematical derivation of the relationship has been presented by several authors but a concise and clear version is given by Easton and Barrett [4] which is summarised here. The Radon projection, $\lambda(p, \phi)$, is defined as a 2-D integral transform parametrised by projection angle ϕ :

$$\lambda(p, \phi) = \int \int_{-\infty}^{\infty} f(\mathbf{r}) \delta(p - \mathbf{r} \cdot \mathbf{n}) dx dy \quad \{3.12\}$$

where the δ -functional acts along the normal to the projection angle to reduce the area integral of $f(\mathbf{r})$ to a set of line integrals over the projection variable p by use of the sifting property of the δ -functional. 1-D Fourier transformation of $\lambda(p, \phi)$ then gives:

$$\Lambda(v, \phi) = \int_{-\infty}^{\infty} \left[\int \int_{-\infty}^{\infty} f(\mathbf{r}) \delta(p - \mathbf{r} \cdot \mathbf{n}) dx dy \right] \exp(-j2\pi p v) dp \quad \{3.13\}$$

which on changing the order of integration becomes:

$$\Lambda(v, \phi) = \int \int_{-\infty}^{\infty} f(\mathbf{r}) \left[\int_{-\infty}^{\infty} \delta(p - \mathbf{r} \cdot \mathbf{n}) dp \right] dx dy \quad \{3.14\}$$

and then reduces to:

$$\Lambda(v, \phi) = \int \int_{-\infty}^{\infty} f(\mathbf{r}) \exp(-j2\pi \mathbf{n} v \cdot \mathbf{r}) dx dy \quad \{3.15\}$$

which is equivalent to the 2-D Fourier transform of $f(\mathbf{r})$ evaluated at spatial frequencies $\mathbf{k} = \mathbf{n}v$ ie. along a direction corresponding to an angle ϕ from the k_x frequency domain axis and orthogonal to the direction of projection.

In the present context, the central-slice theorem may be employed to account for the localisation of the simulated correlation peak despite the non-uniformity of the frequency domain signal. Since these are related by a 2-D Fourier transform, radial sections through the correlation plane correspond to the 1-D Fourier transform of projections of the frequency domain random signal. Each point on a particular projection, eg. $\Lambda(v, \phi_0)$ in Figure 3.10, results from the integral of the signal lying on a line orthogonal to it, as described by :

$$\Lambda(v, \phi) = \int \int_{-\infty}^{\infty} F(\mathbf{k}) \delta(v - \mathbf{k} \cdot \mathbf{n}) dk_x dk_y \quad \{3.16\}$$

However, since this is a random signal, if this integral is performed over a large enough data sample the sum will be approximately the same for each data line on the projection, the value becoming more nearly constant the larger the data sample. This will also be true for the other projection directions. Thus it can be seen that the projections $\Lambda(v, \phi)$ will be approximately uniform in value due to the averaging effect of the line integral acting over the random data. This constant has a 1-D Fourier transform which approximates a δ -functional ie.

$$\frac{\sigma_f^2}{2} \int_{-\infty}^{\infty} \exp(-j2\pi p v) dv = \frac{\sigma_f^2}{2} \delta(0) \quad \{3.17\}$$

the approximations being that the line integral and projection are of a finite extent. The finite limit on the line integral results in fluctuations of the height of each

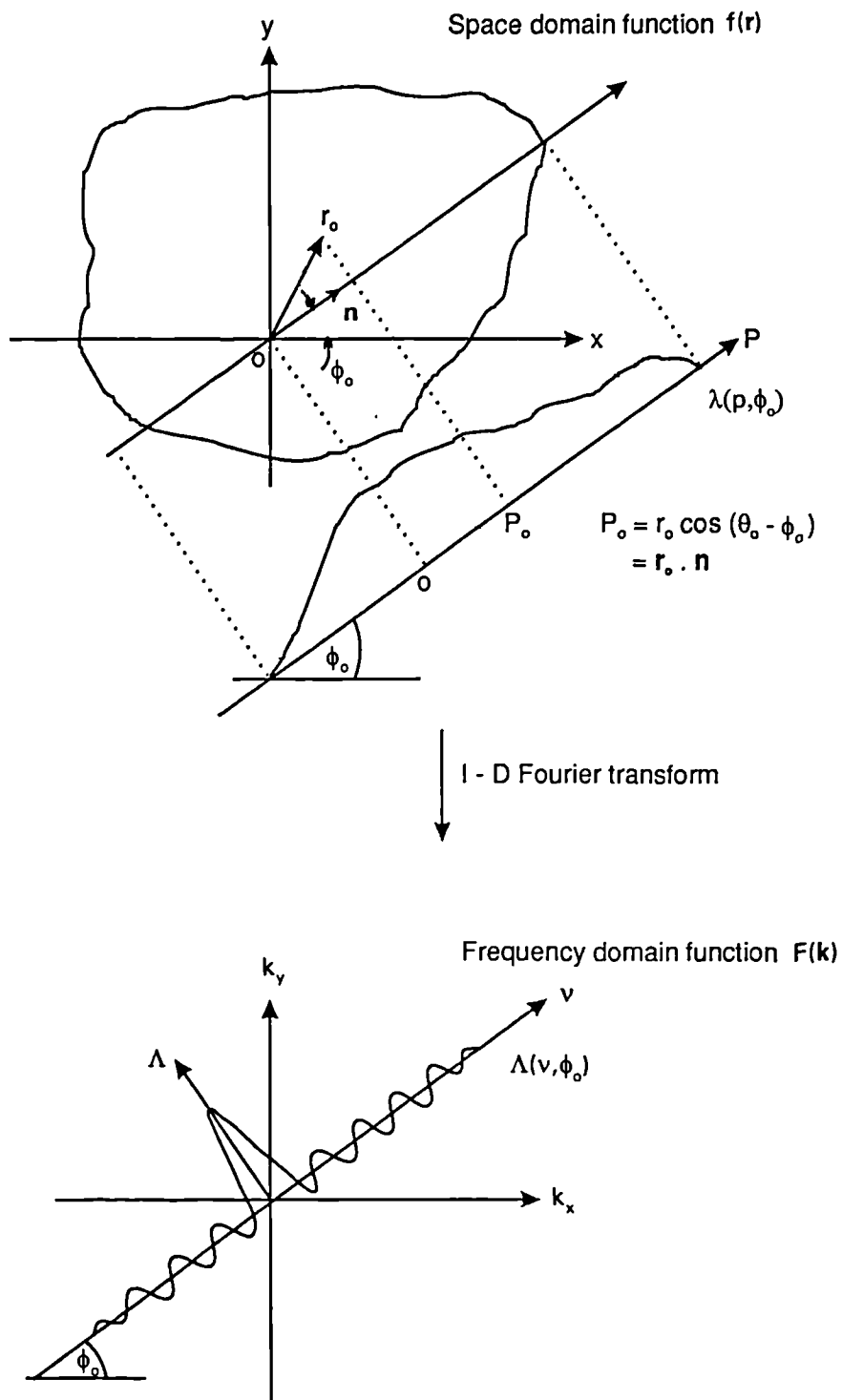


Figure 3.9 Depiction of the central-slice theorem with notation

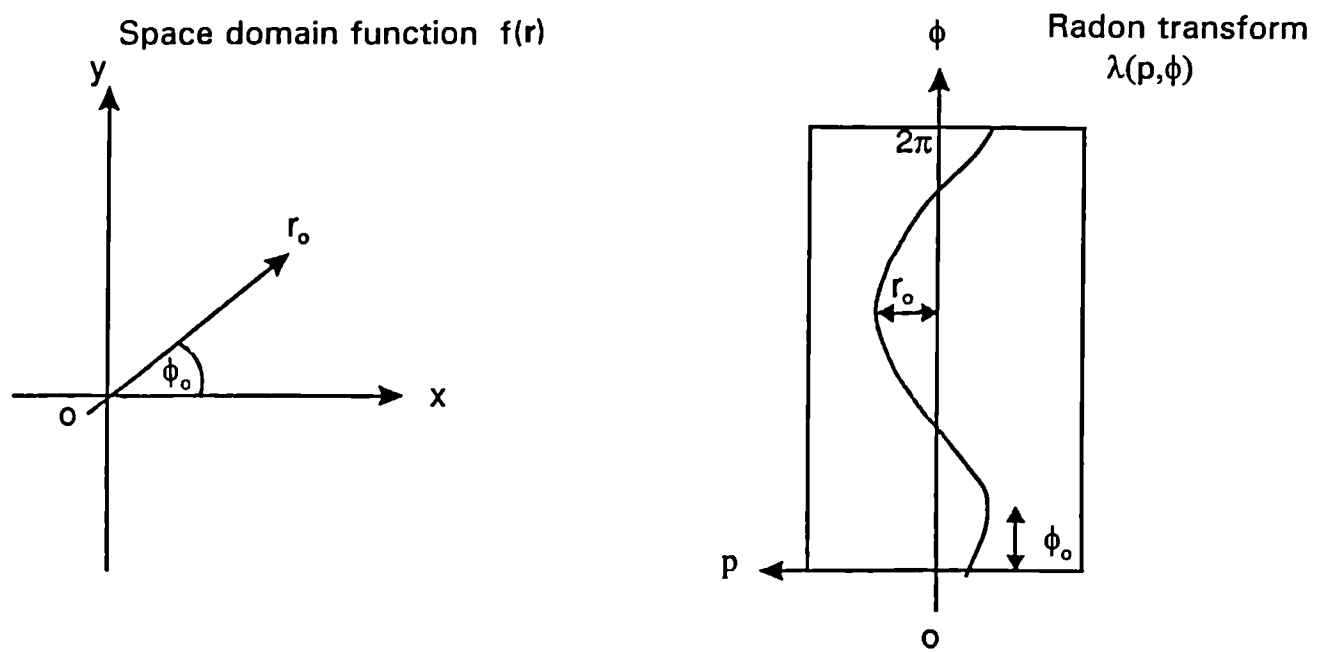


Figure 3.10 The Radon transform

projection from the expected value of $\frac{\sigma_f^2}{2}$ (see Equation {3.7}). The finite limit on the Fourier transform results in a sinc function approximation to the δ -functional. (In the case of the digital simulation the discrete Fourier transform gives rise to a single data point representing the δ -functional as described previously in Section 3.1. Thus the 2-D correlation peak is constrained in all radial directions to be impulse-like despite the large deviations from uniformity of the frequency domain signal occurring in the simulated frequency domain distributions shown in Figures 3.3 (a) and (b).

The Radon transform forms the basis of many applications which require the recovery of a function from its projections, the most widely used being the technique for reconstruction of cross-sections of the anatomy of a living subject from x-ray projections, known as computer assisted tomography (CAT). However, direct back projection is most commonly used for reconstruction in tomographic computations due to a number of practical difficulties in using the central-slice theorem. With the recent wide availability of specialised DSP chip sets capable of performing the required discrete Fourier transform operations very rapidly, the central-slice theorem method of reconstruction becomes more attractive. In addition, another technique may benefit from application of the theorem.

The Hough transform is regarded as a powerful pre-processing operation in machine vision applications [5]. In the most common form of the parametrization it is essentially equivalent to the Radon transform operating on an edge enhancement of the initial input image. The Radon transform has the property that a point in Cartesian space is mapped to a sinusoid in Radon space of phase and amplitude governed by the x,y location of the point, as depicted in Figure 3.11. Further, it can be shown that all points lying on a straight line in Cartesian space give rise to sinusoids in Radon space that cross at a point which corresponds to the parametrisation of the line [5]. Thus thresholding the Radon (Hough) output yields local maxima that indicate the position of line segments in the image. Since the maxima arise from an integration of all points contributing to the line they are only gradually degraded by breaks in the line, which permits the transform to be robust against the imperfect results usually generated when the various edge operators are applied to real world imagery.

The main difficulty with the Hough transform is its computational complexity, direct calculation of the sinogram requiring the evaluation of the modulus and arctan of each point in the image and a recording of the corresponding sinusoid. However, the inverse of the central-slice theorem may be employed to evaluate the Hough transform by the sequence: 2-D Fourier transformation of the image, a polar to Cartesian

mapping and subsequent inverse 1-D Fourier transformation (in addition, edge enhancement can be accomplished by multiplication in the frequency plane which is available after the 2-D Fourier transform). Using the computational savings of the FFT algorithm to effect the Fourier transforms results in a much faster implementation. For instance, on an IBM 3090 (with vector processor) the direct evaluation of the Hough transform of a 512x512 image required 234.4 seconds CPU. In contrast, the method using the central-slice theorem just outlined required only 14.8 seconds CPU, this code being in no way optimised to run particularly quickly. The recent availability of digital signal processing chip sets capable of rapidly performing the FFT (see Chapter 5 for further details) together with a look-up table for the polar to Cartesian mapping, make a video-rate Hough transform feasible. In addition, a non-scanning optical implementation is also possible employing a recently patented r - θ to x - y mapping mirror system together with 2-D and 1-D (cylindrical) Fourier transform lenses.

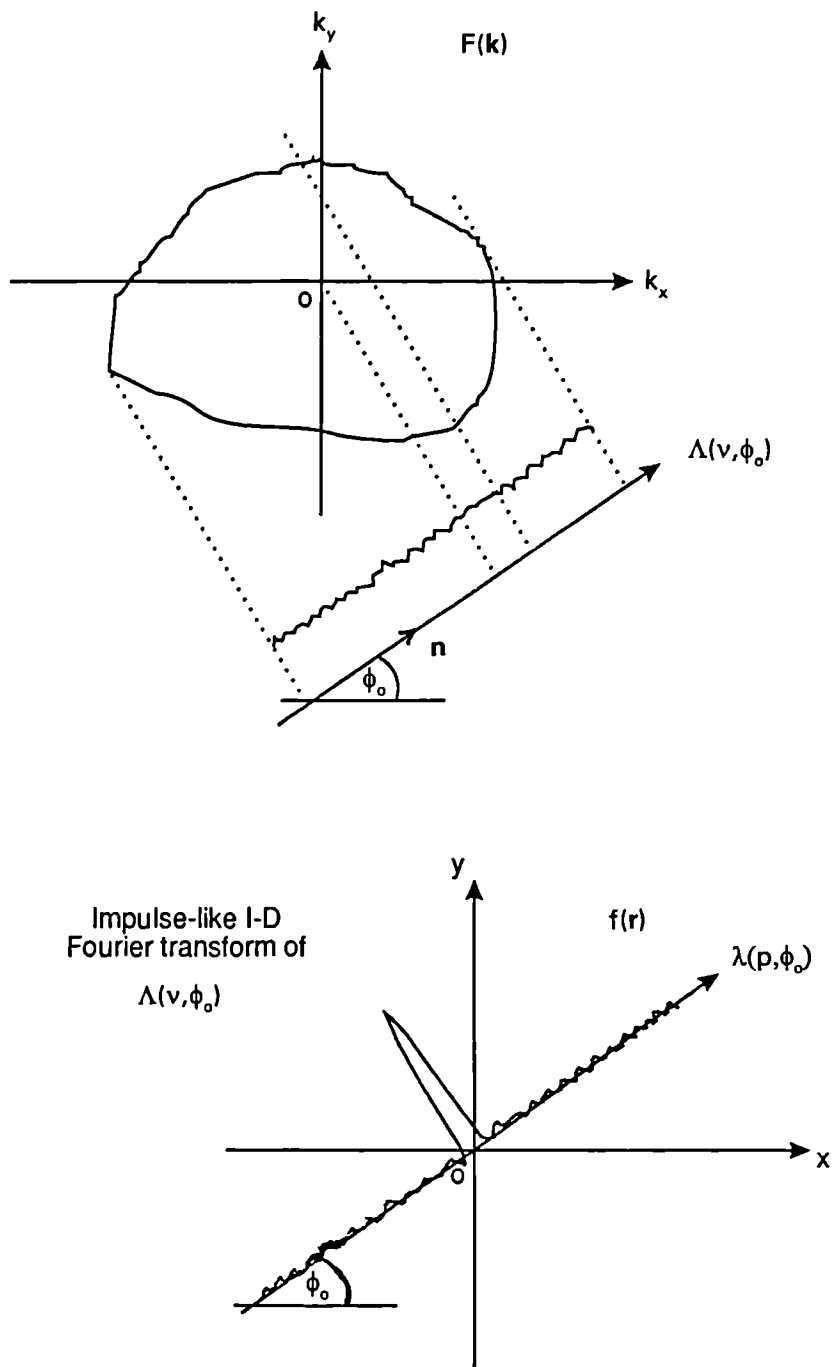


Figure 3.11 Application of the central-slice theorem to a random frequency domain function

3.3 Use of discrete frequency removal to enhance the correlation filter performance

3.3.1 Simulation of the joint transform correlator

The maintenance of the localisation of the correlation peak despite severe amplitude disruption to the frequency domain field distribution can be made use of in order to enhance the performance of a correlation filter; the discrimination ability of the filter can be improved while simultaneously reducing the peak fall-off with alterations of the target object from the view used to generate the filter. Simulations to demonstrate this application are given in this section.

The transfer function used in the simulations was that generated by a photorefractive based joint transform correlator configuration. The optical arrangement to implement this is shown in Figure 4.1 and was the configuration used in the experimental work described in Chapter 4. This arrangement has some advantages for the bandpass filtering work carried out which will be discussed in that Chapter.

Since the photorefractive correlator is operated non-degenerately, the interaction of the HeNe read-out beam may be ignored and the photorefractive interaction in the steady-state described by the simple transfer function:

$$m(\mathbf{k}) = \frac{2U_1^*(\mathbf{k})U_3(\mathbf{k})}{|U_1(\mathbf{k})|^2 + |U_3(\mathbf{k})|^2} \quad \{3.18\}$$

where the notation is identical to that used in Chapter 2. The difference between Equation {3.18} and Equations {2.1} and {2.2} is that the beam $U_3(\mathbf{k})$ is modulated with the reference image rather than being a plane wave as in the implementation more closely analogous to the Vander Lugt type correlator. Thus $U_3(\mathbf{k})$, the Fourier transform of the reference object, interacts with the Fourier transform, $U_1(\mathbf{k})$, of the input scene to generate the overall grating modulation described by Equation {3.18}. This has some important implications for the performance of the correlator, such as the limitation on maximum possible update speeds, which are not directly relevant to the present discussion.

Again, as in Chapter 2, a linear relation between $m(\mathbf{k})$ and the diffracted field is taken, the space charge electric field in the photorefractive crystal being assumed not to saturate. Thus if a plane wave illuminates the crystal, the field diffracted by the

hologram will have a complex distribution given by Equation {3.18}. Upon further Fourier transformation by L_2 , the cross-correlation between $U_1(\mathbf{k})$ and $U_3(\mathbf{k})$ will thus be generated at plane P_4 . In the particular case of $U_1(\mathbf{k})$ and $U_3(\mathbf{k})$ being identical, Equation {3.18} reduces to a constant. Thus the HeNe read-out beam remains plane in amplitude and unperturbed in phase after diffraction from the unmodulated grating formed in the hologram by the two identical inputs. The subsequent Fourier transform of this distribution yields an on-axis peak which is the impulse response of the clear aperture of the crystal (since this is generally much smaller than that of lens L_2) and corresponds to the auto-correlation of the input distributions. If one of the inputs is slightly translated (under large translation Bragg matching effects must be considered due to the volume nature of the hologram formed in the crystal) a linear phase factor is introduced into Equation {3.18}. The read-out wave is thus deflected by a corresponding angle and the resulting correlation peak shifted off-axis by an amount proportional to the movement of the input.

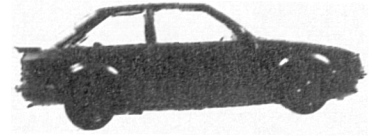
Figure 3.12 shows typical views of the input objects used in the simulations. The discrimination ability between the same aspect views of the Ford Escort and Mazda RX7 cars and tolerance to alterations of the aspect of the Escort were used in tests of the filter performance. The view aspects are referred to by the angular increment from a front view, for example E090 being a left hand side view of the Escort.

An important difference of the joint transform configuration from that of the Vander Lugt correlator is that the auto-correlation function, being simply the impulse response of the optical system, is independent of the functional form of the input object. Thus the auto-correlation function will always be the Fourier transform of the amplitude distribution of the HeNe beam used to read-out the unmodulated grating formed in the photorefractive volume hologram by an identical input and reference function.

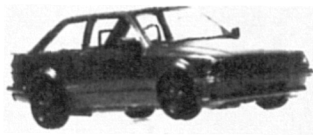
It is necessary to relate the various physical dimensions used in the experimental implementation of the correlator to the simulation in order to account for the finite size of the photorefractive crystal and match the modulation of the HeNe reconstruction beam to this. The models of the Ford Escort and Mazda were photographed to produce transparencies which were used as inputs to the experimental correlator. A high resolution emulsion was employed (Kodak orthochromatic copy film) so that a small enough image could be used to avoid, as far as possible, the effects of a limited field of view in the reference arm of the correlator caused by Bragg dephasing effects in the volume hologram. For example, the Ford Escort image measured 2.03 mm between wheel bases, a 46.8 times demagnification from



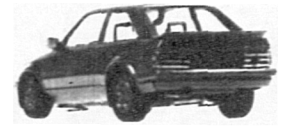
E000



E270

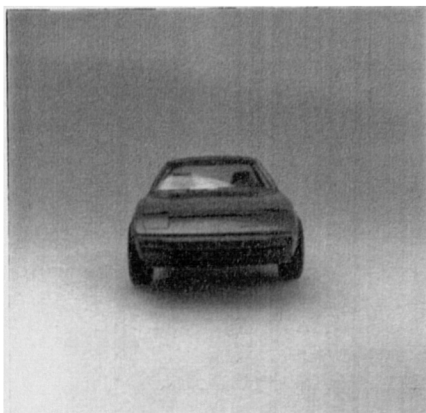


E315

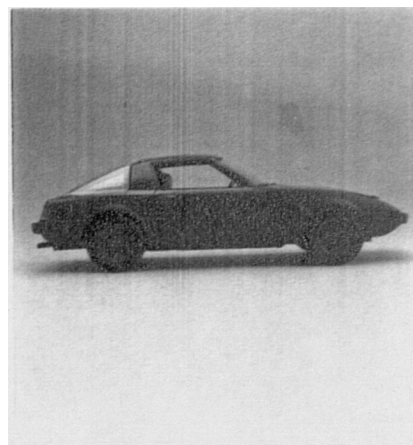


E135

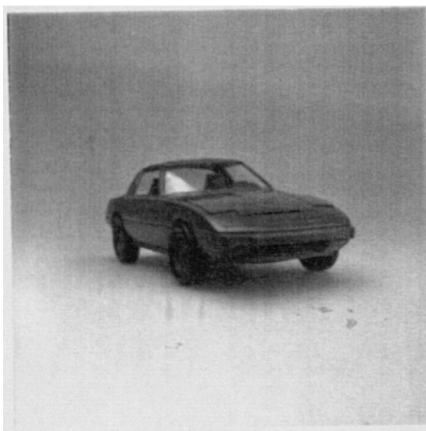
Figure 3.12(a) Example views of the Escort car test imagery



M000



M270



M338



M158

Figure 3.12(b) Example views of the Mazda car test imagery

the $1/25^{\text{th}}$ scale model. The model was digitised for input to the simulation in a 256×256 frame (stretched from 256×171 to 256×256 pixels to maintain a uniform x and y sampling rate in the framestore) such that there were 110 pixels between the wheel bases. The model was thus digitised to a corresponding maximum spatial frequency of 27.1 cycles/mm when related to the scale of the photographic transparencies used in the experimental work. Further physical dimensions of importance are the $9 \text{ mm} \times 9 \text{ mm} \times 1.4 \text{ mm}$ size of the Bismuth Silicon Oxide (BSO) crystal and the 600 mm focal length of the Fourier transform lenses used in the optical correlator. Thus the maximum spatial frequency of the correlator may be calculated from:

$$k_{x_{(\max)}} = \frac{x_{f_{(\max)}}}{\lambda f} \quad \{3.19\}$$

which, when using a Argon ion laser on the 514.5 nm line with $x_{f_{(\max)}} = 4.5 \text{ mm}$, gives $k_{x_{(\max)}} = 14.6 \text{ cycles/mm}$. The finite aperture of the BSO crystal will thus act as a low pass filter. In the simulation, the 256×256 array is zero supplemented to a 512×512 array before discrete transformation to avoid overlap of the correlation signal from neighbouring data cells, as is the standard procedure when computing discrete correlations or convolutions [1]. This has the effect of interpolating the frequency plane signal rather than increasing the maximum spatial frequency calculated (since the sampling rate of the input signal remains unaltered). Thus the maximum spatial frequency computed remains 27.1 cycles/mm. To simulate the experimental situation, spatial frequencies beyond 14.6 cycles/mm were, therefore, set to zero before the complex multiple with the reference signal and subsequent discrete Fourier transformation to the correlation plane. This resulted in the computed auto-correlation function spreading from the single on-axis non-zero data point (representing the δ -functional transform of the uniform frequency plane field distribution) to the sinc(x) function shown in Figure 3.13. As stated previously this auto-correlation shape is independent of the input function in the joint transform correlator, the spread of the sinc function being controlled only by the extent of the frequency plane aperture. Of course, experimentally, the intensity of the correlation plane is observed ie. the modulus squared of the sinc function, which is described analytically by :

$$I(x, y) = \left(\frac{l_x l_y}{\lambda f} \right)^2 \text{sinc}^2 \left(\frac{l_x x}{\lambda f} \right) \text{sinc}^2 \left(\frac{l_y y}{\lambda f} \right) \quad \{3.20\}$$

where l_x, l_y are the dimensions of the frequency plane aperture.

This is shown in Figure 3.13(b). It can be seen that the sidelobe structure is very small, the first two maxima of the $|\text{sinc}(x)|^2$ function being just visible. They have a value 0.0471 of the central peak and a separation given by $2.86\pi\lambda f/l_{x,y}$. As will be seen in the experimental results presented in Chapter 4, the auto-correlation function observed at the output plane of the optical correlator appears as a very localised point of light when the BSO is illuminated by a uniform HeNe beam, the sidelobe structure being too weak to be significant.

In practice a uniform reconstruction beam provides far too high a sensitivity to intra-class variation of the input object since it leads to a filtering effect equivalent to the inverse filter, uniform weighting being given to all spatial frequencies out to k_{max} in the frequency plane (see Chapters 1 and 2 for a detailed discussion of the inverse filter and its multiplexed equivalent, the MACE filter). Thus it is necessary to attenuate the high frequency regions of the spectrum during the correlation process. In the non-degenerate four wave mixing joint transform correlator this is readily accomplished by amplitude modulating the read-out HeNe beam before diffraction from the volume hologram or, more simply, by using its already Gaussian TEM_{00} profile and expanding this to the correct size determined to be suitable for the pattern recognition task. Further, the addition of a Gaussian attenuation at the centre of the crystal can be used to generate a bandpass filter. A detailed consideration of the filters that can be generated by these techniques will be given in Chapter 4. In the present context it is relevant in that a Gaussian profile is used to modulate the reconstruction beam in the simulation prior to the removal of individual frequency terms. This removes high frequency components that are known *a priori* to be of very little value since they will only be phase matched for inputs identical to the reference function. Figure 3.14(a) shows the frequency plane Gaussian modulation of the reconstruction beam. This is of the same size relative to the digital data array as that used experimentally; the distance to the $1/e$ point of the maximum amplitude of the Gaussian beam (measured as the $1/e^2$ intensity point) is equivalent to 0.95 mm which is represented by 29.074 data points in the array. (The vertical bands in Figure 3.14(a) indicate the 9 mm aperture of the BSO crystal, represented by 275 data points). Figure 3.14(b) shows the Fourier transform of this Gaussian which is itself Gaussian since:

$$F[\exp(-2\sigma^2\pi^2\mathbf{k}^2)] = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\mathbf{r}^2}{2\sigma^2}\right) \quad \{3.21\}$$

This then represents the ACF when the bandwidth is reduced by the frequency plane apodisation. The ACF is now broader but this is a required trade-off as discussed in Chapters 2 and 4. Since it is the intensity that is measured at the correlator output the peak will, in fact, appear narrower as shown by Figure 3.14(c) which is the modulus squared of the field values plotted in Figure 3.14(b).

The ACF peak intensity will, however, fall with reduction of bandwidth if the power of the read-out beam is kept constant since the correlation response becomes less localised and the available power is distributed over a larger area of the output plane. A similar effect on the correlation peak localisation due to alteration of the frequency plane weighting also occurs in the Vander Lugt photorefractive based correlator configuration as demonstrated by the simulation results given in Chapter 2. The system implications of this are briefly considered there.

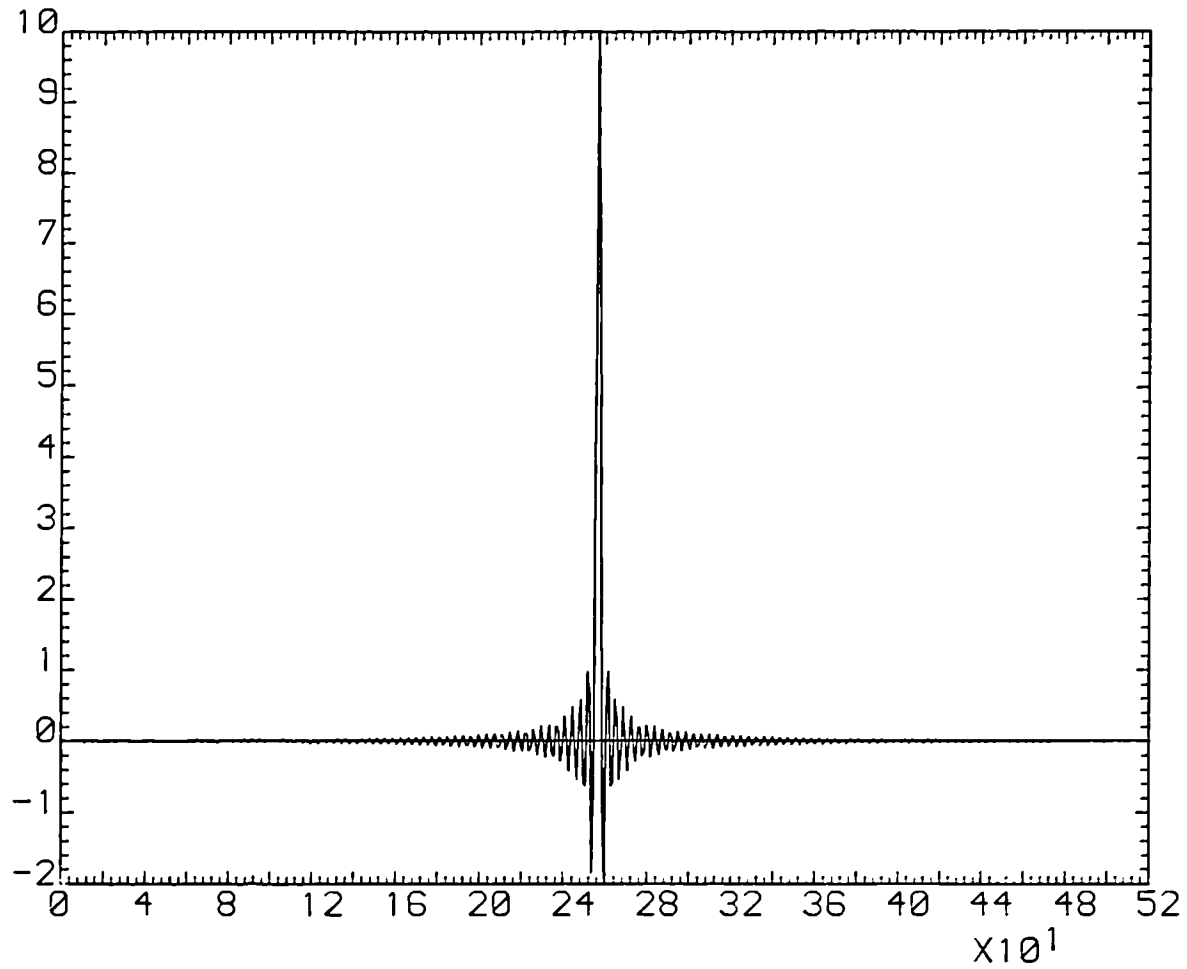


Figure 3.13(a) Auto-correlation function generated by the joint transform correlator

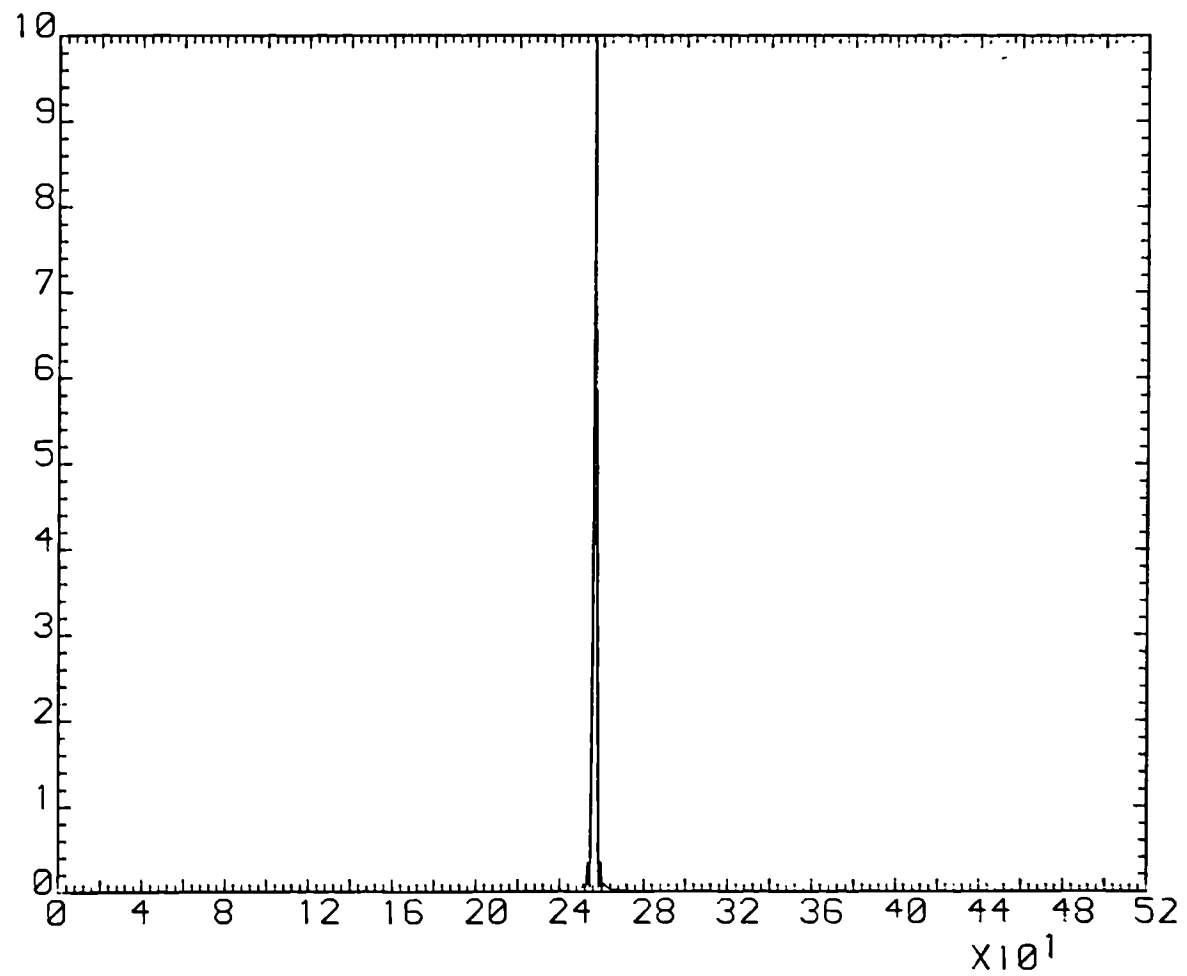


Figure 3.13(b) Intensity of auto-correlation function

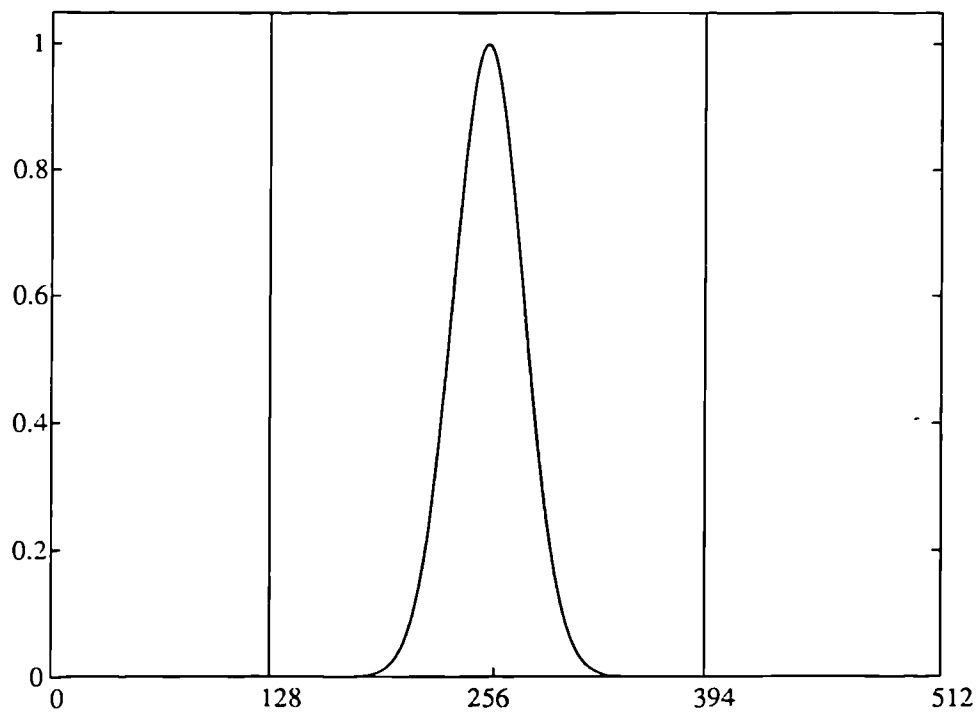


Figure 3.14(a) Frequency plane Gaussian apodisation

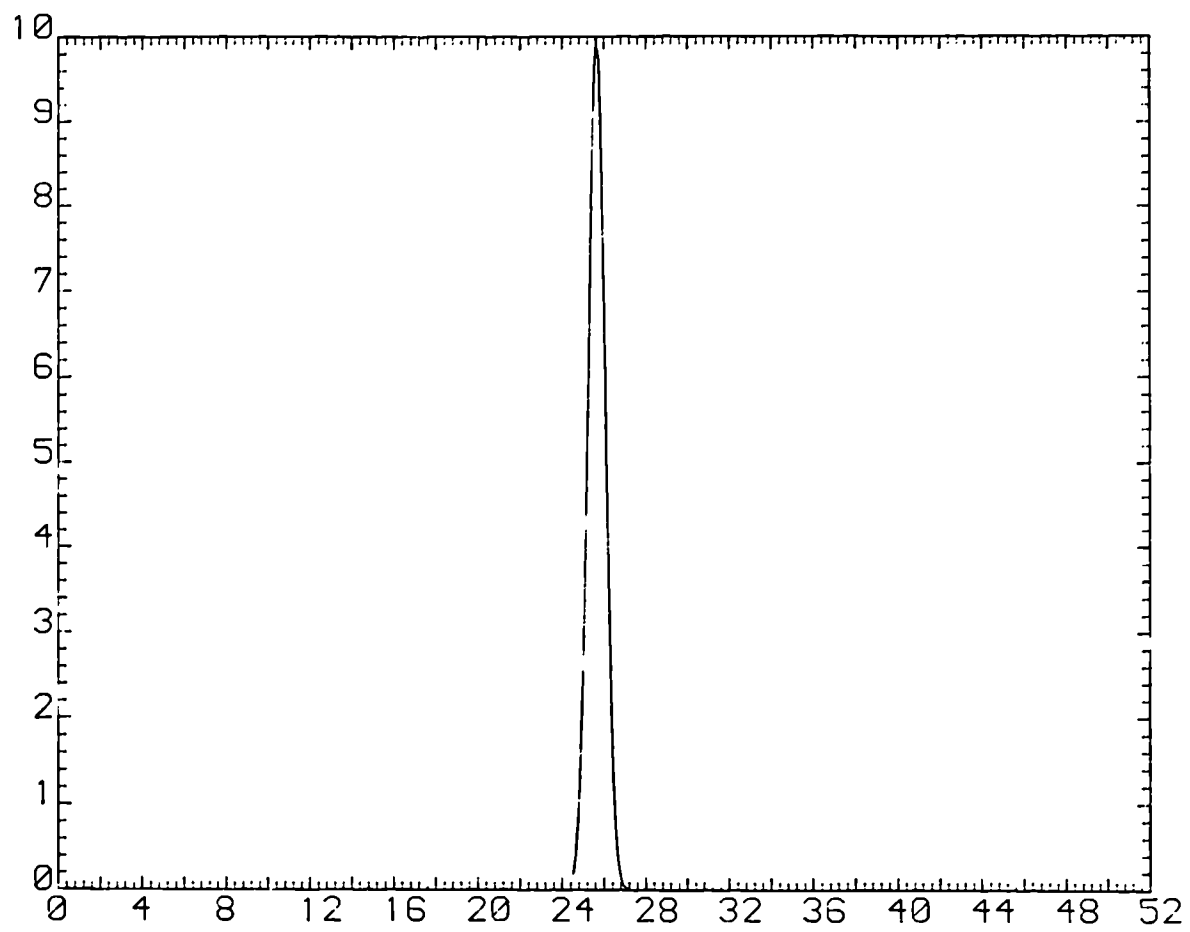


Figure 3.14(b) Fourier transform of the Gaussian function shown in Figure 3.14(a)

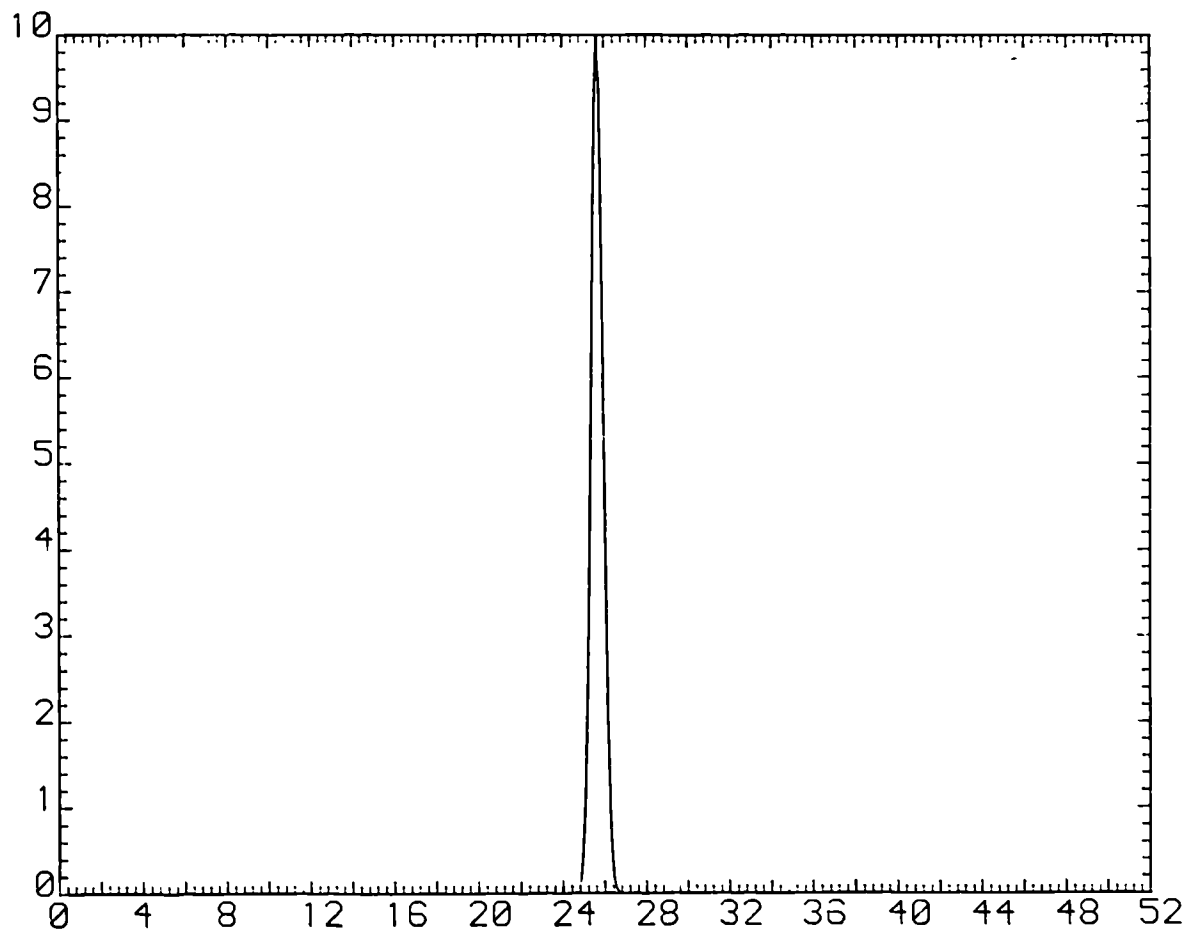


Figure 3.14(c) Intensity of auto-correlation function generated by bandlimited correlator

3.3.2 Joint transform correlator response to the test input imagery

As explained in the previous section the shape of the auto-correlation response is independent of the exact form of the image used as input to the joint transform correlator but is dependent on the bandpass characteristics of the frequency plane filter. The response resulting from uniform illumination of the full aperture available results in the ACF having the functional form of the sinc(x) function shown in Figure 3.13(a). This in fact was computed using image E270 as the reference and test input. Maintaining this image as the reference, the cross-correlation performance of the correlator was initially examined when operating without high frequency attenuation of the reconstruction beam ie. with a transfer function equivalent to an inverse filter.

Figure 3.15(a) shows a cross-section through the maximum of the field values resulting from the cross-correlation of E090 with M090 (ie. a broadside view of the Ford with that of the Mazda). The peak is normalised with respect to the ACF field peak value of E090 as are all the results shown in Figure 3.15. As may have been expected from the discussion of the properties of inverse filter in Chapter 1, the discrimination ability of the filter is very good, the cross-correlation peak field value being only 0.14 of the ACF amplitude, resulting in a detected intensity of less than 2%. However, Figures 3.15(b) and (c) reveal the unsuitability of the full bandwidth correlation for the orientation independent recognition of a 3-D object. They show the field values resulting from the cross-correlation of the Ford with views altered by $11\frac{1}{4}^\circ$ in orientation. This is done with both a rear view and a side view as the reference. Alteration in the appearance of the image occurs more rapidly with orientation changes of the car away from the rear view than they do for the same angular increment away from the side view. However, in both cases the filter is so sensitive that a correlation response is completely lost. Thus the view independent recognition of the Ford would require a very large number of filters with small increments in orientation between each view, especially if altitude as well as azimuth orientation changes had to be accommodated. The number of filters required would become too large even for a correlator system capable of a rapid filter search since the memory requirements to store all of the filter data would become impractically large. This requirement could be reduced by multiplexing many images onto a single filter as is done in the MACE SDF [6]. However, the tolerance to loss of correlation of the inverse filter to non-training set images seems so poor that the practicality of this approach must also be called into question.

Simulations were, therefore, run with a filter of reduced bandwidth. The frequency plane Gaussian of size 29.074 data points to the 1/e amplitude point was used as the apodizing function. Since the correlation plane distributions produced in the simulations were relatively localised, the effects of overlap of signal energy from neighbouring data cells due to the circular nature of the discrete correlation was minimal. For this reason the input array was not zero supplemented in the following simulations, leading to an output correlation plane data array the same size as that of the input array ie. 256x256 data points. The ACF peak height produced by the Fourier transform of the frequency plane Gaussian will be proportional to the integral of this function ie.

$$\text{ACF peak} \propto \int_0^\infty 2\pi r \exp\left(-\left(\frac{r}{w}\right)^2\right) dr \propto \pi w^2 \quad \{3.22\}$$

(where w is the distance to the 1/e point of the Gaussian). Since for the discrete Fourier transform :

$$g(0,0) = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} G(i,j) \quad \{3.23\}$$

when $w = 29.074$, the intensity of the computed ACF in the unsupplemented 256x256 array will, therefore, be : $I_0 = \frac{\pi}{256} (29.074)^2 = 107.61$.

The cross-correlations and modified ACF generated after the frequency removal procedure will be seen to be very much lower than this value. The technique is employed, however, to maintain a more uniform intra-class correlation response while increasing the difference between this and inter-class cross-correlations. The simulation results which follow are example results which demonstrate the improvements possible using this method.

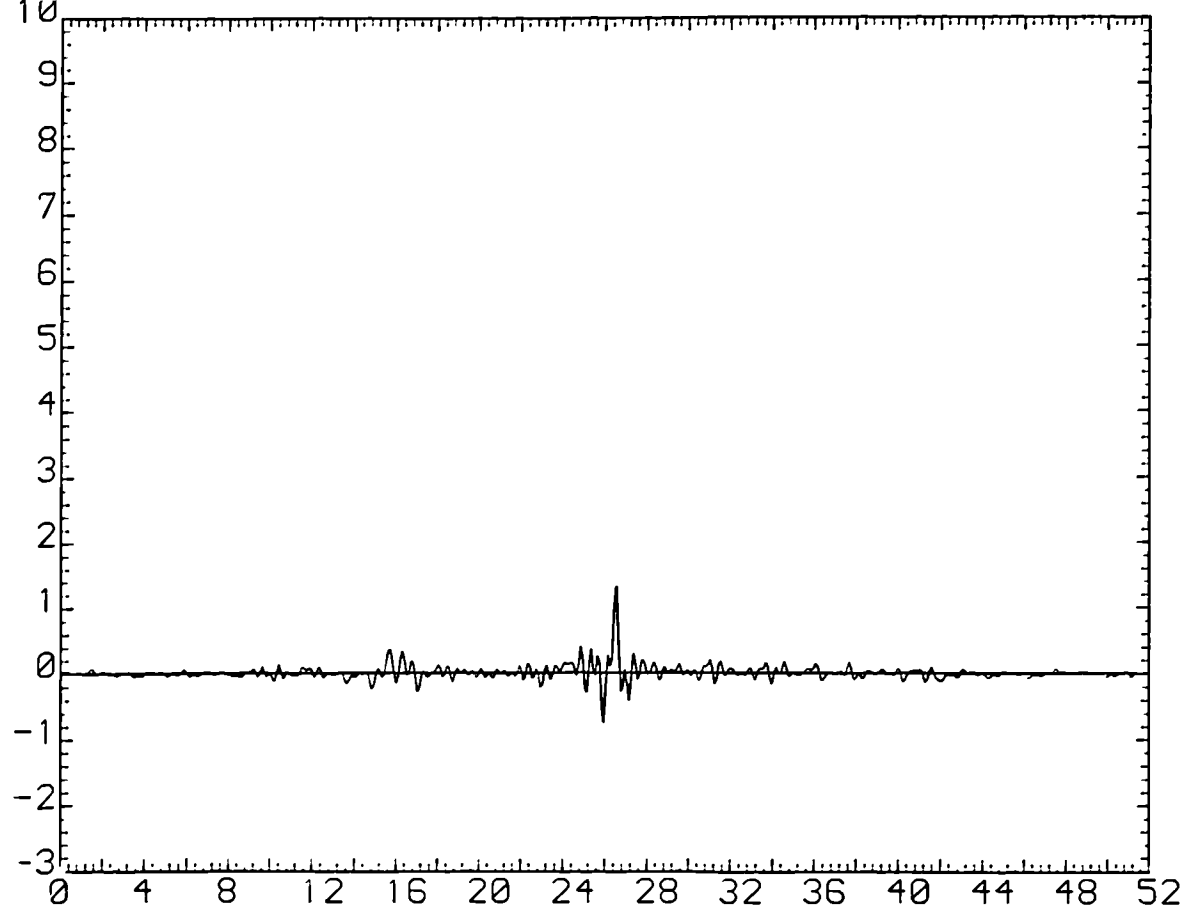


Figure 3.15(a) Full bandwidth cross-correlation of E090 with M090

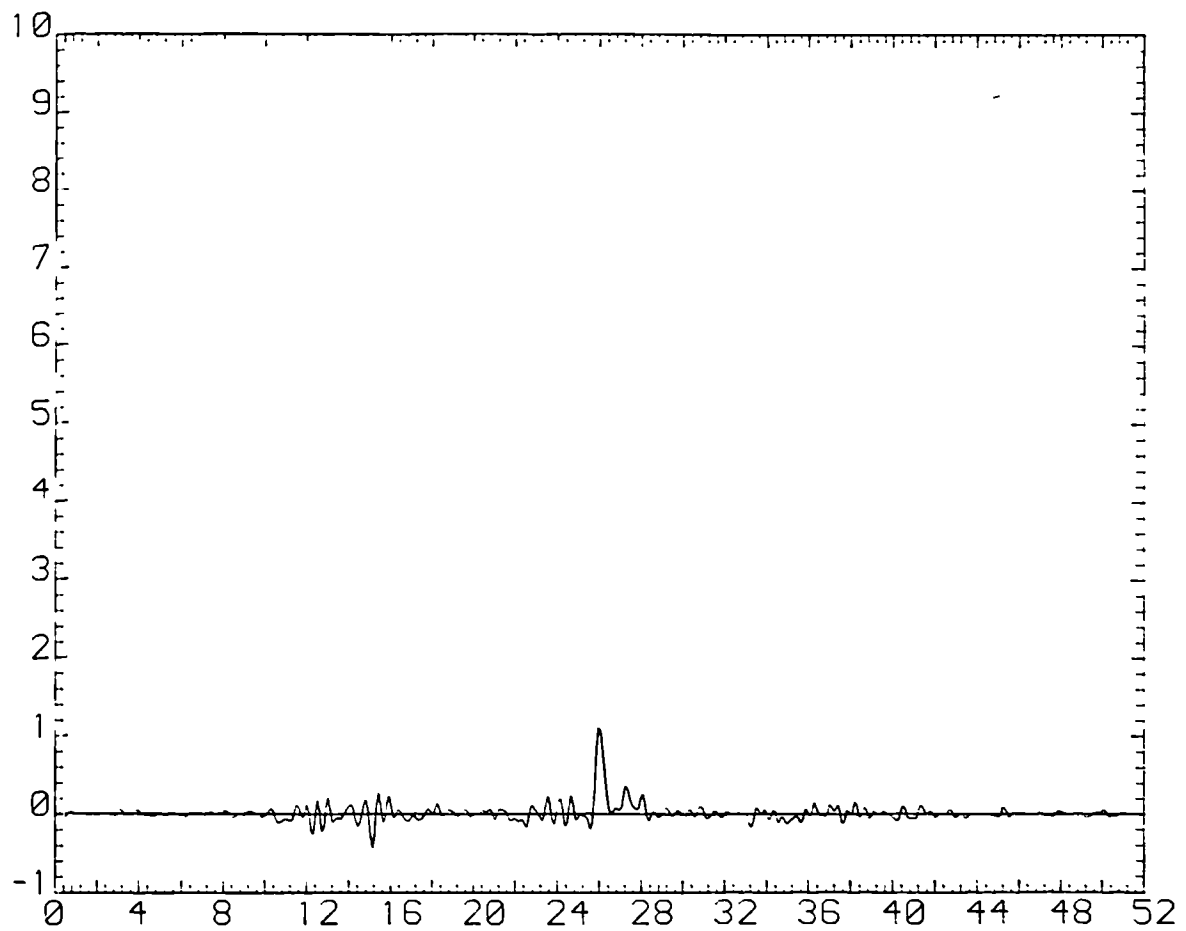


Figure 3.15(b) Full bandwidth cross-correlation of E090 with E101

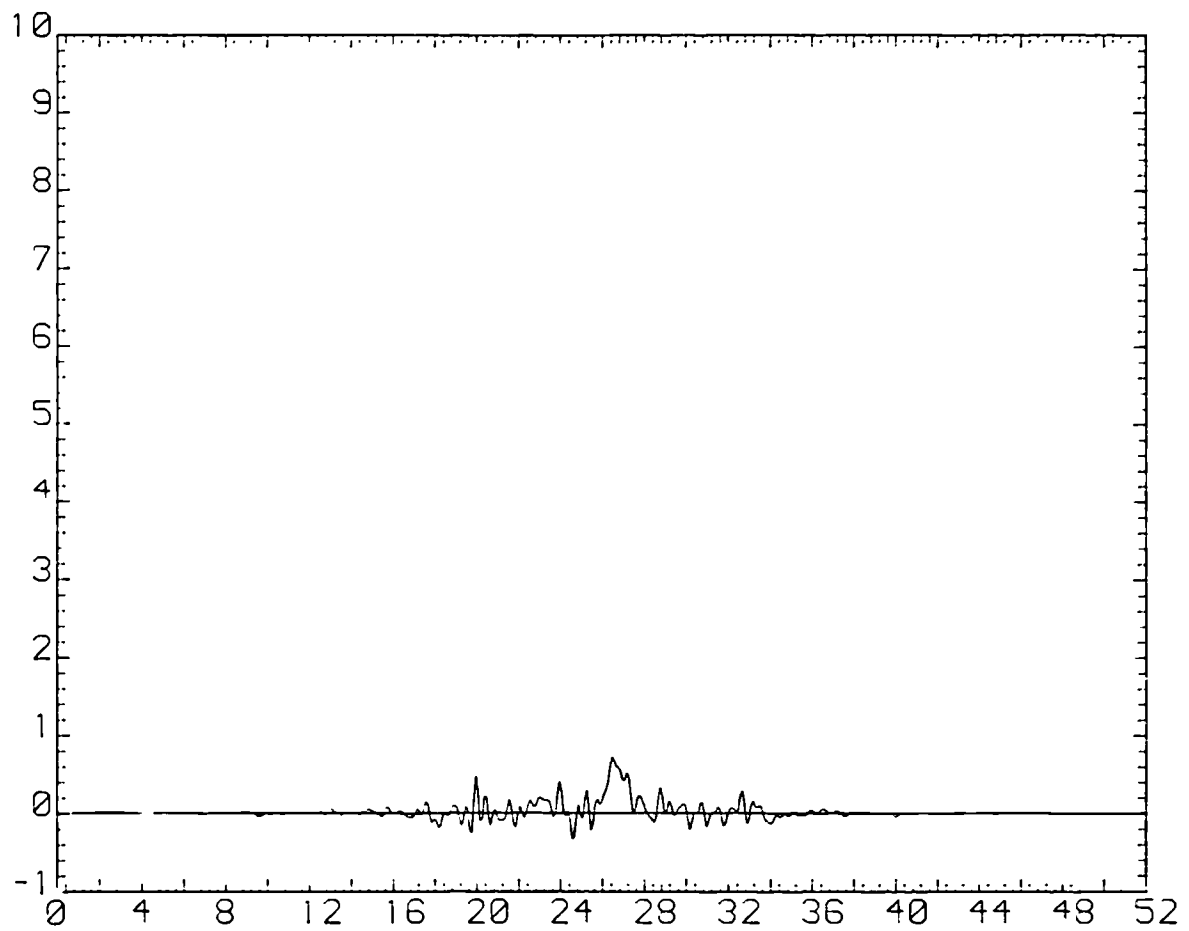


Figure 3.15(c) Full bandwidth cross-correlation of E180 with E191

3.3.3 Application of the discrete frequency removal technique

Figure 3.16 shows cross-sections and an isometric of the cross-correlation between E000 and E349 ie. between the front view of the Escort and a $11\frac{1}{4}^\circ$ rotated view. In comparison to the cross-correlation response when a full bandwidth is employed, the cross-correlation resulting when the frequency plane has a Gaussian apodization is sharply localised but still has a peak intensity which is only 7.3% that of the ACF (see Figures 3.16(a) and (b)). In addition, there is an appreciable sidelobe structure. Nevertheless, the complete collapse of the signal which occurred when a full bandwidth was employed has been prevented and a usable cross-correlation obtained.

The smaller reduction of the cross-correlation peak from that of the ACF in the bandlimited case is due to the high frequency terms being absent and so not contributing the ACF in the first place, rather than any increase in the absolute value of the cross-correlation peak intensity *per se*. Thus with a uniform beam of unit modulus illuminating the crystal, the ACF will have a peak intensity of $\left(\frac{1}{256}(256 \times 256)\right)^2 = 65536$ in the digital simulation as compared to 107.61 for the Gaussian illumination. (Experimentally the reduction will not be so great since the energy of the illuminating HeNe beam is approximately constant when adjusted to the correct size for the required frequency weighting. However, as mentioned above, a reduction of the correlation peak will occur due to the available energy being distributed over the broader correlation response resulting from a more localised frequency plane reconstruction beam.)

Since only the lower frequency terms will be significantly phase matched in the cross-correlation, it is these alone which will contribute to the peak. The higher frequency terms occurring over the relatively large perimeter area of the frequency plane will be randomly phased and hence only contribute random fluctuations to the correlation plane intensity distribution. In the joint transform correlator configuration these terms are, however, of unity modulus and so make a very large contribution to the ACF when phase matched. In contrast, in the Vander Lugt configuration the natural fall-off in amplitude with frequency of the spectrum of the input leads to a considerable apodisation of the frequency plane signal. Hence there is a reduced fall-off of the ACF peak with input image alterations from that of reference function, as has been noted in Chapters 1 and 2 in the comparison of the MSF and other filter types. It is to reduce this high frequency bias in the joint transform configuration that substantial

frequency plane apodisation must be employed, the Gaussian modulation being ideal for the reasons already explained.

A severe problem arises, however, when excluding the high frequency terms from the correlator in that the inter-class discrimination, when compared to the intra-class cross-correlation performance, is inadequate for reliable pattern recognition. In the specific example under consideration, the cross-correlation of the front views of the Escort and Mazda is actually *higher* than the cross-correlation of E000 with E349 (shown in Figure 3.16). The isometric of Figure 3.16(c) is normalised with respect to the E000/M000 cross-correlation. Cross-sections and an isometric of the inter-class signal are shown in Figures 3.17(a), (b) and (c) from which it can be seen that it has a localised peak with minimal sidelobe disruption and is indistinguishable from an intra-class response.

To improve the performance of the filter, discrete frequency components from within the Gaussian envelope of the frequency domain function are removed. These are chosen so as to maximise the difference in the cross-correlation of the intra-class and inter-class view with the expectation that the essentially randomly located discontinuities introduced into the Gaussian envelope by this procedure will not affect the localisation of the correlation peak, as demonstrated by the simulated correlation results presented in the previous section.

The method is implemented by the following steps :

1. Align the input images with the reference image such that the peak of the cross-correlation occurs on axis in each case.
2. Compute the spectrum of the intra-class and inter-class image and perform the required phase additions to the conjugated phase of the reference function spectrum; multiply the result by the Gaussian envelope function.

(These first two steps determine the complex field values that emerge just behind the frequency plane filter.)

3. Compute the sum of the real components for the two field distributions to determine the initial cross-correlation heights.

4. At each discrete frequency location of the two complex arrays subtract the intra-class real field component from the corresponding inter-class component and arrange the results in an ordered list of descending magnitude. Maintain an index of their locations in the array.
5. Remove each frequency component in turn (note that these will occur in pairs of equal magnitude due to the Hermitian symmetry of the field) and recompute the two field sums.
6. Continue Step (5) until the inter-class sum has reached a specified low value (maybe zero but in practice perhaps a small positive value, for the reasons to be discussed below).
7. Record in an integer array the indices corresponding to the location of the components that have been removed and at which points the transfer function of the filter should be zero. This array is then multiplied by the calculated field values of Step (2) to obtain the maximum discrimination filter. Fourier transformation of the resulting field values yields the corresponding intra-class and inter-class cross-correlation signals.

Figures 3.18 and 3.19 show the results of the procedure applied to the cross-correlations of E000 with E349 and M000 and should be compared to the unmodified cross-correlations shown in Figures 3.16 and 3.17. The intra-class cross-correlation peak has only been reduced slightly (from 7.9 to 7.5) but with a much reduced sidelobe (from 0.25 to 0.067 of the peak) and somewhat improved localisation, the full width at half maximum (FWHM) being 3.95 as compared to 5.1 data points. However, the inter-class cross-correlation (shown in Figure 3.19 normalised with respect to the E000/E439 cross-correlation) has been almost totally suppressed on-axis and, due to its remaining localised, the sidelobe structure does not disrupt the discrimination; the cross-sections in Figures 3.19(a) and (b) are taken through the maximum which occurs at location (126,128) and has a value of 0.133 of the intra-class cross-correlation. The remaining sidelobe structure is due to the relatively broad base of the original cross-correlation function which results from the Fourier transform of the bandwidth reducing Gaussian envelope applied in the frequency domain.

In common with the SDF algorithm, the present method only controls the on-axis correlation response and so relies on a relatively localised correlation peak for its

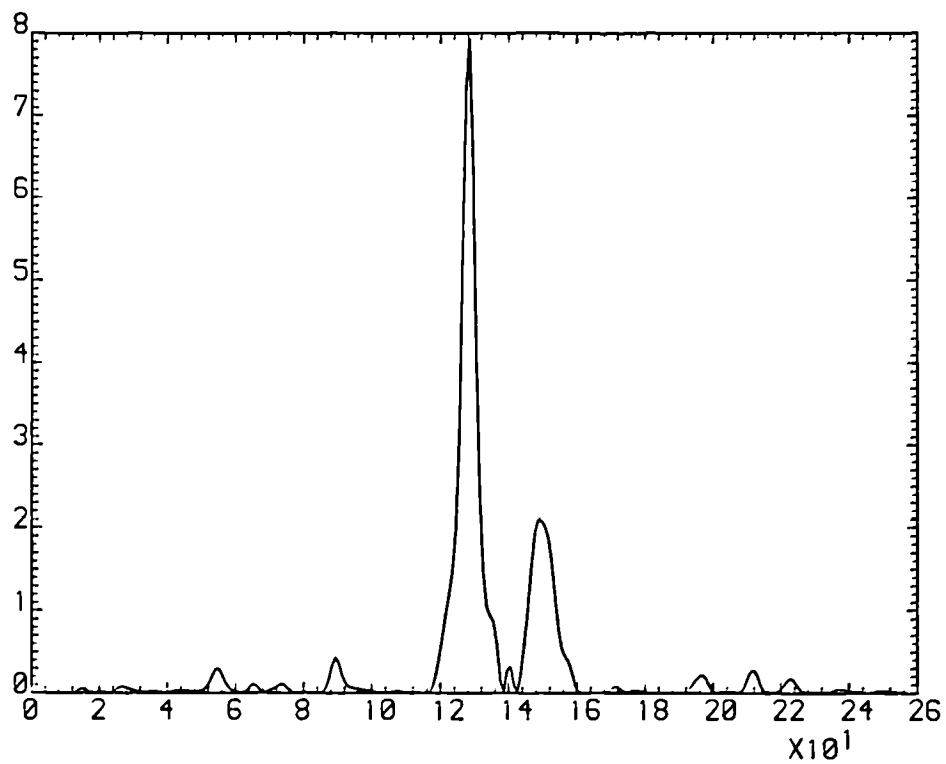


Figure 3.16(a) Cross section at $y = 0$ through intra-class cross-correlation of E000 and E349

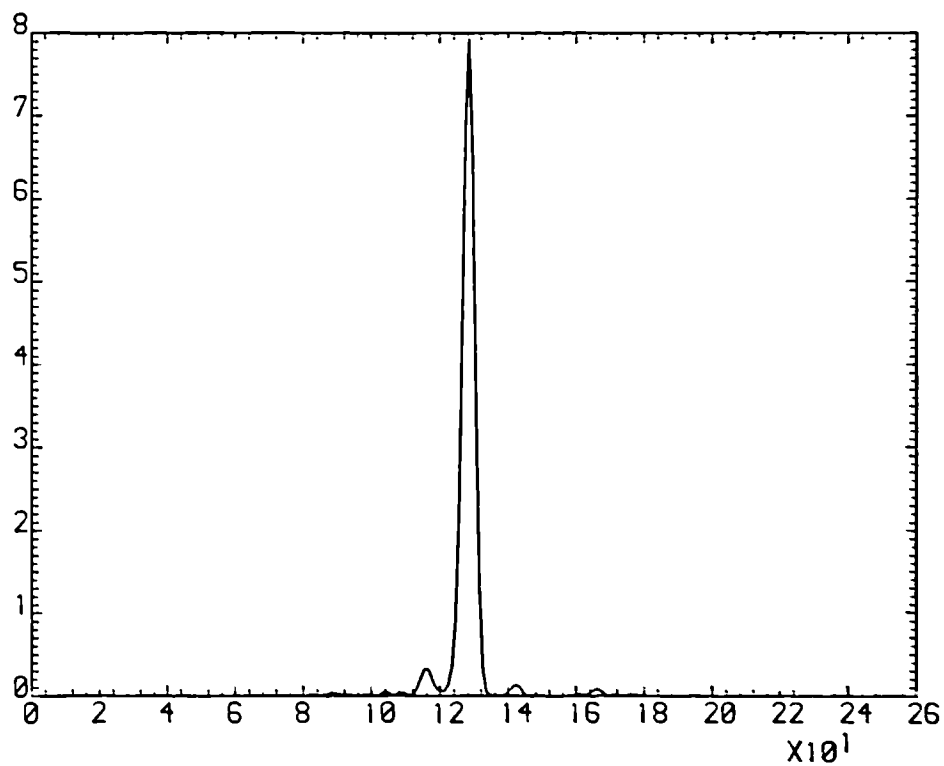


Figure 3.16(b) Cross section at $x = 0$ through intra-class cross-correlation of E000 and E349

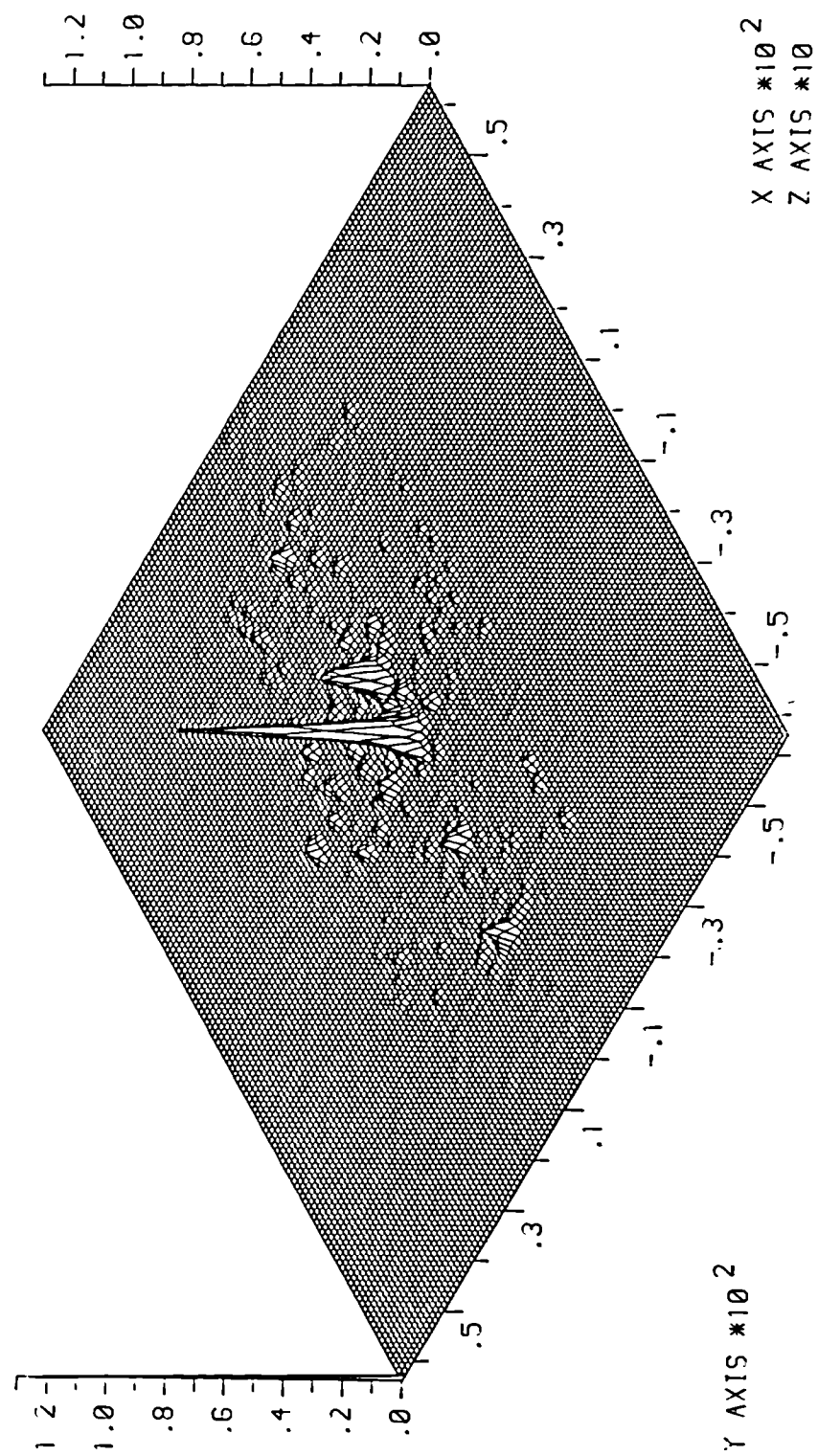


Figure 3.16(c) Isometric of intra-class cross-correlation of E000 and E349

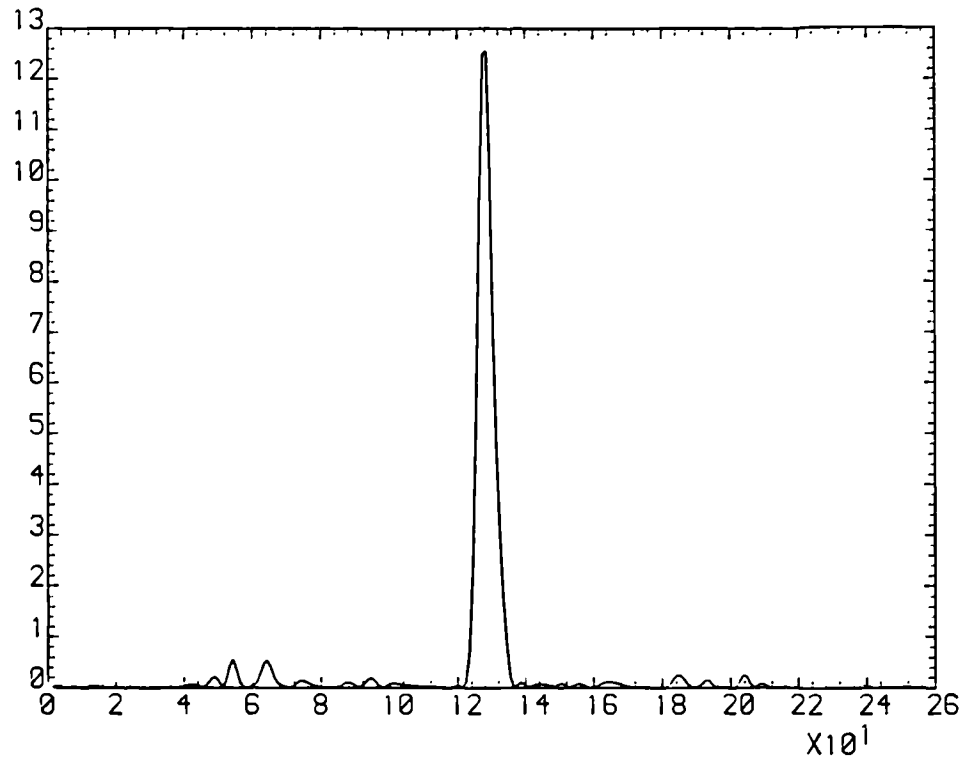


Figure 3.17(a) Cross-section at $y = 0$ of inter-class cross-correlation of E000 and M000

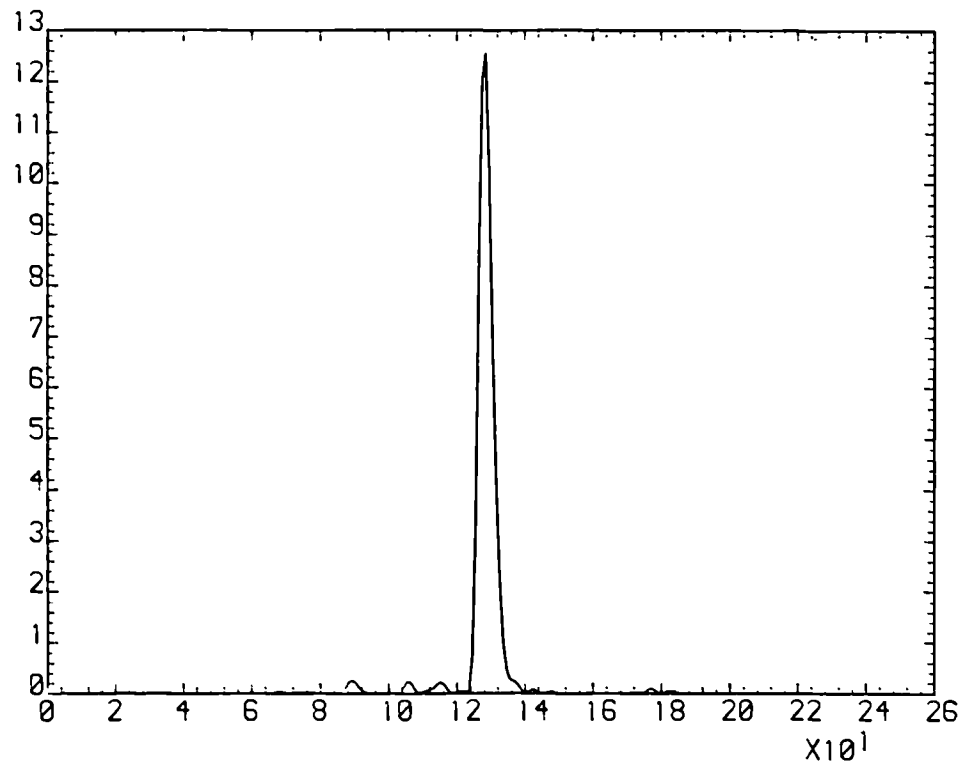


Figure 3.17(b) Cross-section at $x = 0$ of inter-class cross-correlation of E000 and M000

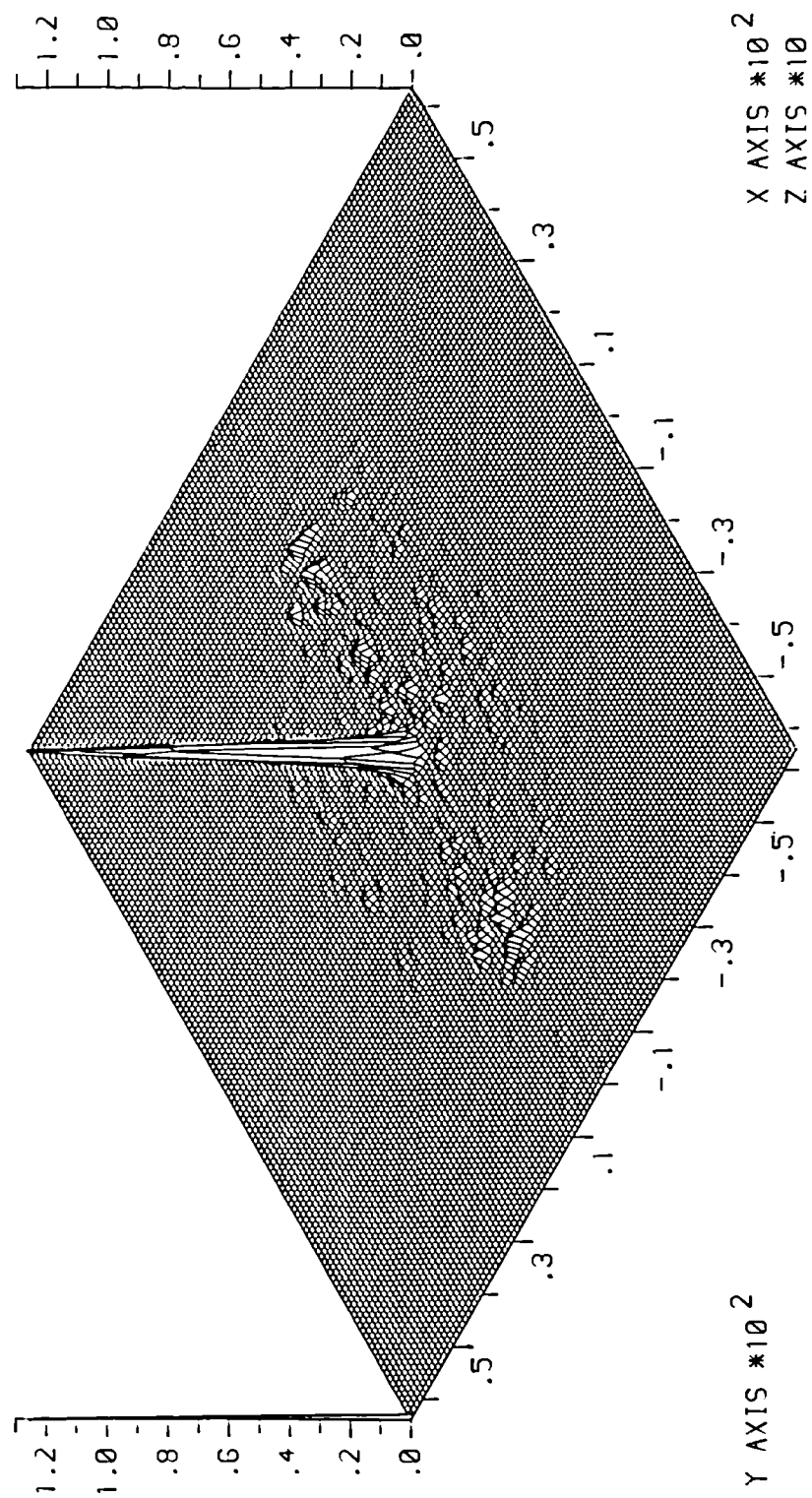


Figure 3.17(c) Isometric of inter-class cross-correlation of E000 and M000

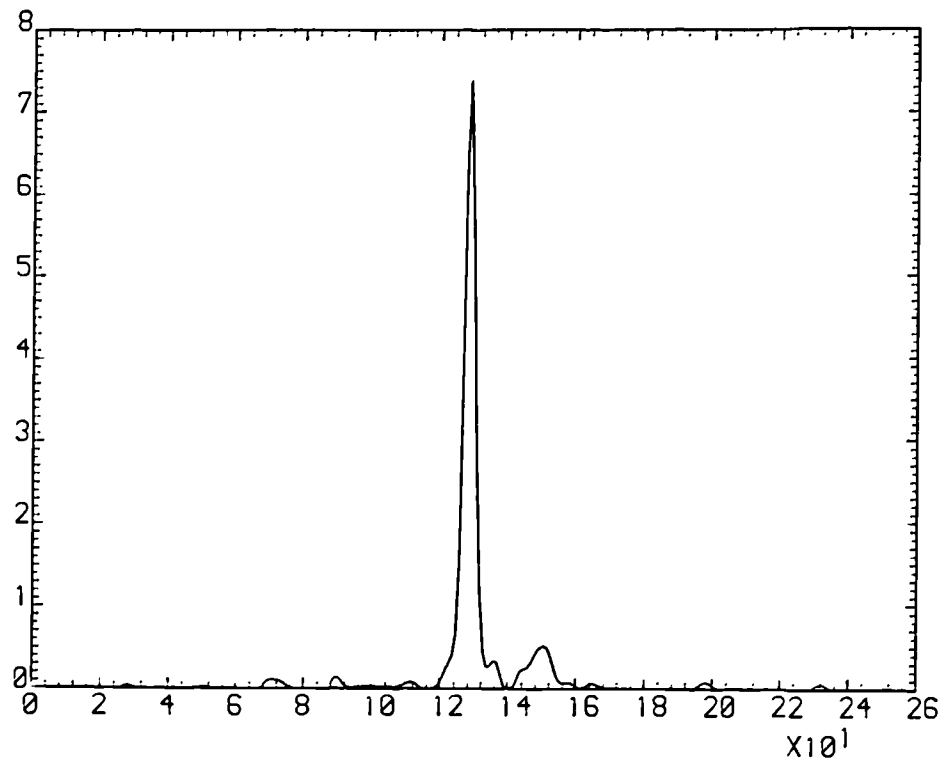


Figure 3.18(a) Cross-section at $y = 0$ of intra-class cross-correlation of E000 and E349 after discrete frequency removal

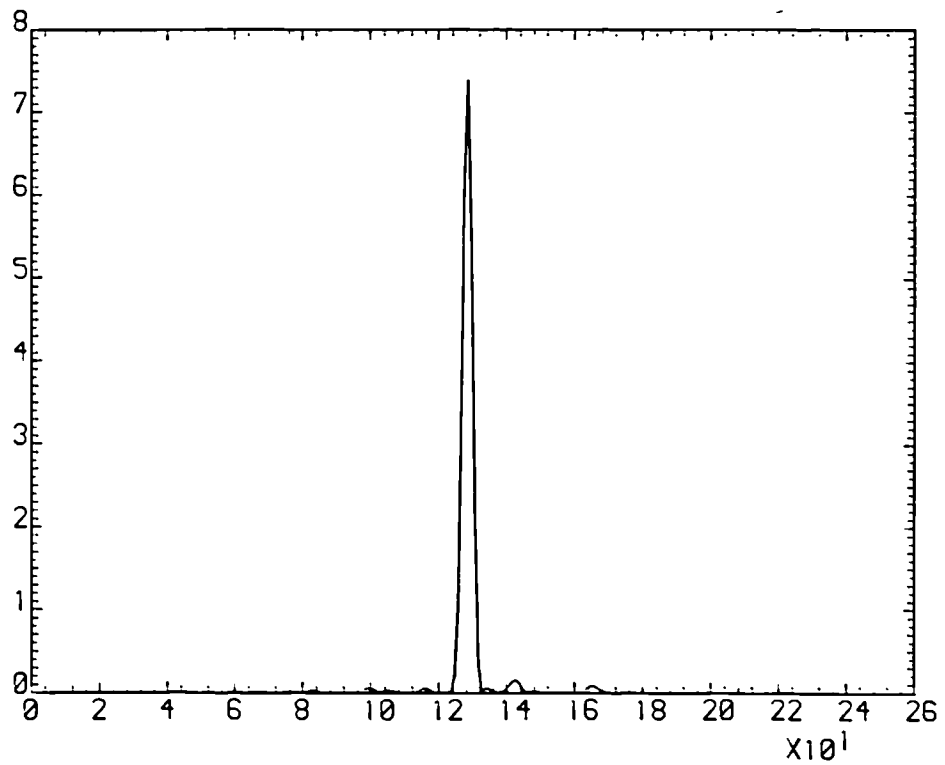


Figure 3.18(b) Cross-section at $x = 0$ of intra-class cross-correlation of E000 and E349 after discrete frequency removal

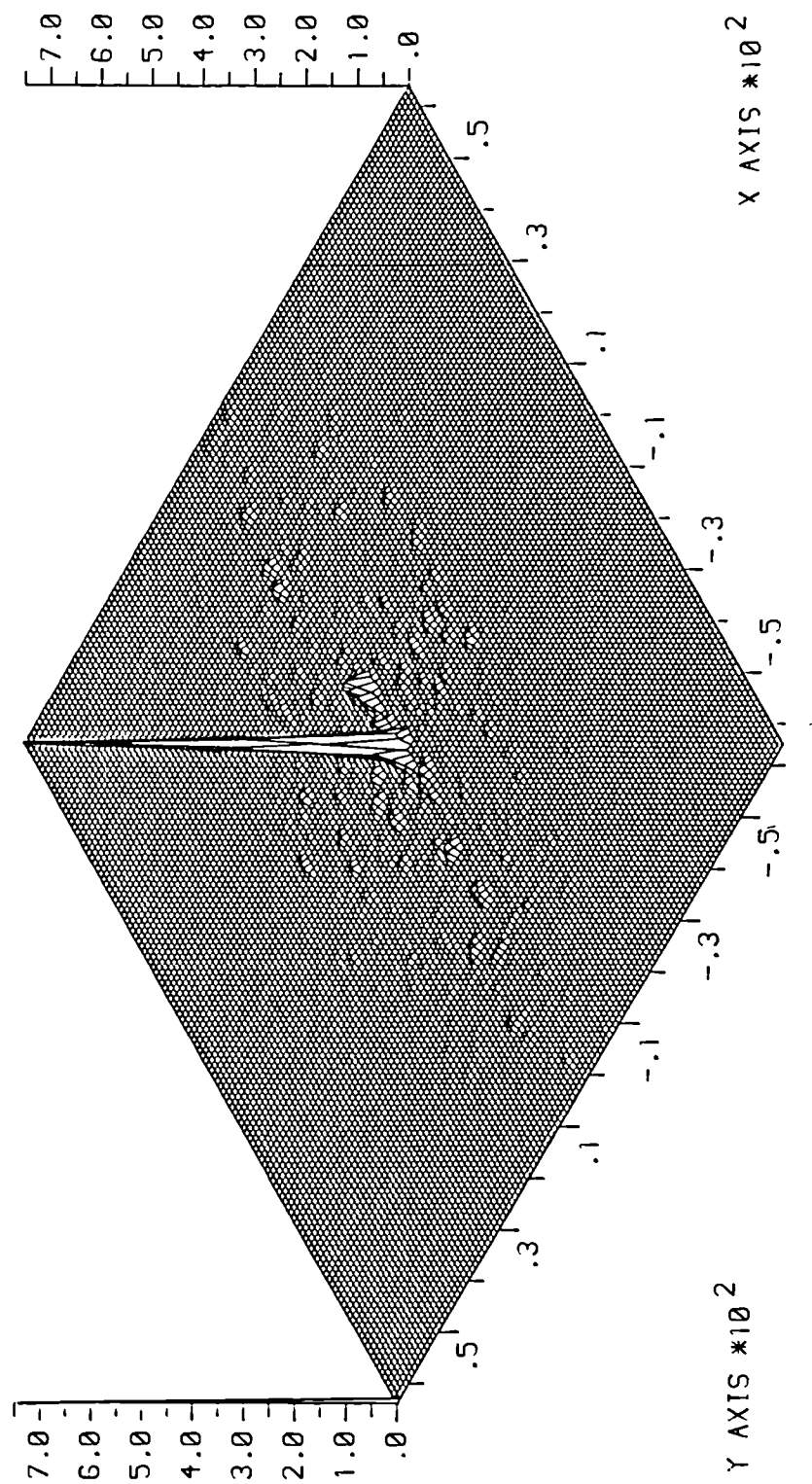


Figure 3.18(c) Isometric of intra-class cross-correlation of E000 and E349 after discrete frequency removal

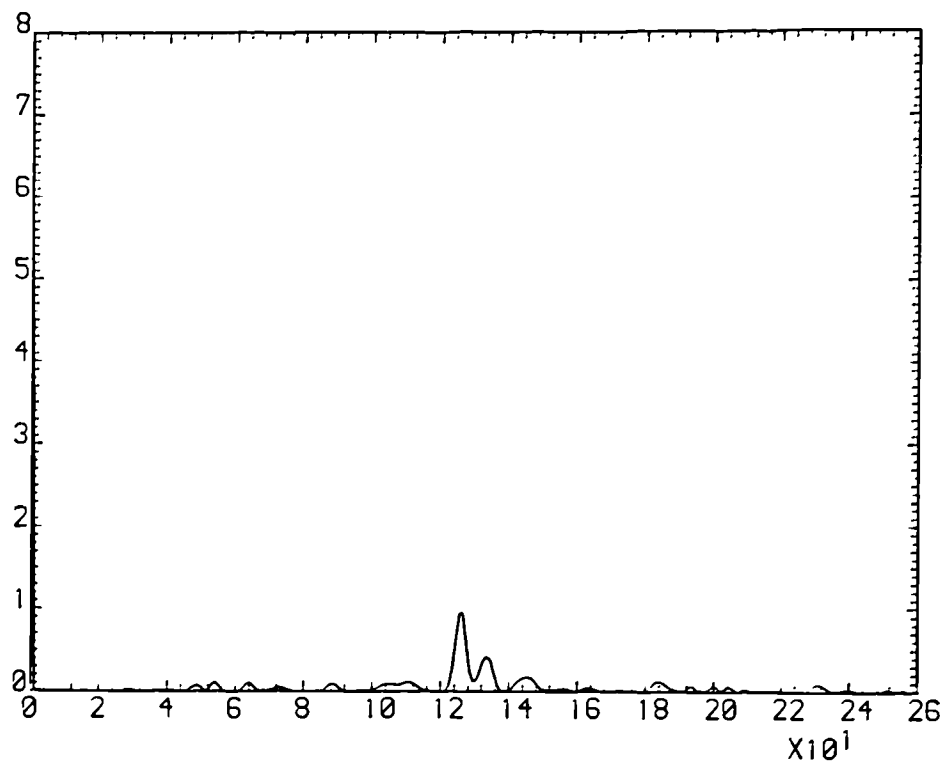


Figure 3.19(a) Cross-section at $y = 0$ of inter-class cross-correlation of E000 and M000 after discrete frequency removal

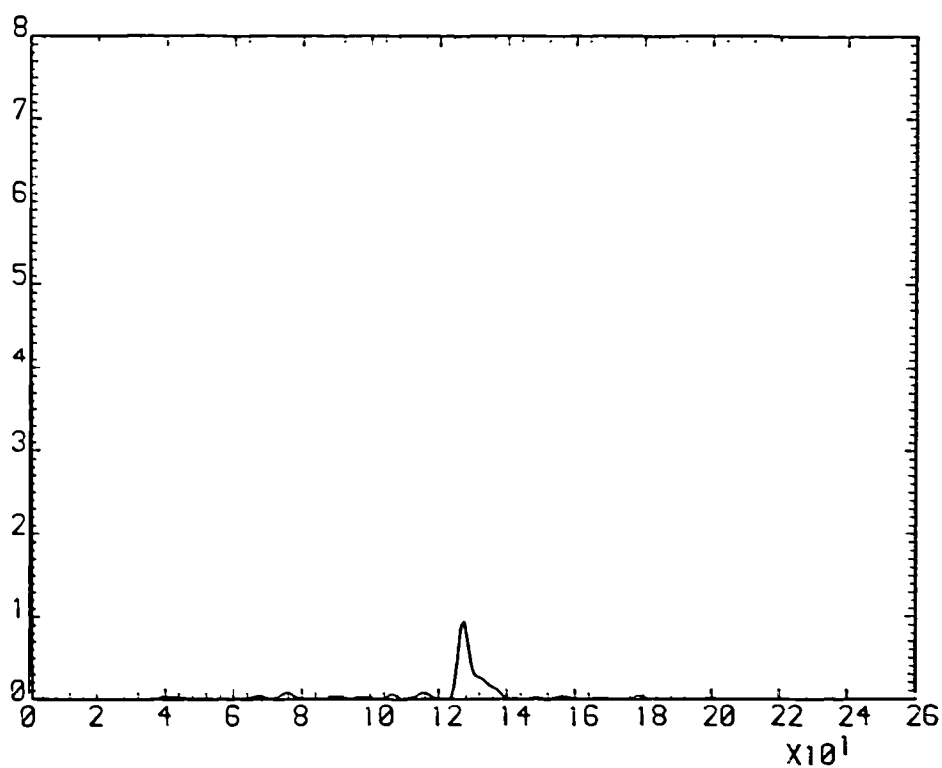


Figure 3.19(b) Cross-section at $x = 0$ of inter-class cross-correlation of E000 and M000 after discrete frequency removal

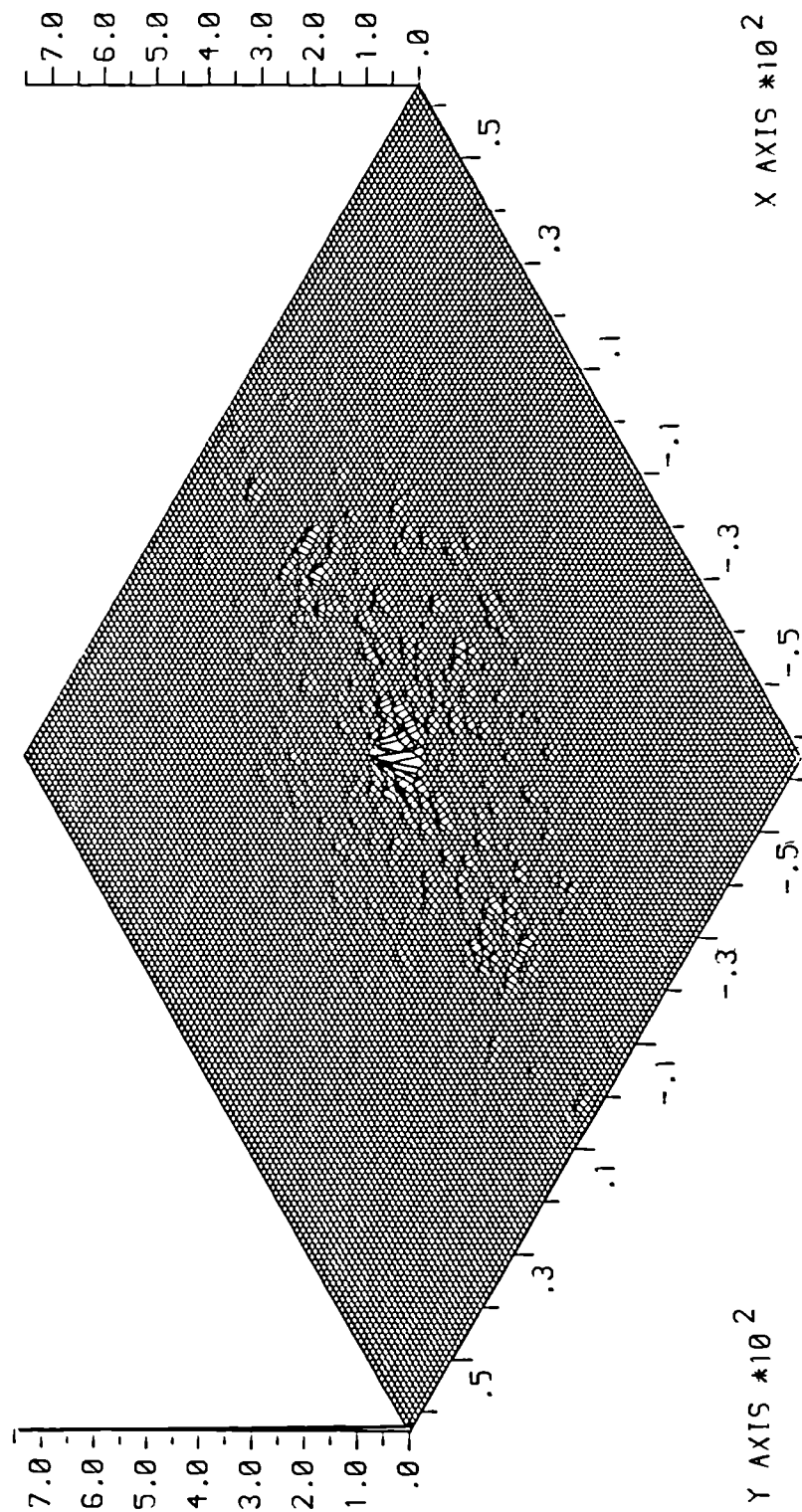


Figure 3.19(c) Isometric of intra-class cross-correlation of E000 and M000 after discrete frequency removal

effectiveness. The result shown in Figure 3.19 demonstrates that it works adequately well with the widths of correlation that result from a frequency plane bandwidth limited to the degree required for useful intra-class tolerance to be achieved. Figures 3.20(a) and (b) show cross-sections through the ACF resulting from the modified filter. The central peak is the relatively localised Gaussian resulting from the Fourier transform of the frequency plane Gaussian envelope. The discontinuities introduced into the frequency plane have led to some disruption of the correlation plane field values away from the peak which are of relatively low spatial frequency as compared to those noted in the correlation simulations of the previous section (eg. see Figure 3.17). This is due to the low-pass filtering effect of the frequency plane Gaussian.

Figure 3.20(c) and (d) show cross-sections and an isometric of the intensity of the ACF. The sidelobe disruption is almost imperceptible, the distribution appearing the same as that before frequency removal. Thus the sharp discontinuities introduced into the Gaussian envelope function have led to minimal disruption of the localisation of the ACF as the simulations of the previous section suggested would be the case. Of course, the ACF peak value has fallen due to removal of the selected frequency components as these are all phase matched for the auto-correlation and so contribute to the ACF peak. The maximum of the modified ACF is 0.1245 of the unaltered ACF peak intensity. However, it is still 3.52 times larger than the intra-class E000/E349 cross-correlation peak. Thus if the filter is not exactly matched to the input a large reduction in the correlation signal will result. However, in practice, in an unconstrained pattern recognition task only rarely will the input exactly match a reference template. Thus the most important requirement of the filter is to produce a correlation peak that remains localised and of adequate magnitude to maintain discrimination against non-target input objects. The simulation results on the car images suggest the frequency removal technique improves the filter performance in these respects.

The example chosen for examination of the effectiveness of the frequency removal technique was particularly demanding in that the inter-class cross-correlation was initially higher than that of the intra-class. As can be seen from the examples shown in Figure 3.12 the front views of the Escort and Mazda are of the same basic size and shape, the wheels providing a particularly strong similarity.

When the Escort is rotated even slightly from the direct head-on orientation, the horizontal geometry alters very quickly as does the wheel shape and apparent spacing; only the vertical spacing of features remains unaltered and probably it is these that

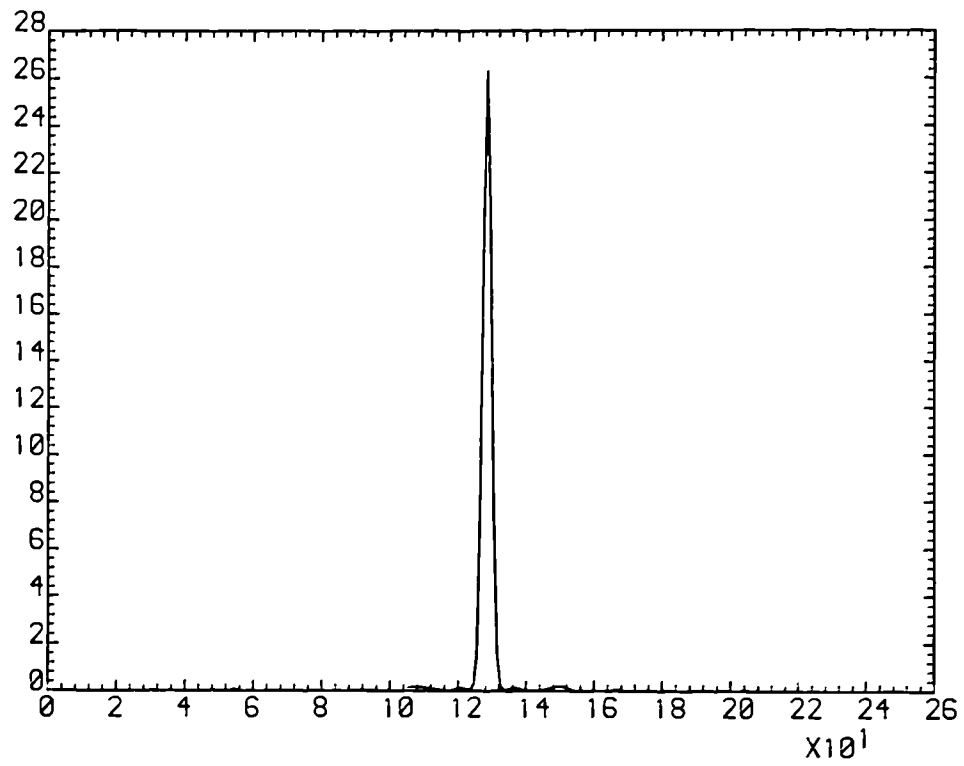


Figure 3.20(a) Cross-section at $y = 0$ of auto-correlation function of E000 after discrete frequency removal

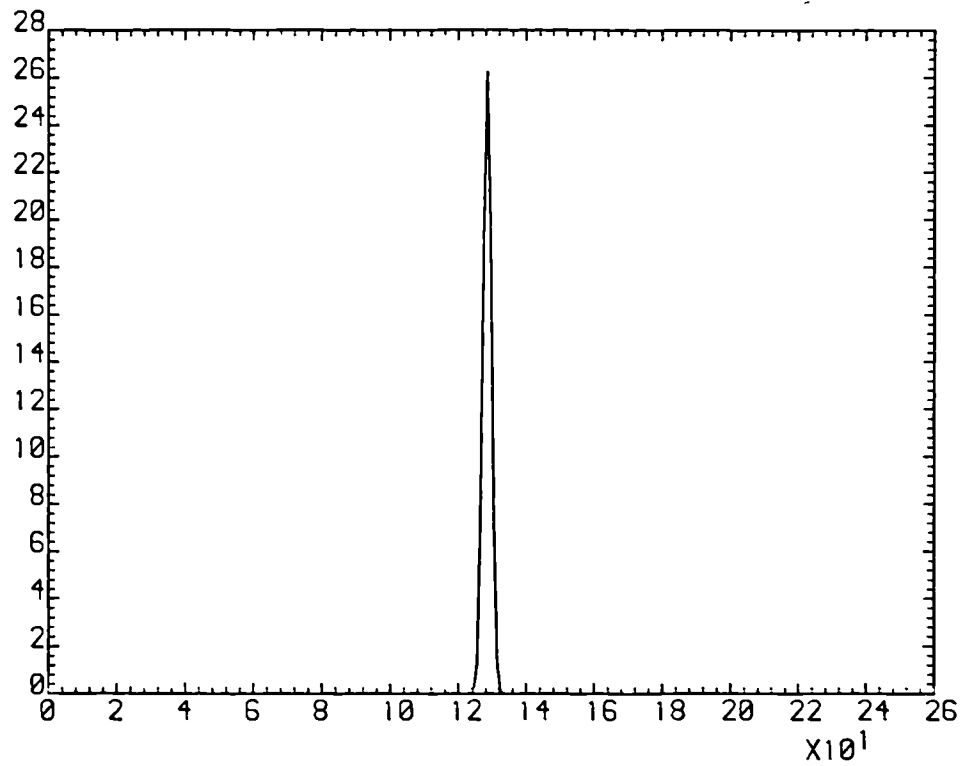


Figure 3.20(b) Cross-section at $x = 0$ of auto-correlation function of E000 after discrete frequency removal

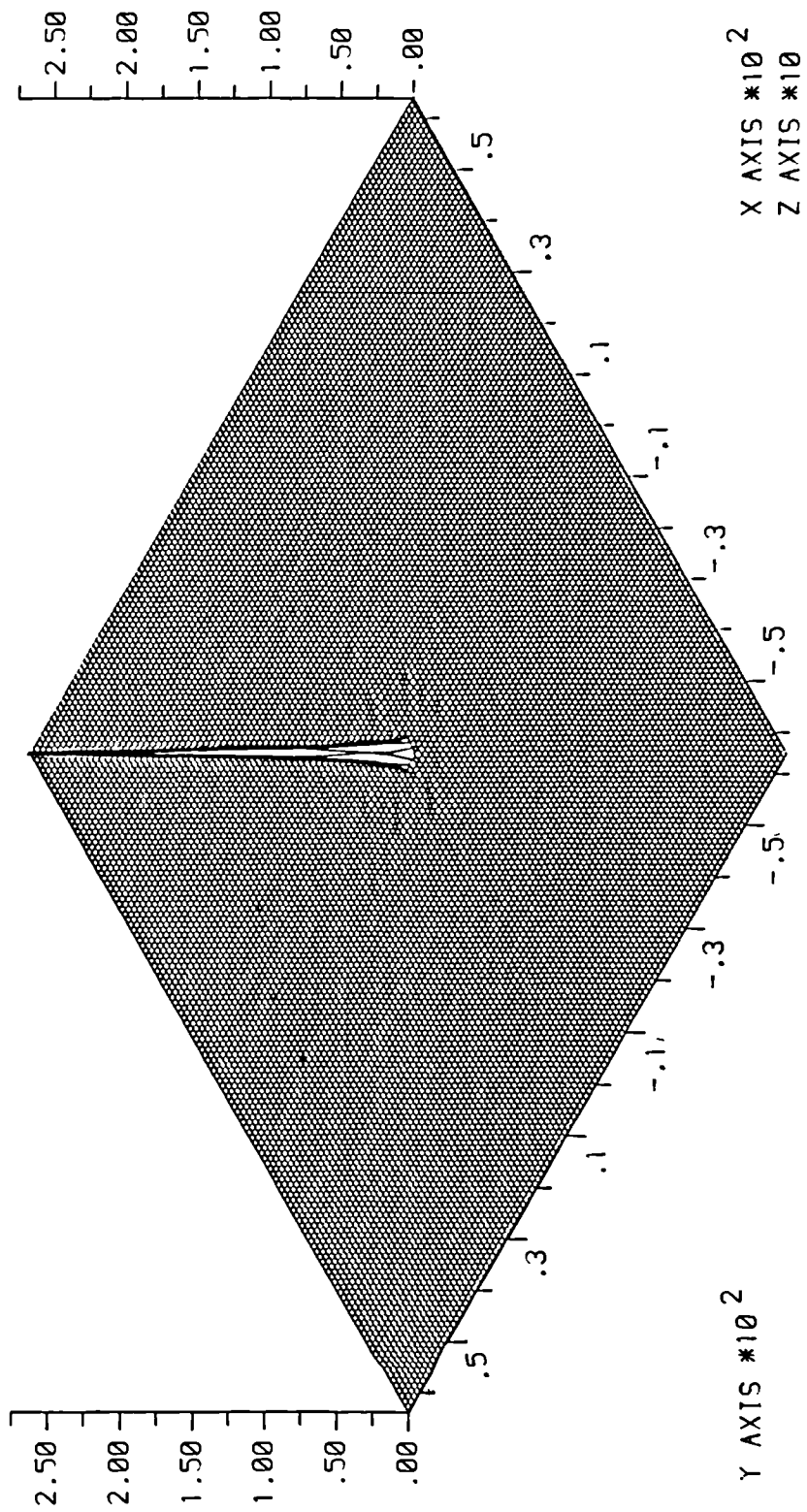


Figure 3.20(c) Isometric of auto-correlation of E000 after discrete frequency removal

provide for the maintenance of some correlation signal. In addition, it must be remembered that the correlation was performed on images of much reduced bandwidth than those shown in Figure 3.12. (This point will be considered further in Chapter 4 when the experimental work on bandpass correlation is described. Experimentally obtained reconstructions of bandlimited images are shown in Figure 4.16 of that Chapter.)

A maximum discrimination between two intra-class and inter-class images is set by the intrinsic difference between the images. This is shown clearly by the example under examination. As frequency components are removed the discrimination improves to a maximum but then remains constant with further removal of components. This occurs as all the terms are exhausted which have negative real components for the M000 cross-correlation but positive real components for the E349 cross-correlation. From that point on the terms removed have positive real components for both cross-correlations, the difference in magnitude becoming smaller the further the process proceeds. This effect is shown in Figure 3.21 for the discrimination of E349 and M000 using the filter matched to E000. In this graph, the sum of the real field components emerging directly behind the filter (equivalent to the magnitude of the on-axis cross-correlation peak) are plotted against the number of frequency terms removed from the ordered list of real component magnitude differences produced in Step (4) of the procedure outline above. The inter-class cross-correlation starts above that of the intra-class cross-correlation but initially falls rapidly, the two correlations becoming equal after the removal of about 200 frequency components. The fall of the inter-class sum can be seen to follow an approximately exponential decay whereas that for the intra-class builds to a maximum after a further 300 components have been removed and then remains approximately constant until a total of 1000 components are zeroed. From this point on, in order to obtain a further decrease in the inter-class response, frequency components with a positive real component must be removed from the intra-class signal, leading to a decrease in the cross-correlation response. However, at this point the inter-class sum is still falling relatively rapidly. The difference between the two sums continues to increase as further components are removed, although progressively more slowly as the inter-class sum flattens out after the removal of about 4000 components. It is at this point that the removal is halted since a continuation would result in a decrease in the difference between the two correlation responses, the intra-class sum falling faster than that of the inter-class. Thus, for the bandwidth over which the filter is operated, the maximum possible difference between the two signals has been realised (although of course the *ratio* between the two cross-correlations could be significantly increased by pushing the inter-class curve further towards zero). It is for this reason that the inter-class response

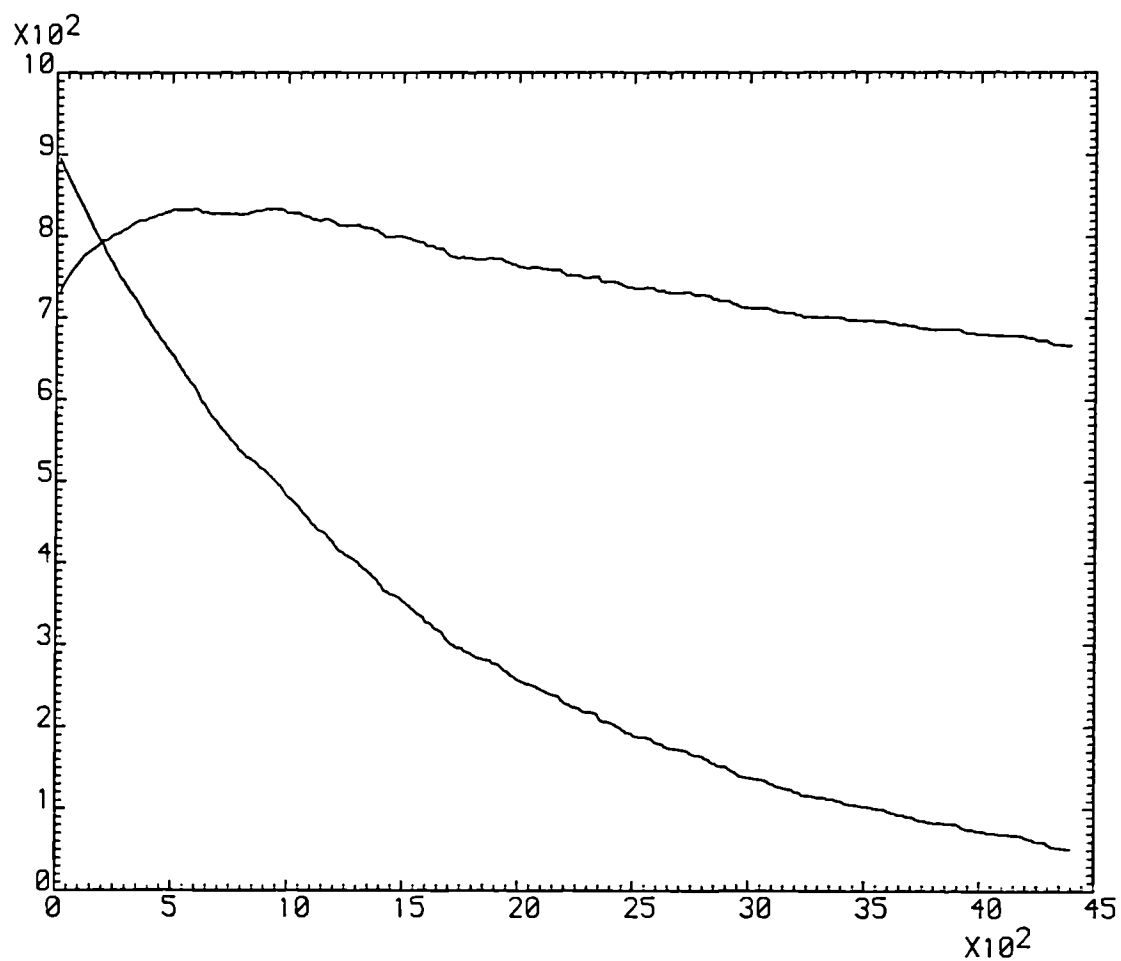


Figure 3.21 Plot of the relative intra-class and inter-class cross-correlation magnitudes of E000 with E349 and M000 as discrete frequency components are removed

is not forced to a zero value. When the modulus squared of the field is taken to simulate the observable correlation plane distributions, the plots shown in Figure 3.19 indicate that the residual on-axis positive field value leads to only minor disruptions of the inter-class response, discrimination being readily achievable.

As a further example of the improvement that can be achieved by the use of selective frequency removal, the process was applied to the maximization of the discrimination between the Escort rear view, E180, and that of the Mazda while simultaneously maintaining a high intra-class cross-correlation with an Escort view rotated by $11\frac{1}{4}^\circ$ from the direct rear view. Again, as with the head-on view, this is a situation in which correlation is quickly lost with small changes of orientation due to the rapid changes of aspect that occur. In contrast to the previous example, the inter-class cross-correlation is initially lower with respect to the intra-class response, the rear views of the two cars (see Figure 3.12) having quite distinctive differences. However, the wheels are a strong common feature, particularly in low-pass filtered images (see Figure 4.16).

Figure 3.22 shows cross-sections through the maximum of the intra-class cross-correlation intensity. It can be seen from these graphs that the peak initially occurs off-axis, is broadened in the x-direction (FWHM=9 data points) and has a significant sidelobe which is 0.286 of the maximum. The peak height is only 0.293 of the unmodified ACF peak intensity. Thus, as expected, the change in aspect away from the rear view has led to a considerable degradation from the auto-correlation response, even with the reduced bandwidth being employed. Figure 3.23 shows cross-sections and an isometric of the inter-class cross-correlation. The isometric shows the correlation plane to be strongly disrupted by secondary peaks. In addition, the maximum peak is of a localised nature which could readily lead to a false recognition, its magnitude being 0.86 that of the intra-class peak shown in Figure 3.22 (the isometric being normalised with respect to this).

The frequency removal procedure was applied to these results with the improvements shown in Figures 3.24 - 3.26 resulting. Figure 3.24 shows the enhanced intra-class cross-correlation response. The peak intensity has been increased by 2.16 times over the initial cross-correlation, the increase resulting in a reduction of the relative sidelobe magnitude to 0.09 of the peak. The FWHM has also reduced to 5.6 pixels due to increased phase matching in the frequency plane, the amplitude discontinuities introduced by the frequency removal process not having any noticeable deleterious effect. It is interesting to note, however, that the FWHM of the peak in the y-direction is less (3.4 pixels) and is approximately the same as that before frequency removal (compare with

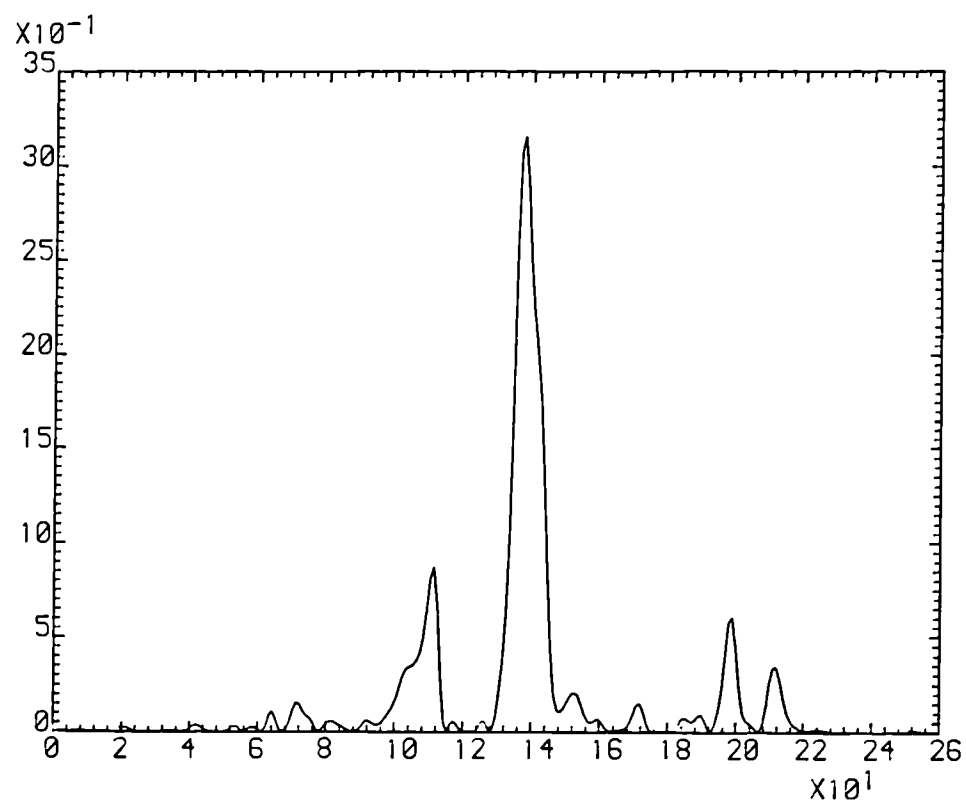


Figure 3.22(a) Cross-section at $y = 0$ of intra-class cross-correlation of E180 and E191

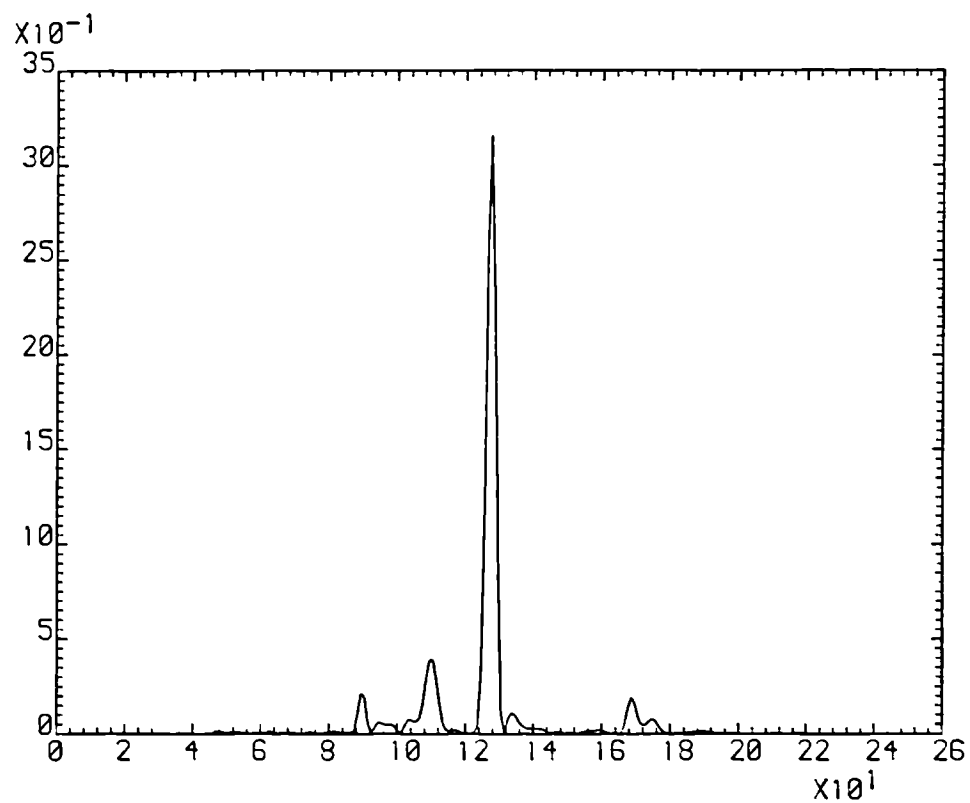


Figure 3.22(b) Cross-section at $x = 0$ of intra-class cross-correlation of E180 and E191

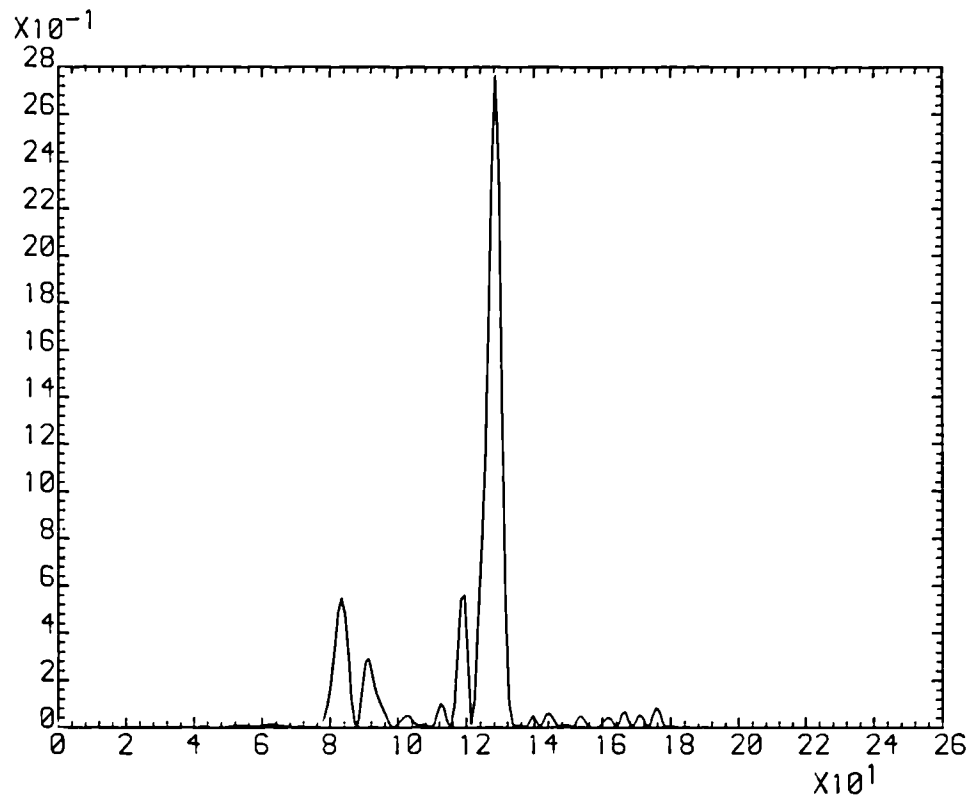


Figure 3.23(a) Cross-section at $y = 0$ of inter-class cross-correlation of E180 and M180

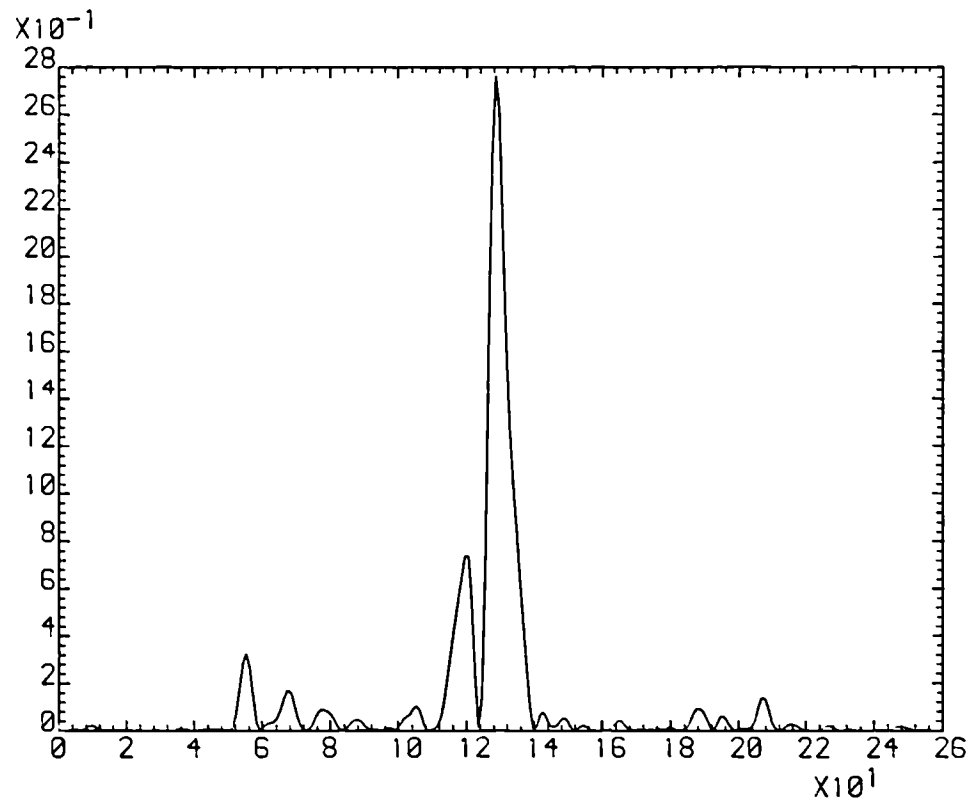


Figure 3.23(b) Cross-section at $x = 0$ of inter-class cross-correlation of E180 and M180

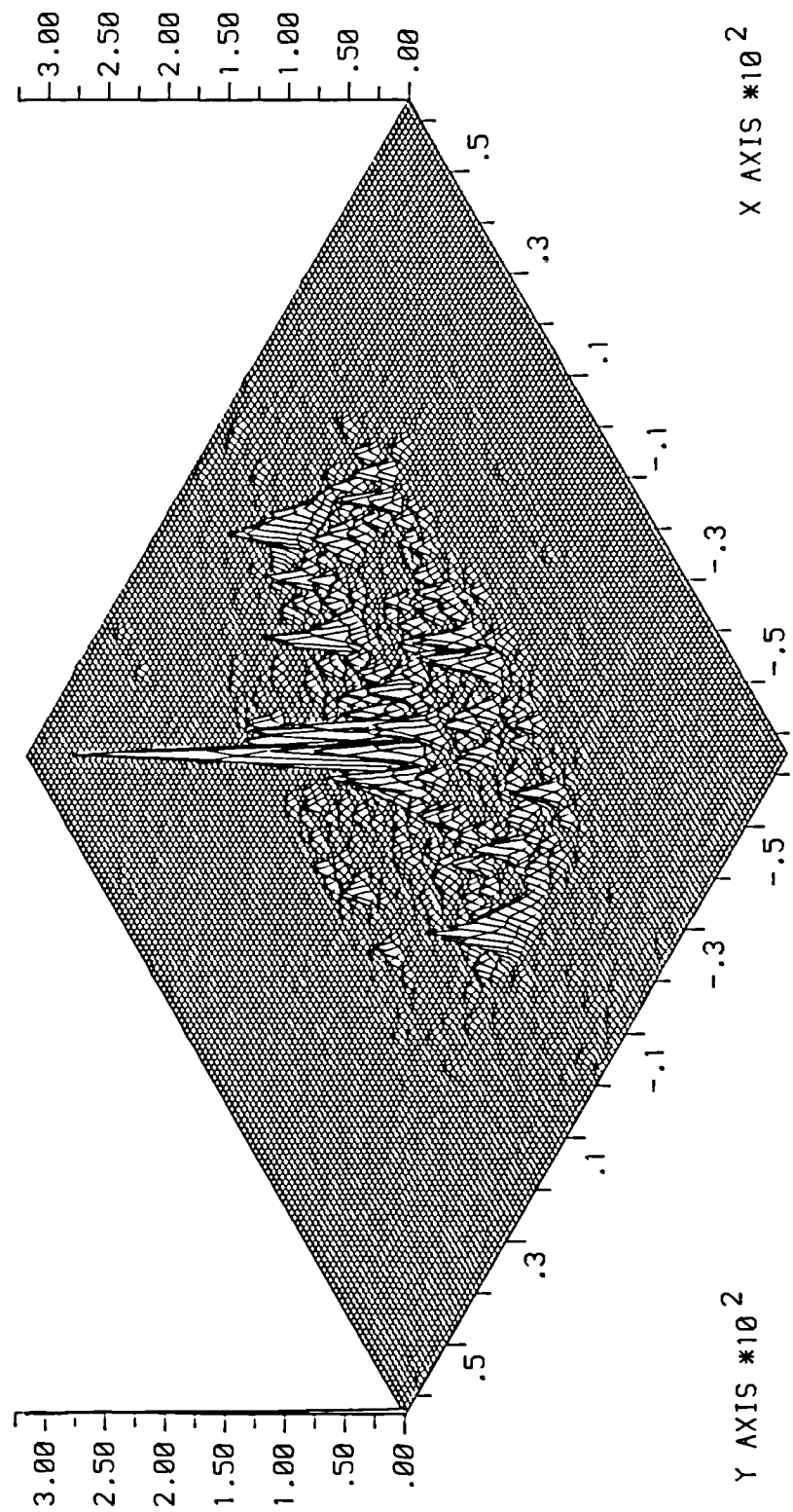


Figure 3.23(c) Isometric of inter-class cross-correlation of E180 and M180

the cross-section shown in Figure 3.22(b)). It has been noted previously that the ACF in the photorefractive joint transform correlator configuration is invariant to the input image distribution, being simply the Fourier transform of the frequency plane modulating functions. In contrast, the cross-correlation results just described indicate that the peak width is modified by the dimensions of the input image, the FWHM of the cross-correlation peaks shown in Figures 3.22 and 3.24 being more extended in the x direction than in the y direction due to the greater length of the re-oriented E191 view of the Escort with which the E180 view is correlated. It is also interesting to note that the spatial frequency content of the car images is higher in the y direction than the x due to the large amount of fine horizontal structures present in the images (for instance, edges of the rear spoiler and roof); this of course leads to an increased magnitude in the higher spatial frequencies in this direction as can be readily seen from the photographs of the power spectra of the car images shown in Figure 4.5 in the next Chapter. The increased extent of the frequency domain field distribution will in turn lead to a more localised correlation peak in the y direction as observed. During the frequency removal process these spatial frequency components may be expected to be preserved since the horizontal lines and their relative spacings are important discriminating features between the two cars but remain constant as the Escort is rotated about an axis perpendicular to the viewing direction. Thus they will contribute to the recognition of this orientation when it is correlated against the rear view. The asymmetry in the cross-correlation FWHM between the x and y directions is maintained in the cross-correlation response of the modified filter to the intra-class view.

Figure 3.25 shows the inter-class cross-correlation response of the modified filter. The centre point of the cross-correlation has been reduced to a near zero value, the highest sidelobes being remnants of the relatively broad base of the original cross-correlation peak; the cross-sections are taken through the maximum ($x = 134$, $y = 125$) to indicate the worst case. This has, however, been reduced to 0.12 of the intra-class cross-correlation, the cross-sections and isometric having been normalised to this value. It is interesting to note the general reduction in the correlation plane clutter that has occurred for the inter-class response of the filter. In comparison of the isometrics shown in Figures 3.23 and 3.25 account must be taken of the different intra-class maxima to which they are normalised since this leads to an apparent reduction of the clutter by 2.18 times. However, this is not enough to account for the large reductions in the off-axis peaks that have occurred. Thus, although the procedure is explicitly only forcing the on-axis correlation plane data point to a small value, the removal of frequency terms with maximum phase disparity between the intra-class and inter-class images appears in this particular case to have also had a beneficial effect in suppressing spurious off-axis cross-

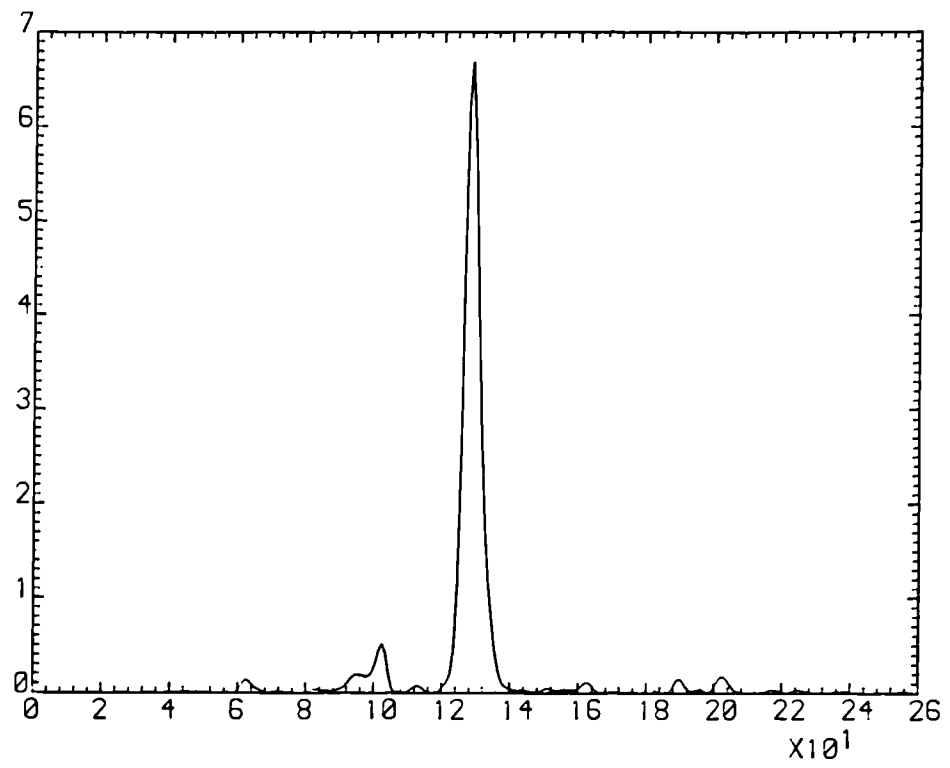


Figure 3.24(a) Cross-section at $y = 0$ of intra-class cross-correlation of E180 and E191 after discrete frequency removal

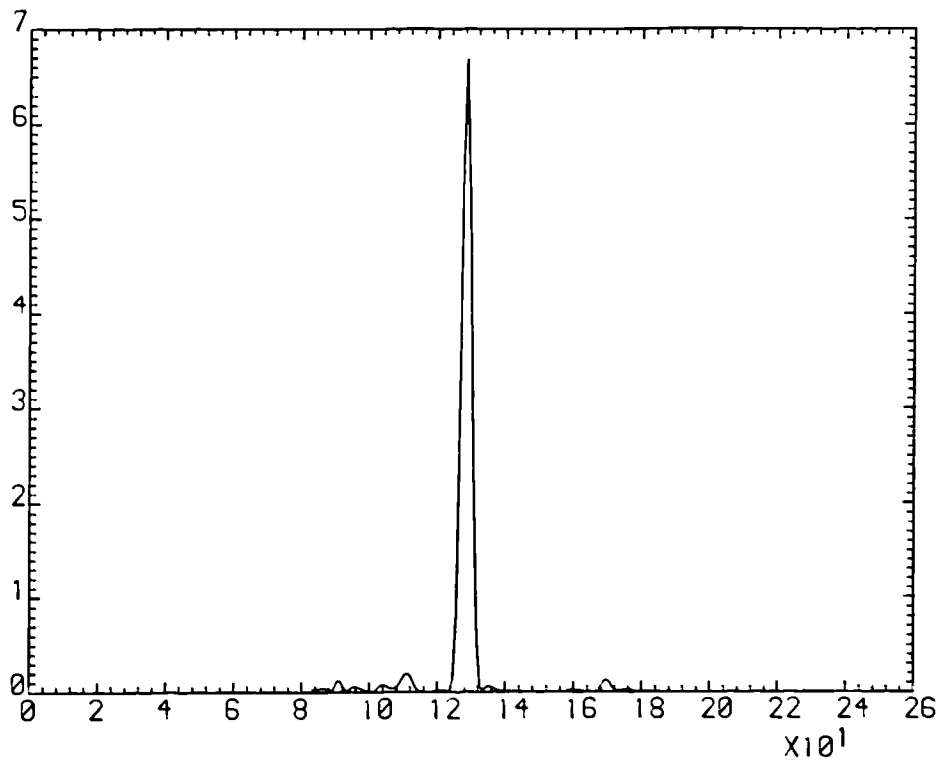


Figure 3.24(b) Cross-section at $x = 0$ of intra-class cross-correlation of E180 and E191 after discrete frequency removal

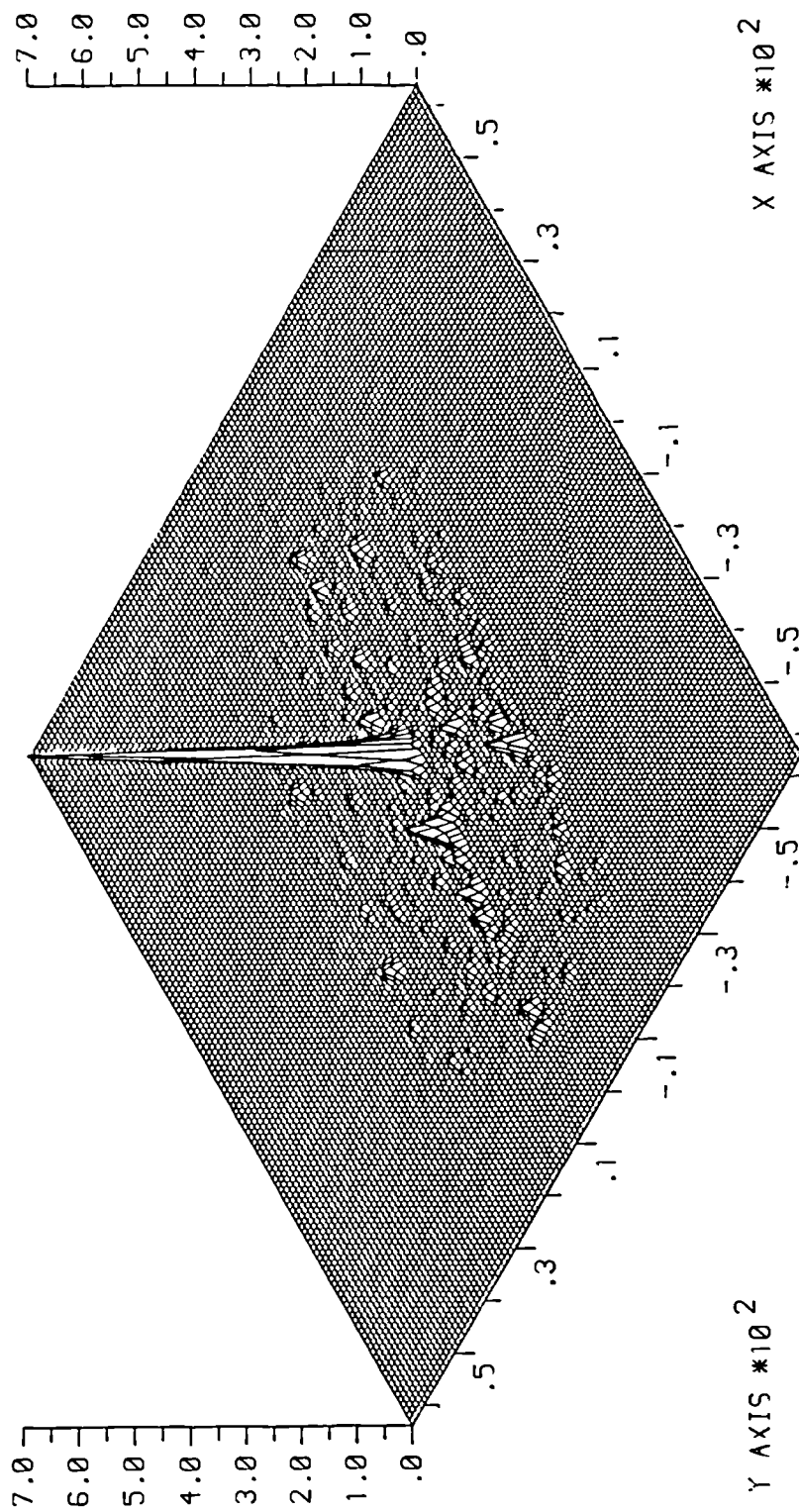


Figure 3.24(c) Isometric of intra-class cross-correlation of E180 and E191 after discrete frequency removal

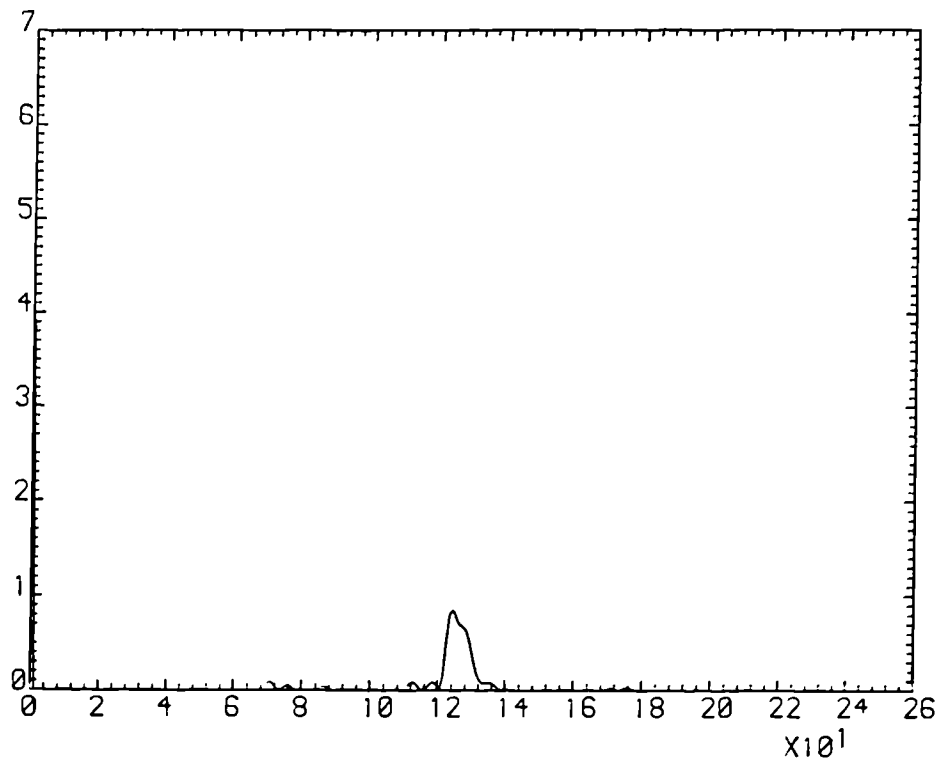


Figure 3.25(a) Cross-section at $y = 0$ of inter-class cross-correlation of E180 and M180 after discrete frequency removal

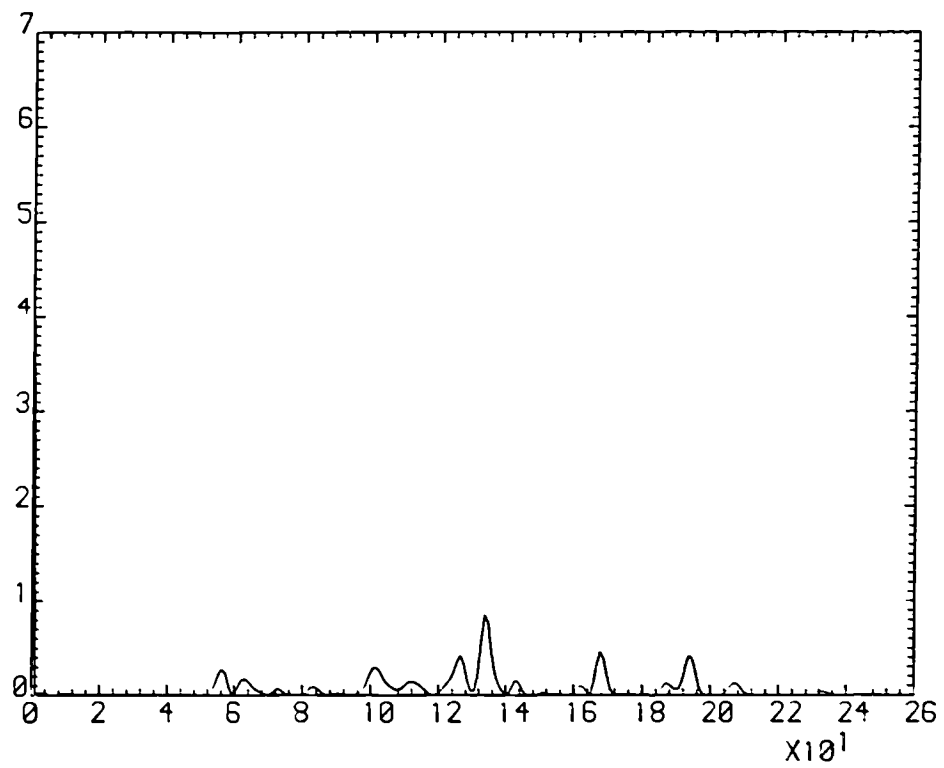


Figure 3.25(b) Cross-section at $x = 0$ of inter-class cross-correlation of E180 and M180 after discrete frequency removal

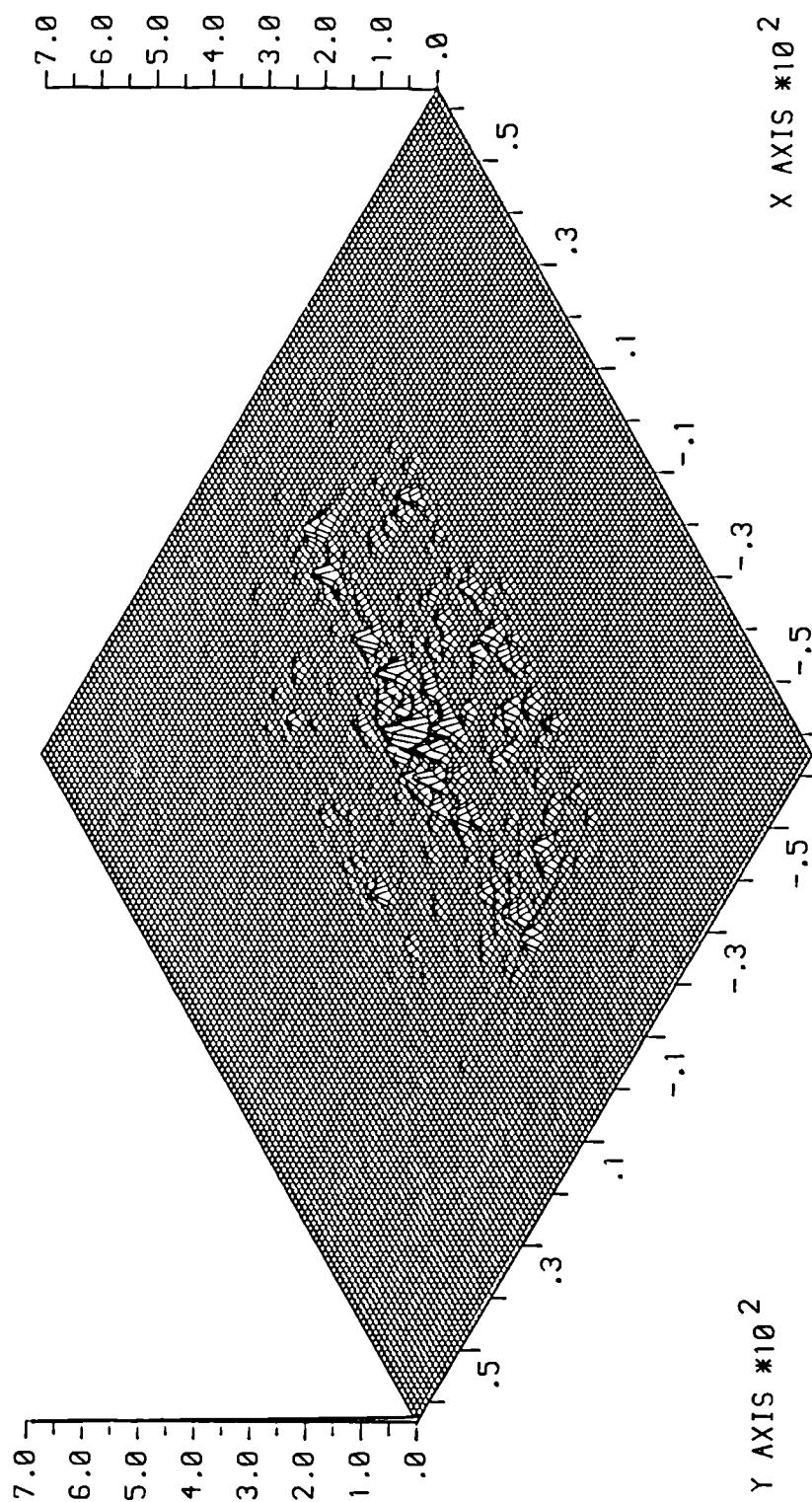


Figure 3.25(c) Isometric of inter-class cross-correlation of E180 and M180 after discrete frequency removal

correlations between the reference and inter-class input image. The simulation runs performed here are not extensive enough to determine whether this is a general property of the filter, further simulations with inter-class images that generate multiple peaked, cluttered cross-correlation responses being required to determine this.

As mentioned earlier, the second example is not as demanding as the first in terms of maximizing the intra-class versus inter-class discrimination since the intrinsic difference between the two images was greater in this case, as indicated by the initial correlation ratios. This is reflected in how readily the inter-class cross-correlation may be forced to zero while maintaining or increasing the intra-class response. Again the evolution of the relative correlation peak heights as the frequency removal process proceeds is shown in Figure 3.26. In contrast to the first example it can be seen that the intra-class response can be rapidly forced completely to zero without any exponential tail-off. This only requires the removal of 1730 components as compared to the removal of 4500 components for the M000 cross-correlation (which still left a sum of 50 although, of course, the starting sum was less than half in the present case - 425 as compared to 907). A major difference with the first example is that in this case the intra-class response does not reach its maximum until 1200 components have been removed and remains approximately constant after this, the final value for the sum of real field components being 660 as compared to a maximum of 675.

This behaviour appears to be the exception rather than the rule, however. Figure 3.27 shows the response obtained when an Escort sideview is correlated with an $11\frac{1}{4}^\circ$ rotated view and the sideview of the Mazda. Since in this orientation the image appearance changes less rapidly with rotation angle, the expected initial higher intra-class cross-correlation is obtained (although comparison with Figure 3.21 reveals it is only of the same magnitude as the unmodified cross-correlation of E000 and M000). Although the initial inter-class response is somewhat lower than this, being 0.42 of the intra-class peak intensity, it cannot be forced to zero as rapidly as for the previous example involving M180 because of the exponential tail-off. In addition, the intra-class response peaks very quickly after the removal of only 600 frequency components. However, it then only decreases gradually with further removal of components, reaching a final value of 908 at the termination of the frequency removal, as compared to a maximum of 1002. The response has in this example, therefore, been maintained at an approximately constant value while the inter-class real field component sum has been decreased to 0.034 of its starting value. Figures 3.28 and 3.29 demonstrate the effective suppression of the inter-class cross-correlation response that can be achieved in this case.

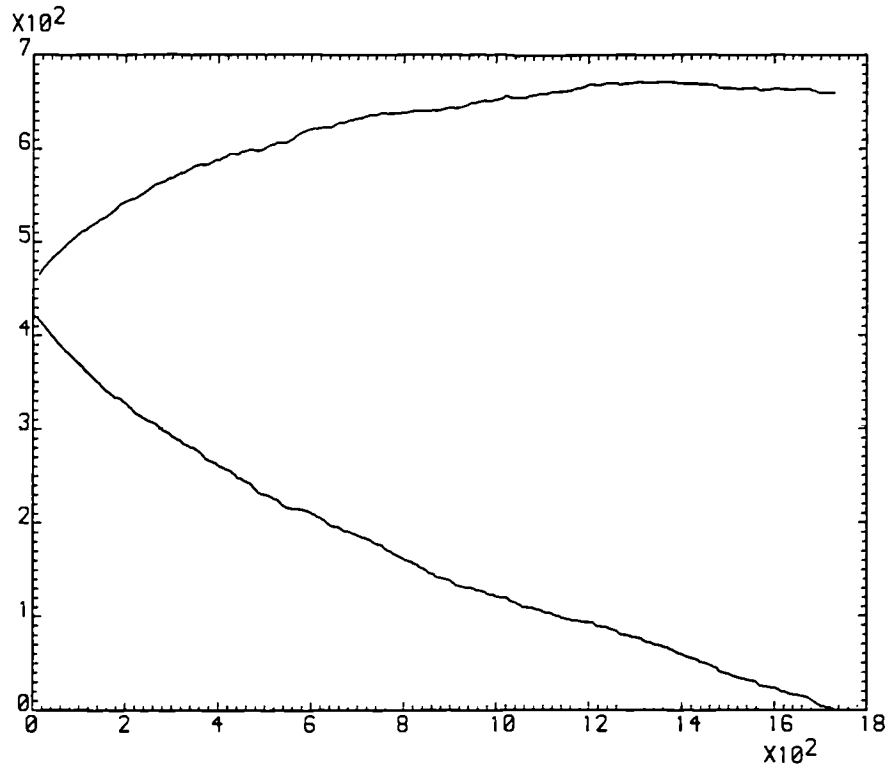


Figure 3.26 Plot of the relative intra-class and inter-class cross-correlation magnitudes of E180 with E191 and M180 as discrete frequency components are removed

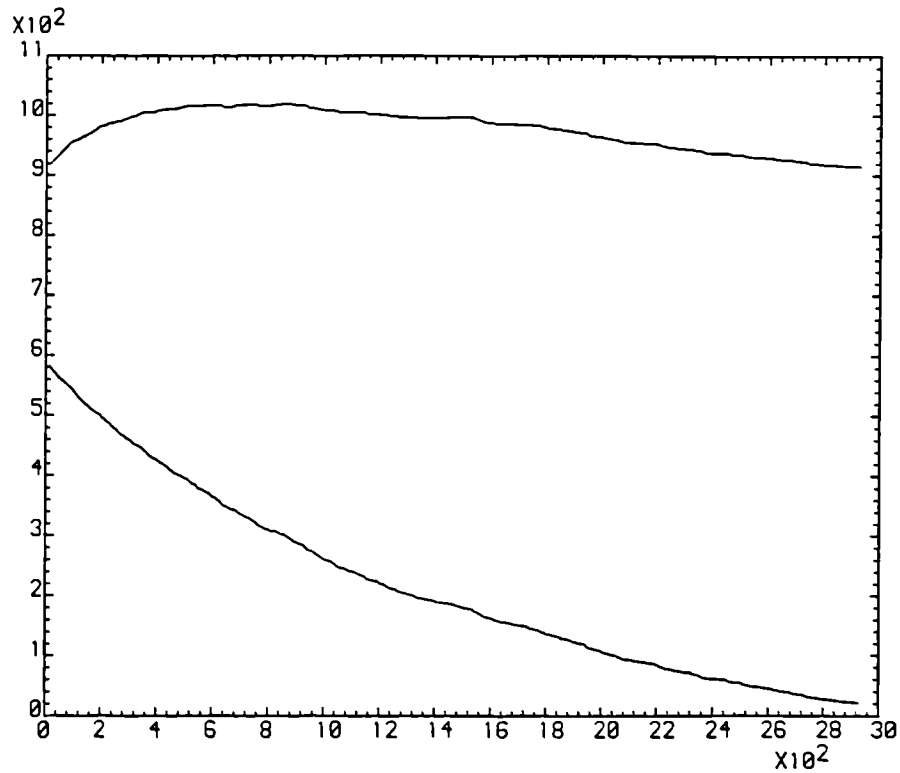


Figure 3.27 Plot of the relative intra-class and inter-class cross-correlation magnitudes of E270 with E281 and M270 as discrete frequency components are removed

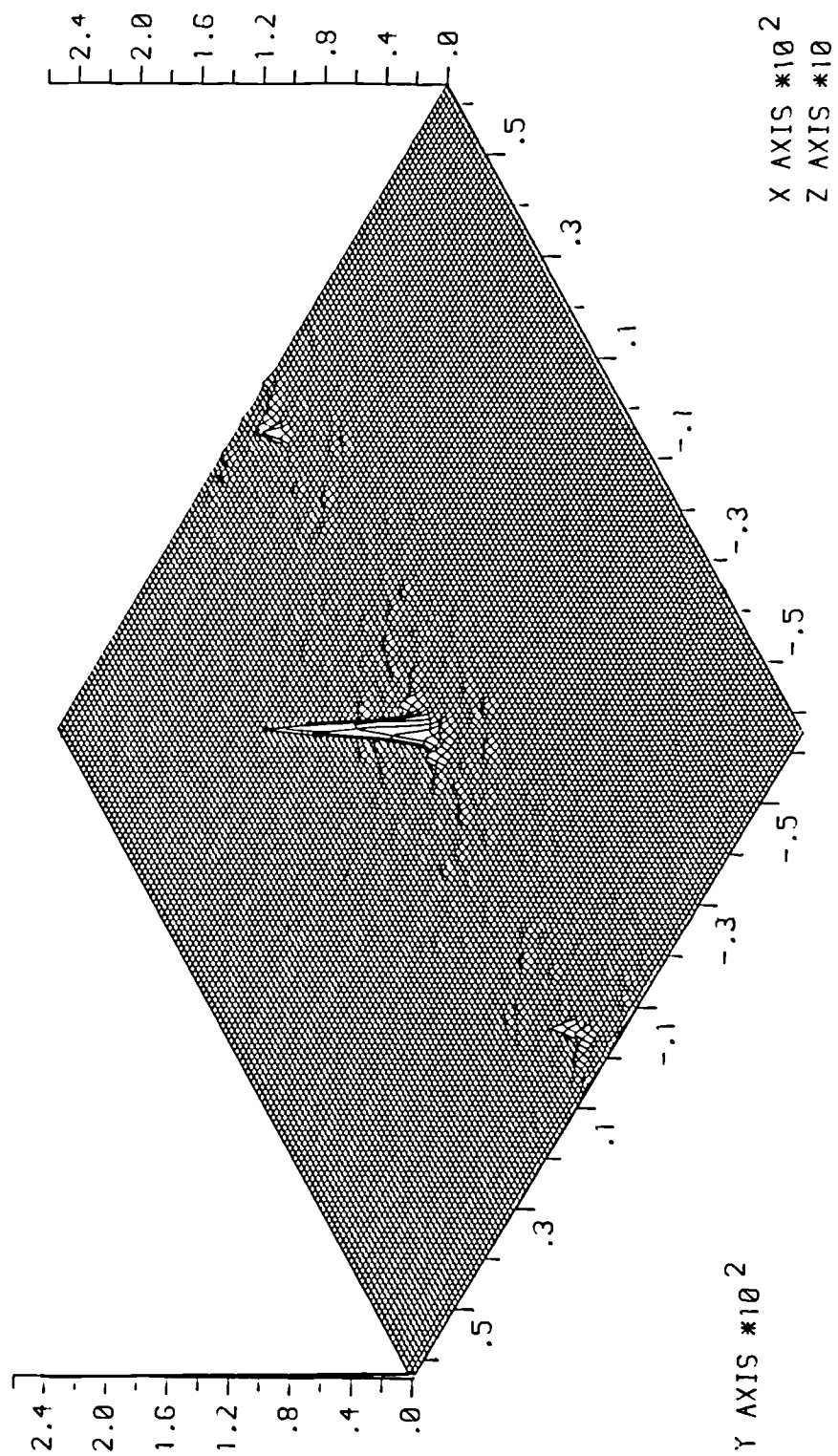


Figure 3.28 Isometric of inter-class cross-correlation of E270 and M270

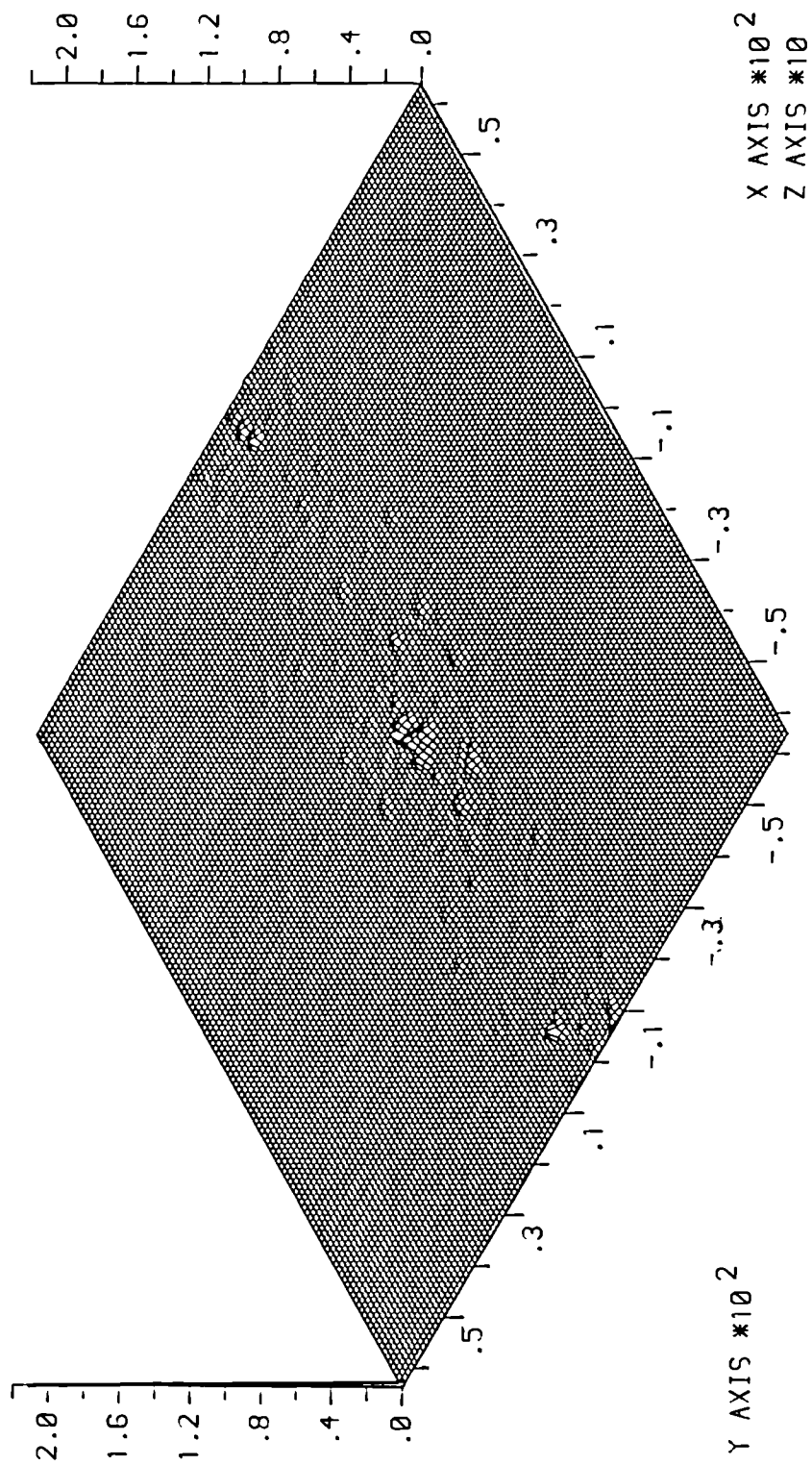


Figure 3.29 Isometric of inter-class cross-correlation of E270 and M270 after discrete frequency removal

Further insight into the way in which the discrete frequency removal technique operates may be gained from examination of the frequency plane field values directly after the filter plane (prior to Fourier transformation to the correlation plane) for the test intra-class and inter-class examples discussed above.

Figures 3.30(a) and (b) show the x and y radial cross-sections in this plane of the real components of the field resulting from the filter used to maximize the discrimination of E000 and M000 while maintaining correlation with the E349 intra-class view. The field distributions shown in Figure 3.30 result from view E000 being used as an input ie. they correspond to the auto-correlation function for this filter. (The field values are thus entirely real and the amplitude cross-sections shown in Figure 3.30 correspond to the modulation transfer function (MTF) of the filter.) The unmodified filter would, of course, generate a Gaussian field distribution under these circumstances which would Fourier transform to a localised Gaussian ACF as shown in Figure 3.30(b). Since the peak height of the correlation function is proportional to the integral of the real field component of the field behind the filter plane, the ACF associated with the field distribution shown in Figure 3.30 will be reduced upon the removal of frequency components, as has been noted previously; the peak height is 0.245 that of the unmodified ACF. However, as shown in Figures 3.20(a)-(c) the localised Gaussian profile is maintained with very minimal sidelobe disruption despite the sharp discontinuities introduced into the frequency plane envelope by the discrete frequency removal process.

Figures 3.30(a) and (b) show that the low frequency portions of the spectrum have been entirely removed. In these regions of the spectrum the real field components of the intra-class and inter-classes images will be of similar magnitude. Removal of these frequency components will, therefore, not greatly increase the difference of the cross-correlation peak heights between these two images. However, their removal will suppress the inter-class cross-correlation since the low frequency inter-class components will have large positive values contributing to the cross-correlation. Their removal will also result in a decrease in the intra-class cross-correlation since they will be large and positive in this spectrum as well. As the order in which the frequency components are removed is dependent on their differences, ie. large negative intra-class components coincident with large positive inter-class components being the first to be removed, the low frequency components will be the last to be removed. Thus in Figure 3.21 only a slight difference of the two curves (corresponding to cross-correlation peak heights) occurs after the removal of 3000 components, the next 1500 being removed from the low frequency regions of the spectra to further suppress the inter-class cross-correlation. Because this

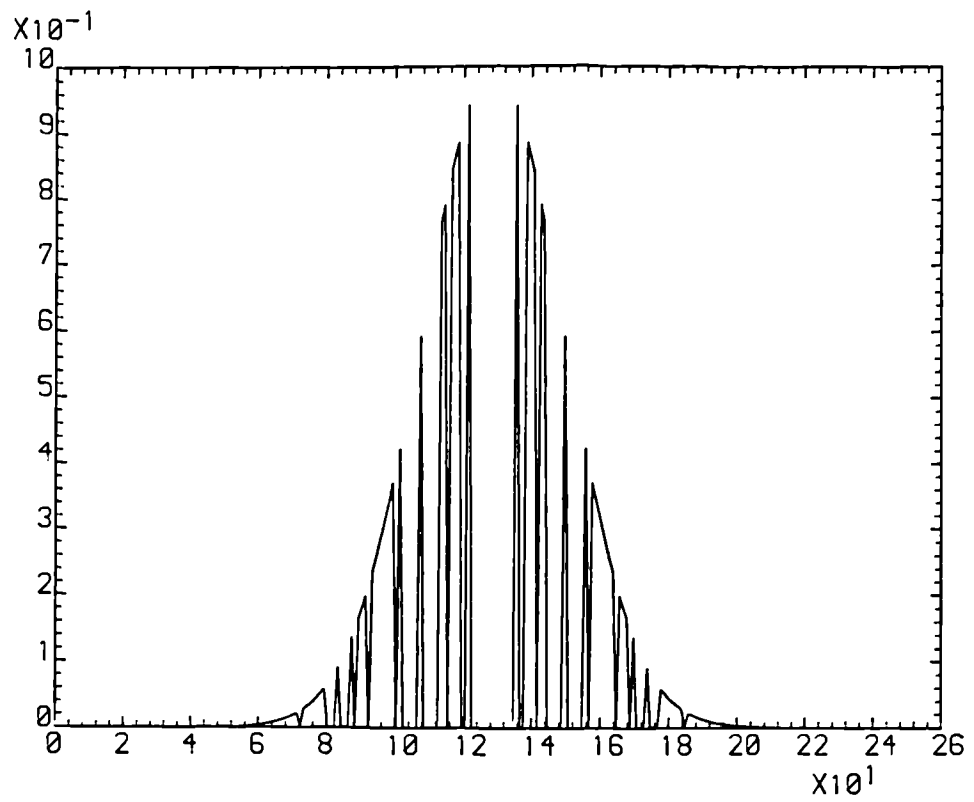


Figure 3.30(a) Cross-section at $y = 0$ of field components directly behind modified E000/E349/M000 filter plane with E000 as input

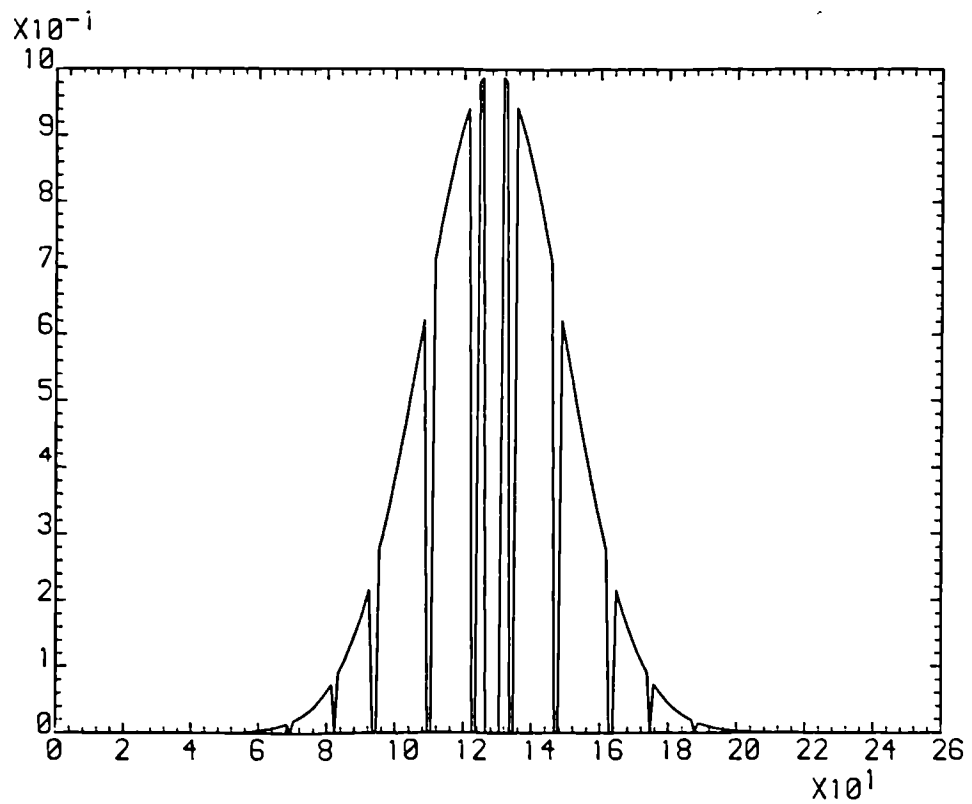


Figure 3.30(b) Cross-section at $x = 0$ of field components directly behind modified E000/E349/M000 filter plane with E000 as input

necessarily results in the removal of large positive values from the intra-class spectrum as well, this correlation will unavoidably fall during this stage of the process. It can be seen that a bandpass filter is automatically generated by the procedure, the width of the low frequency stop increasing in order to maximize the intra-class to inter-class cross correlation ratio. However, some time before this point is reached the maximum difference between the cross-correlations will have been achieved, this requiring only a small low frequency portion of the spectra to be removed.

The frequency components which are removed to enhance the intra/inter-class discrimination occur at higher frequencies and correspond to those which have positive real field-values in the inter-class spectrum but negative values in the intra-class spectrum. Figure 3.30 shows localised regions throughout the mid- and high-frequency regions of the filter spectrum which have been set to zero and correspond to the locations at which the intra-class and inter-class real field values are most different. A feature of the filter is that the zeroed regions are more closely spaced in the x-direction cross-section than in the y-direction cross-section. The reason for this is that, as noted earlier, there is a predominance of detail in the car images of horizontal features, essentially comprising elongated rectangular regions of different spacings. These give rise to a wider distribution of frequencies in the y-direction of the spectrum as clearly shown in the photographs of the power spectra of the car images shown in Figure 4.5. These rectangular features will generate oscillatory spectral components, the period being inversely proportional to the rectangular width. The spectrum of the entire object will, of course, consist of the complex addition of many such component spectral features. Nevertheless, the resulting spectrum has a predominantly oscillatory form along the x and y axes, particularly for the front and rear views of both the Escort and Mazda (see Figure 4.5). Thus cross-sections through the real field components of the spectra of the car images in these directions will pass through many maxima and minima, although the oscillations may be expected to be of a different period and phase in each spectrum.

When the frequency removal process is applied to the spectra, those frequencies will be removed which have real field components that are most different. The purpose of the procedure is to maximize the sum of the real field components of the spectrum of the intra-class image, E349, while minimizing those of the inter-class image M000. Thus, at a particular frequency location, if the E349 spectrum has a large negative real field component and the M000 spectrum has a large positive component, the frequency will be removed as this will lead to a divergence of the intra-class and inter-class cross-correlation values as required. Since the real field components of each spectrum are passing through positive and negative values cyclically, the frequency removal will occur

in those regions where the negative portions of the E349 spectral oscillation coincide with the positive portions of those of the M000 spectrum. These oscillations have a much higher frequency in the x-direction than in the y-direction and so the interval between the zeroed spectral regions may be expected to be considerably less in the x-direction. Hence, within the limiting Gaussian bandpass of the filter, the number of components removed will be greater in this direction. This difference can clearly be seen in Figure 3.30 and also in Figure 3.31 which shows the real field values directly behind the filter designed to discriminate E180 from M180 while maintaining correlation with E191. (Since the field values shown result from the Fourier transformation of E180, the amplitude distribution shown represents the MTF of the filter as in Figure 3.30.)

In Figure 3.31 it is also interesting to note that although significant portions of the low frequency regions have been removed in order to produce the filter, the zero order frequency and two other surrounding low frequency components still remain (comparison of the x and y cross-sections show that, in fact, a narrow sheet of low frequencies remain which is much more extended in the y-direction than in the x-direction). The reason low frequency components remain is that, as pointed out in the discussion of Figure 3.30, the intra- and inter-class low frequencies will have real field components of similar magnitude and so will be the last to be removed by the procedure, if it is necessary that it runs that far in order to suppress the inter-class cross-correlation. Examination of Figure 3.26 shows that the cross-correlation of E180 with M180 is readily forced to zero with little decrease in the intra-class cross-correlation peak height. Thus the procedure can be stopped while the intra-class frequency components have relatively low positive values which will not include the comparatively evenly matched low frequency terms. This discrimination task is easier than that shown in Figure 3.30 (which actually starts with a higher inter-class cross-correlation than an intra-class one) and so can terminate after the removal of less than 1800 terms as opposed to nearly 4500 in the more difficult case, which necessarily had to include large positive intra-class frequency components in order that the inter-class response was effectively suppressed.

The overall effect of the frequency removal procedure in subtracting the positive real field components of the inter-class field coincident with the negative real field components of the intra-class response can be explicitly seen by examining these fields before and after application of the filter. Figures 3.32(a)-(d) show x and y cross-sections through the real and imaginary field components of the distribution arising directly behind the unmodified filter matched to E000 as the reference image when M000 is input to the joint transform correlator simulation (the Fourier transform of these field values would yield the cross-correlation between these two car views). As expected, the real field

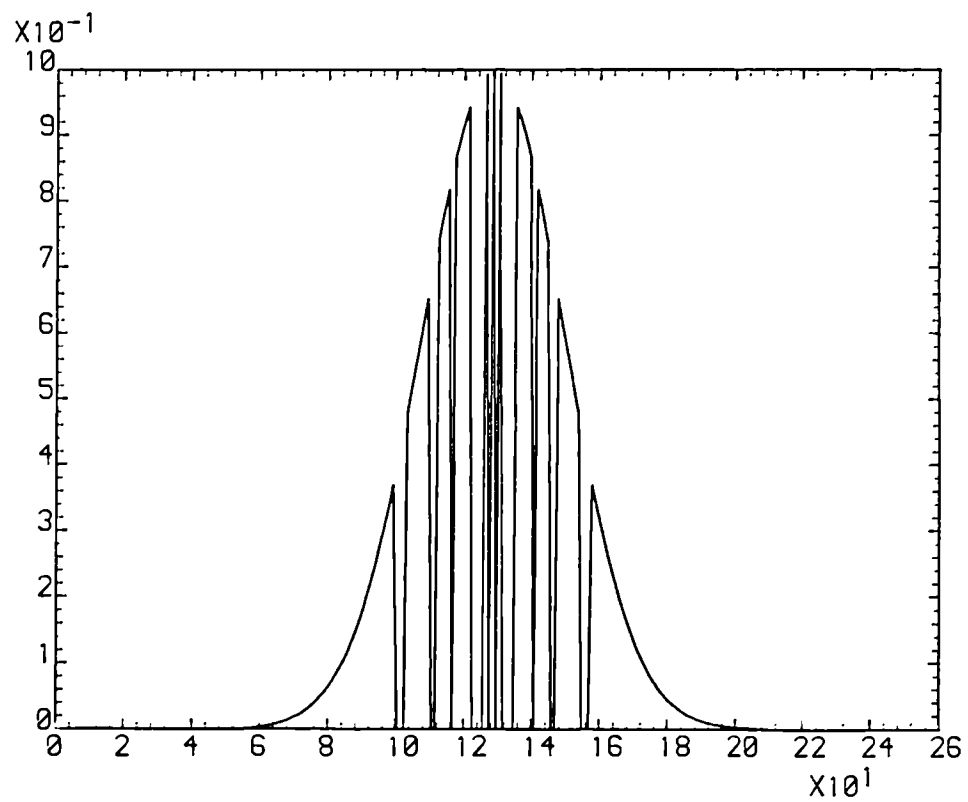


Figure 3.31(a) Cross-section at $y = 0$ of field components directly behind modified E180/E191/M180 filter plane with E180 as input

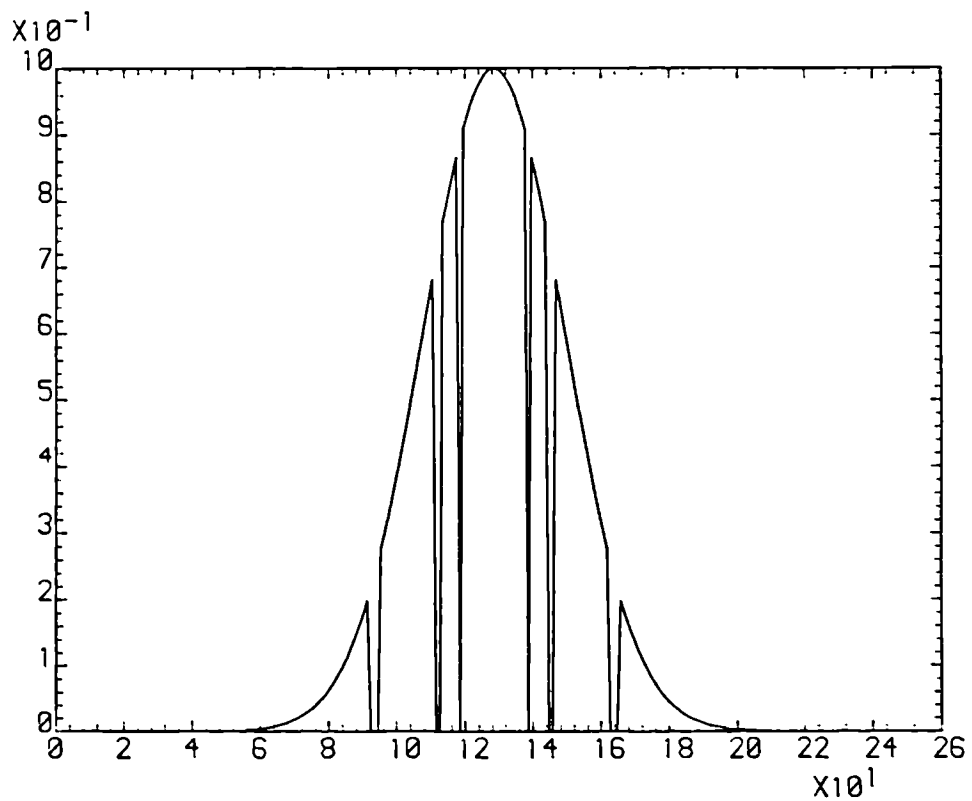


Figure 3.31(b) Cross-section at $x = 0$ of field components directly behind modified E180/E191/M180 filter plane with E180 as input

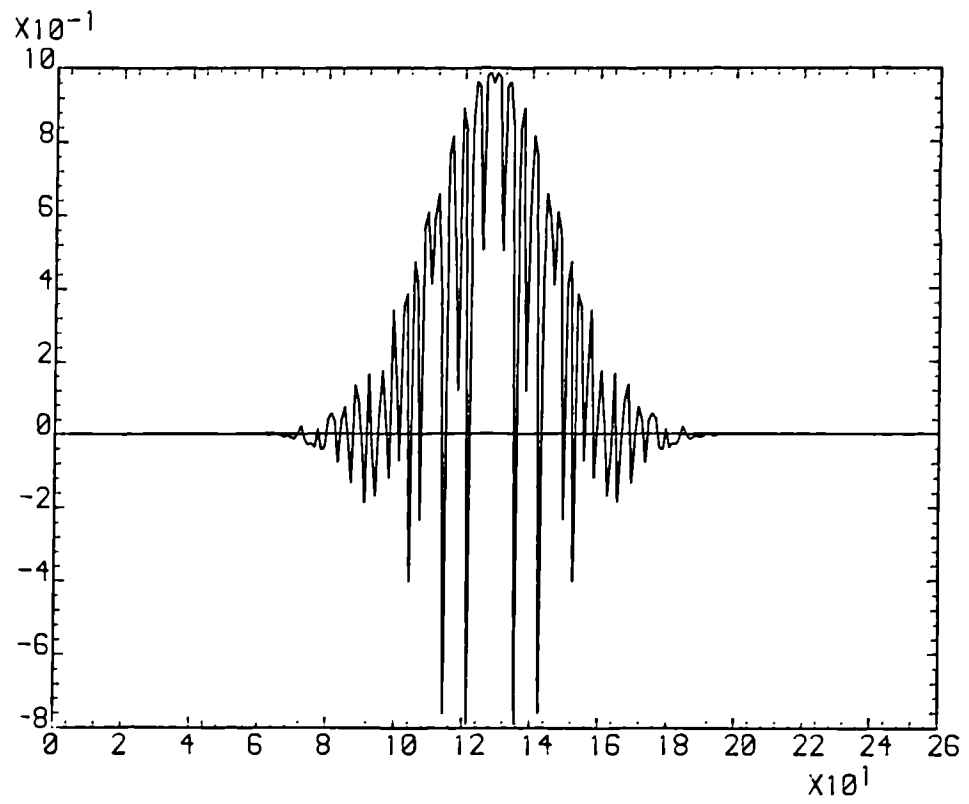


Figure 3.32(a) Cross-section at $y = 0$ of real field components directly behind unmodified E000 filter with M000 as input

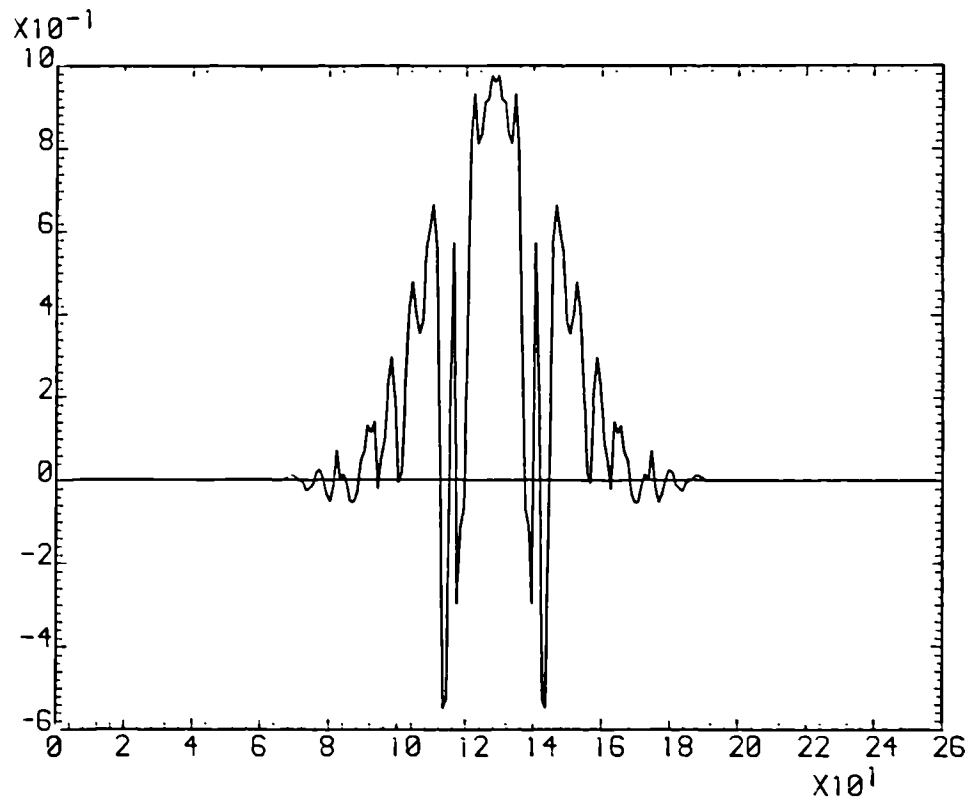


Figure 3.32(b) Cross-section at $x = 0$ of real field components directly behind unmodified E000 filter with M000 as input

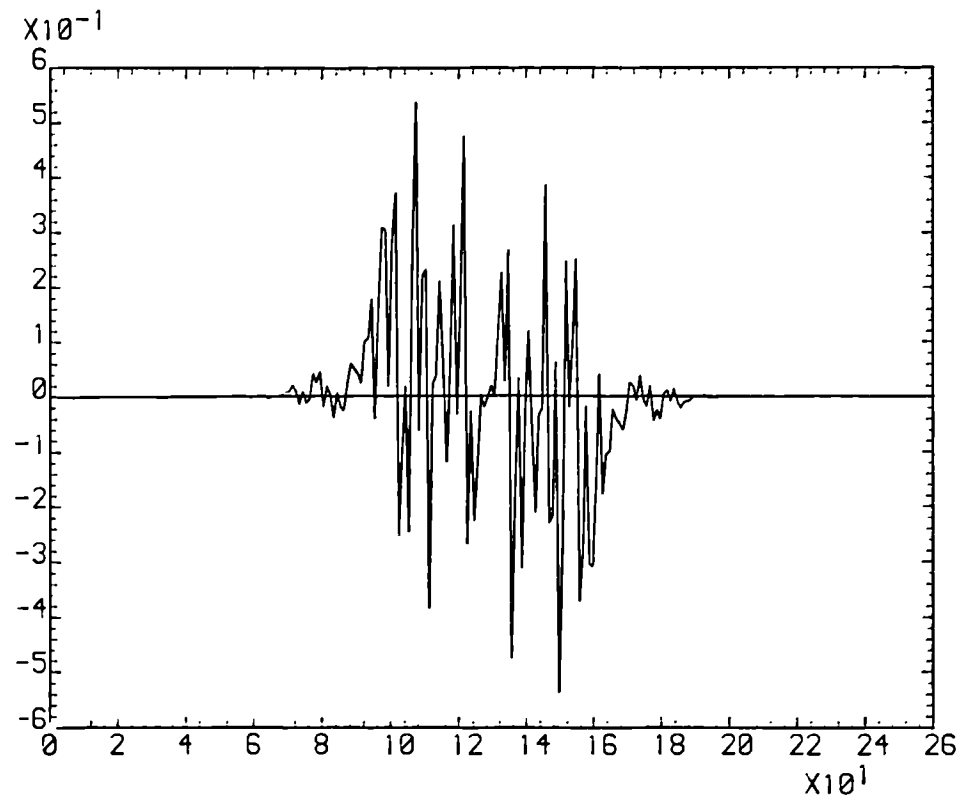


Figure 3.32(c) Cross-section at $y = 0$ of imaginary field components directly behind unmodified E000 filter with M000 as input

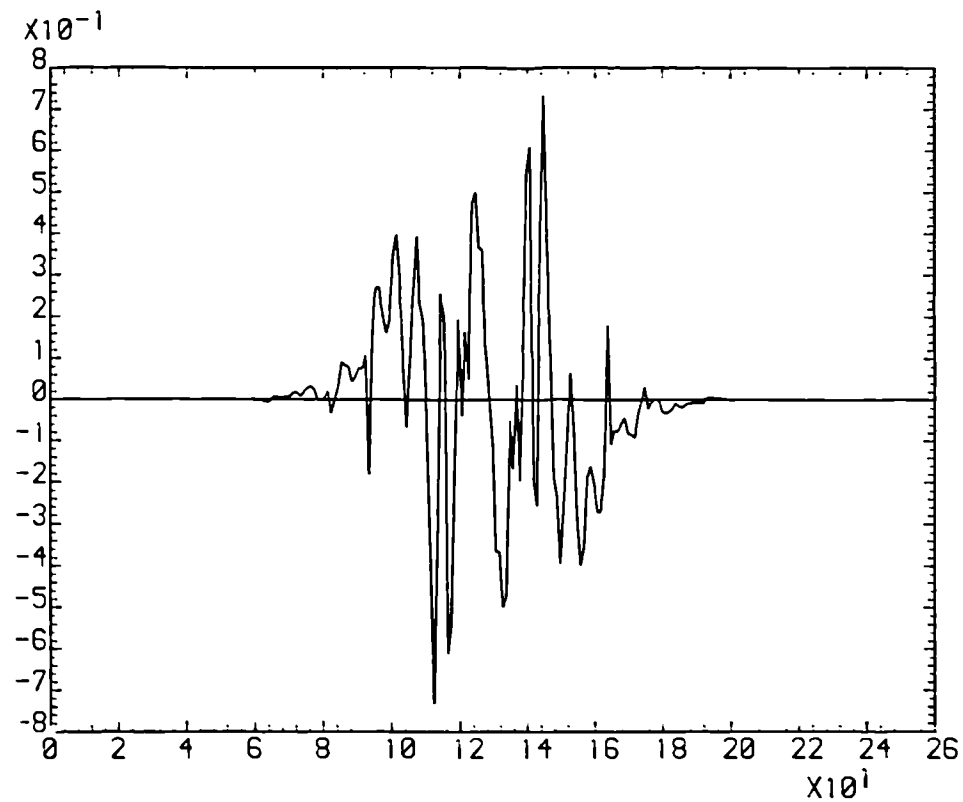


Figure 3.32(d) Cross-section at $x = 0$ of imaginary field components directly behind unmodified E000 filter with M000 as input

components are symmetric and the imaginary components anti-symmetric due to the Hermitian symmetry of the Fourier transforms of the reference and input images. The oscillations of the field values have a higher frequency in the x-direction than in the y-direction due to the structure of the car images, as discussed above. Notice that the real field components have predominately positive values leading to a high cross-correlation peak. (Importantly, M000 has been positioned at the correct location in the input plane such that the peak of its cross-correlation with E000 is centred in the correlation plane and is, therefore, proportional to the sum of the real field components of the distribution shown.) Figures 3.33(a) and (b) show the real field components of the field directly after the filter when the intra-class image, E349, is at the input. This was similarly centred such that the cross-correlation was on axis but, as mentioned above, the sum of the real field components was less than that for the inter-class input, M000.

Figures 3.34(a) and (b) show the real field components emerging from the filter after modification by the frequency removal process, cross-sections of the MTF of which is shown in Figures 3.30(a) and (b). Comparing Figures 3.32(a) and (b) with Figures 3.34(a) and (b) shows that large regions where the real field components are positive have been removed, but the negative components are largely unaffected. Careful examination of Figure 3.32(a) and Figure 3.34(a) reveals that, in fact, some of the smaller negative components of the E000/M000 cross-spectrum have been removed. Further comparison with Figure 3.33(a) indicates that this is because at these locations the E000/E340 cross-spectrum had large negative real field components. Thus removal of the frequency components at these locations results in a divergence of the intra-class and inter-class cross-correlations as required.

The cross-spectrum of E000 and E349, modified by the frequency removal process, is shown in cross-section in Figures 3.35(a) and (b). Comparison to Figure 3.33 shows that, although low frequency positive real field components have been necessarily removed to force the inter-class correlation to zero (resulting in an unavoidable decrease in the cross-correlation peak height), the negative frequency components have been largely eliminated. Those remaining in the x-direction cross-section coincide exactly with a larger negative field component in the E000/M000 cross-spectrum and had not been removed when the frequency removal process was halted.

For completeness, in Figures 3.36-3.39, the corresponding cross-spectral field values of the other example of the application of the frequency removal process that has been examined are shown, namely the inter-class discrimination of E180 and M180 while maintaining intra-class correlation with E191. Comparison of Figures 3.36 and 3.32

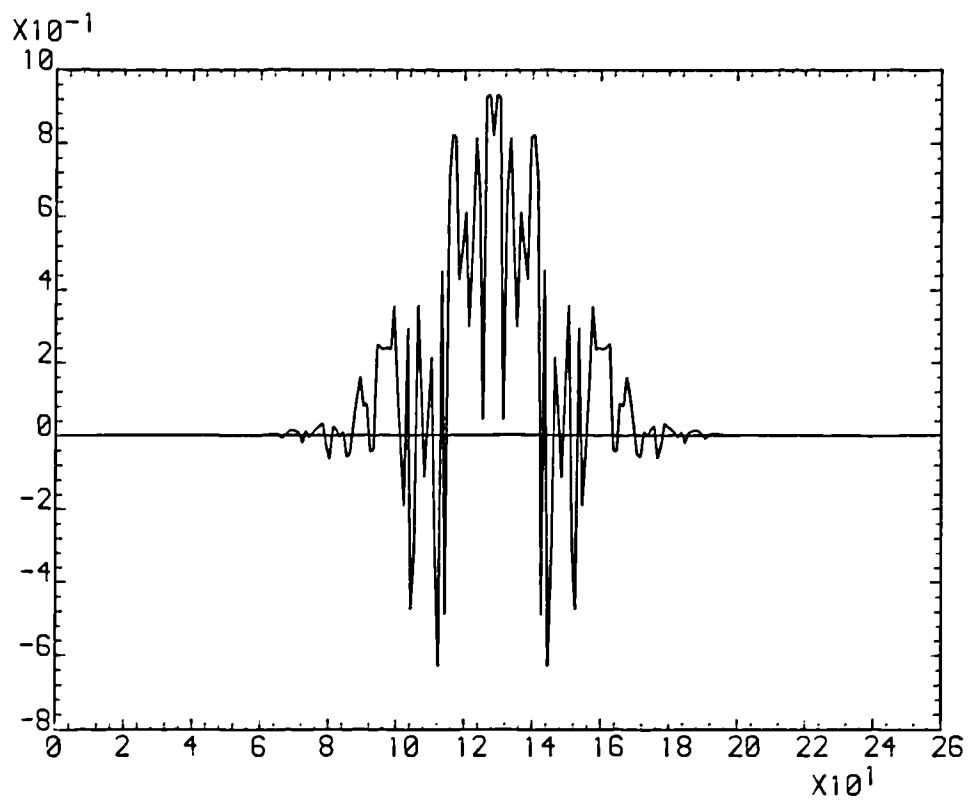


Figure 3.33(a) Cross-section at $y = 0$ of real field components directly behind unmodified E000 filter with E349 as input

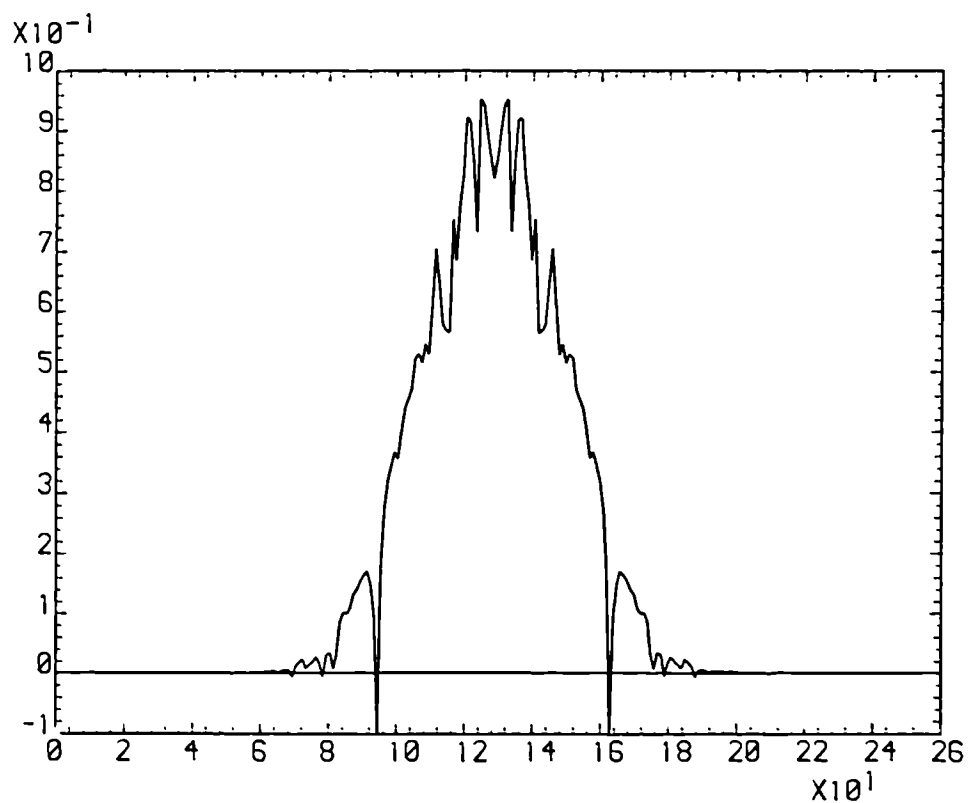


Figure 3.33(b) Cross-section at $x = 0$ of real field components directly behind unmodified E000 filter with E349 as input

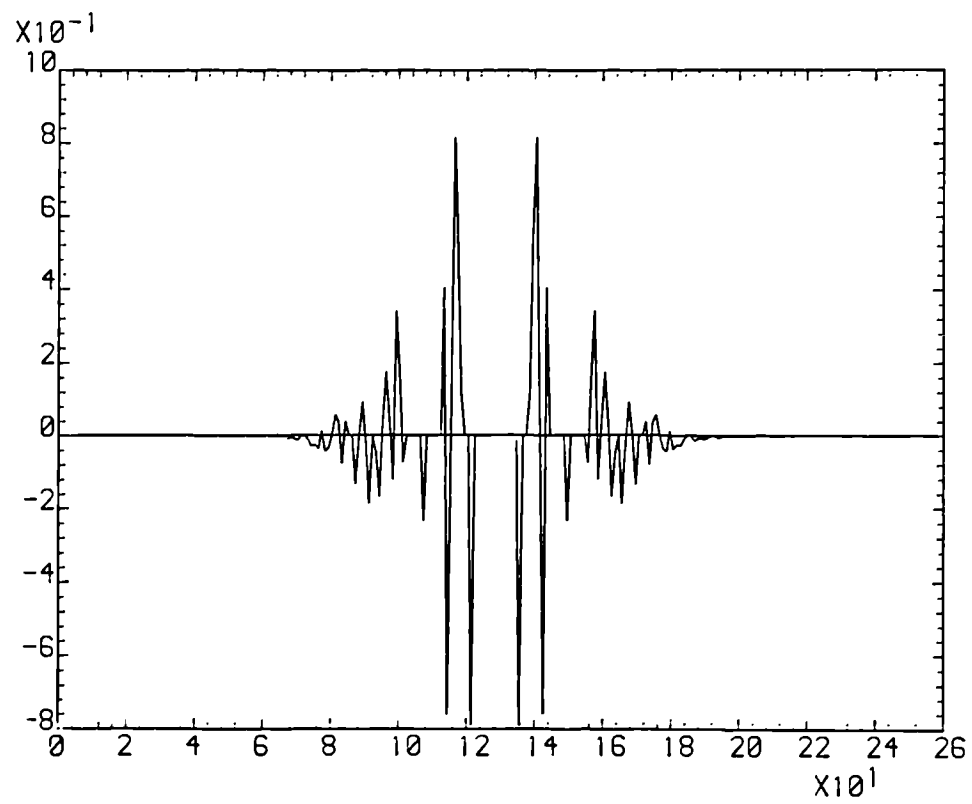


Figure 3.34(a) Cross-section at $y = 0$ of real field components directly behind modified E000/E349/M000 filter with M000 as input

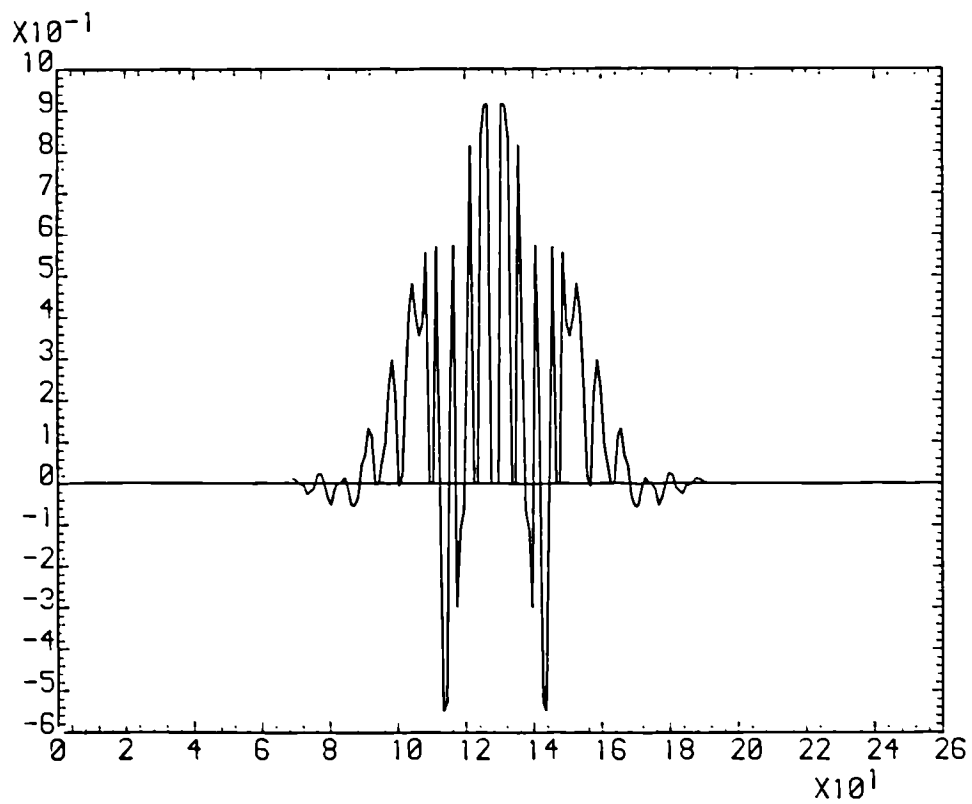


Figure 3.34(b) Cross-section at $x = 0$ of real field components directly behind modified E000/E349/M000 filter with M000 as input

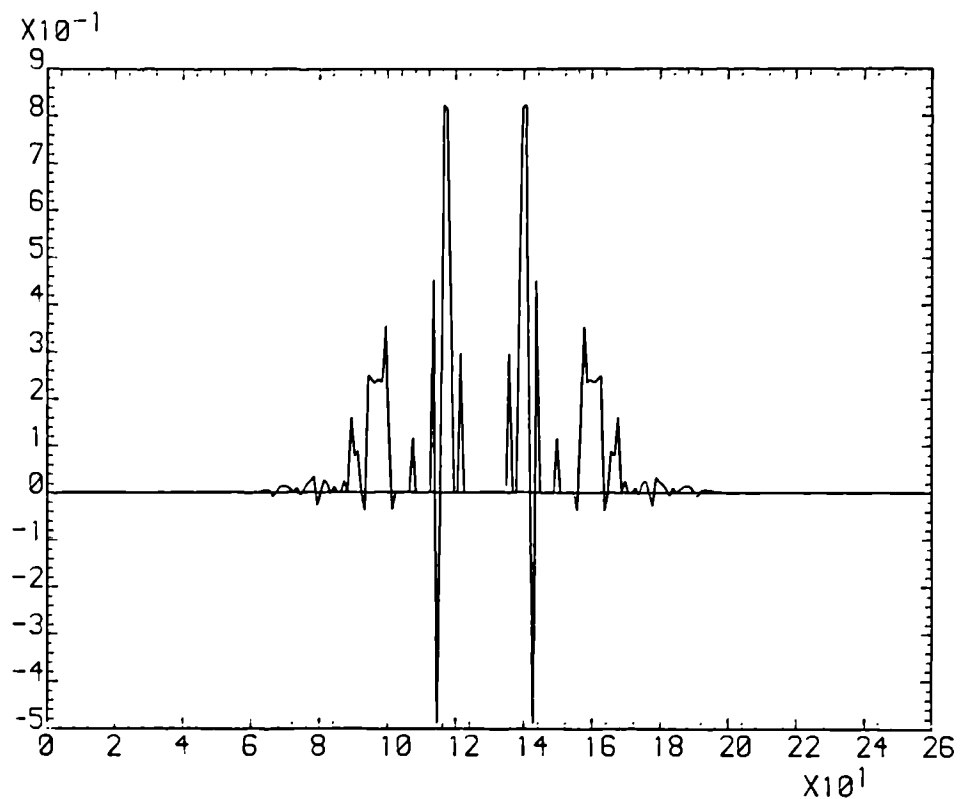


Figure 3.35(a) Cross-section at $y = 0$ of real field components directly behind modified E000/E349/M000 filter with E349 as input

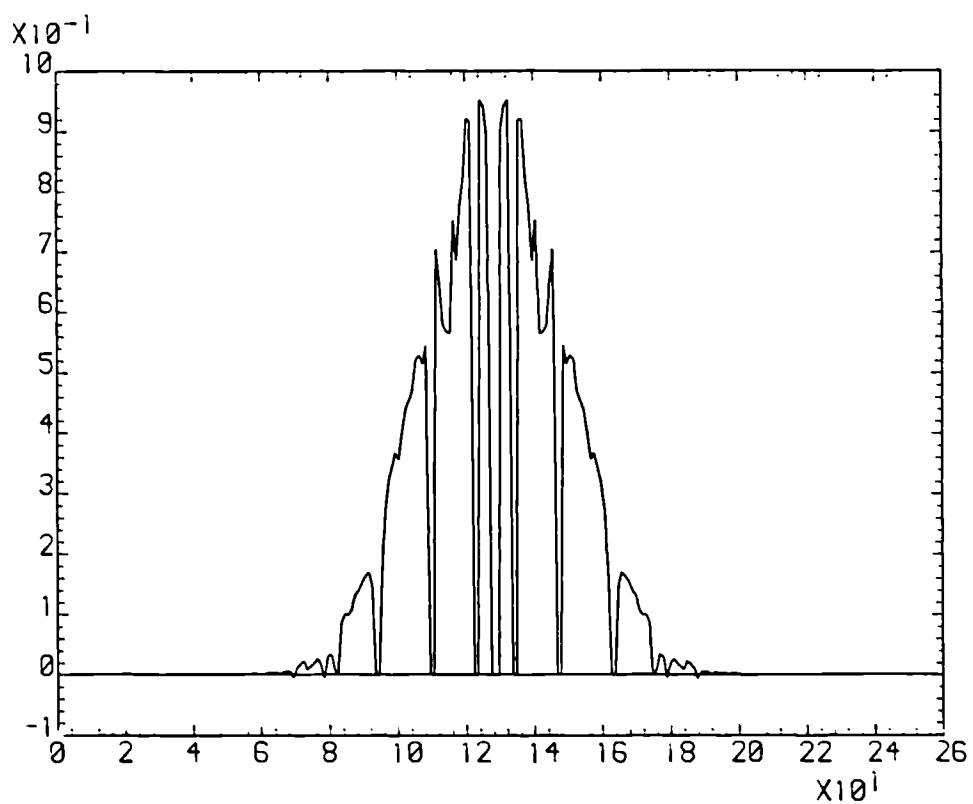


Figure 3.35(b) Cross-section at $x = 0$ of real field components directly behind modified E000/E349/M000 filter with E349 as input

indicates that in this case, in contrast to the previous example, the inter-class cross-spectrum has a greater number of negative real field components than the intra-class cross-spectrum. Indeed, the inter-class cross-correlation is initially lower than that of the intra-class and application of the frequency removal process increases this difference, readily forcing the inter-class response to zero as shown by Figure 3.26 (which was discussed in detail previously.) This is reflected in the appearance of the modified cross-spectral plots shown in Figures 3.38 and 3.39. In Figure 3.38 the inter-class cross-spectral real field values appear to have roughly zero integral in comparison with the unmodified cross-spectrum shown in Figure 3.36, demonstrating that positive regions of the spectrum have been removed to achieve this. In contrast, comparison of Figures 3.37 and 3.39 shows that the negative real field components have been largely removed selectively, (again except where they coincide with large negative field values present in the inter-class cross-spectrum) so increasing the intra-class cross-correlation. Thus, whilst the plots shown are only cross-sections through the cross-spectra and it must be borne in mind that the entire frequency plane contributes to the field sum, they nevertheless confirm and extend an intuitive understanding of the mode of action of the frequency removal procedure.

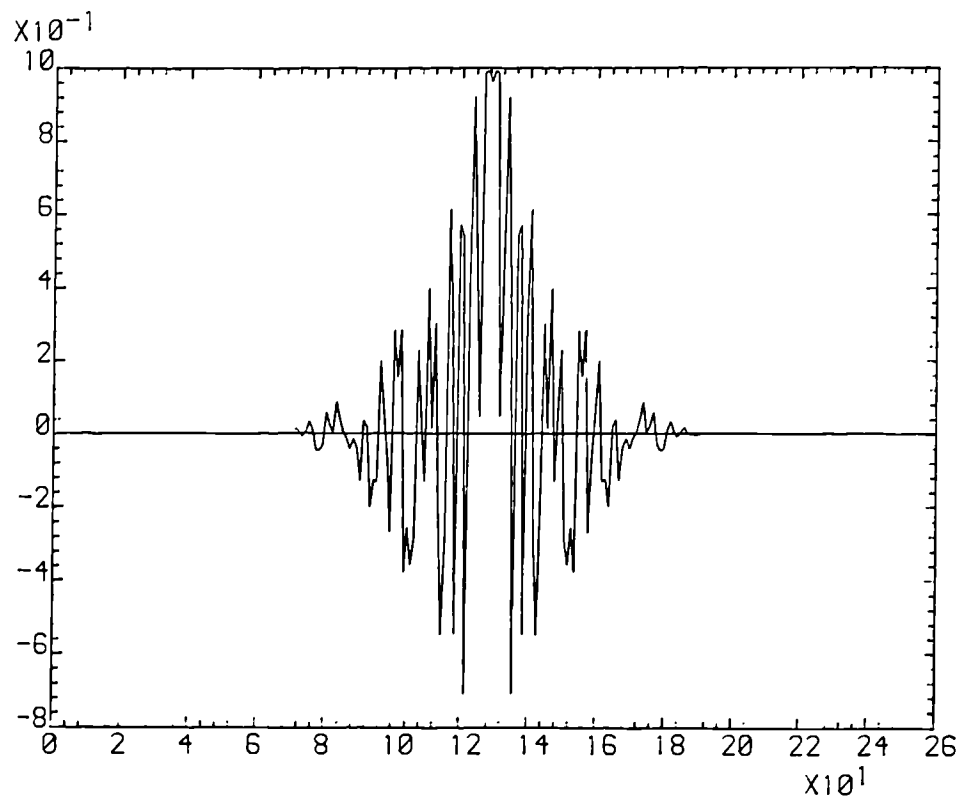


Figure 3.36(a) Cross-section at $y = 0$ of real field components directly behind unmodified E180 filter with M180 as input

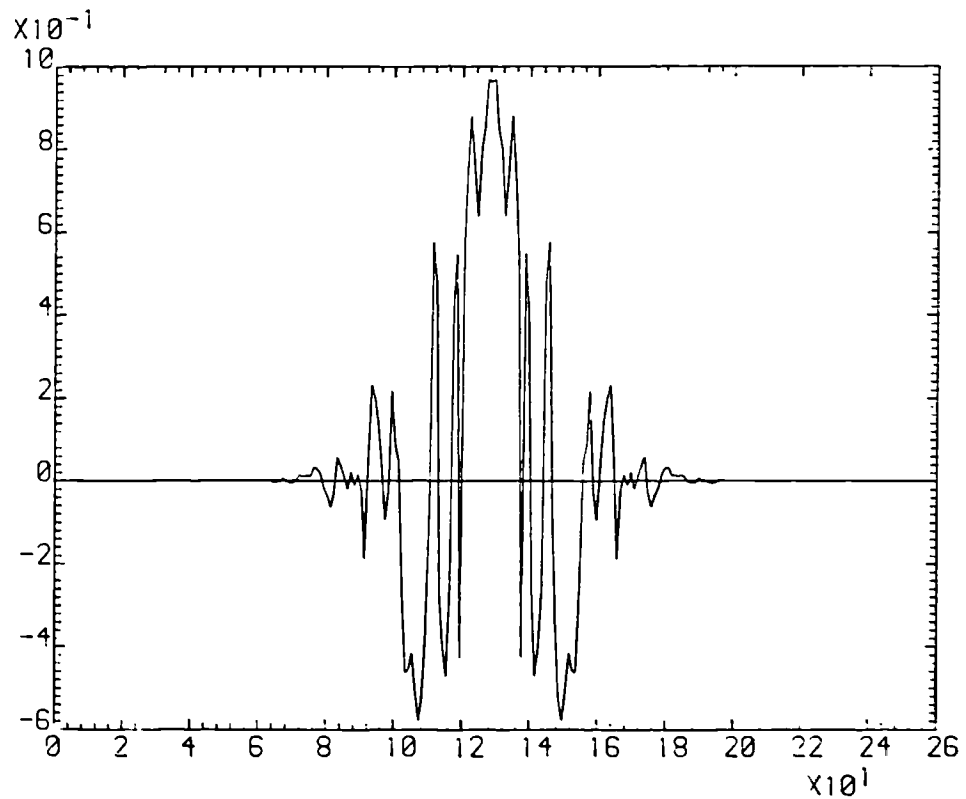


Figure 3.36(b) Cross-section at $x = 0$ of real field components directly behind unmodified E180 filter with M180 as input

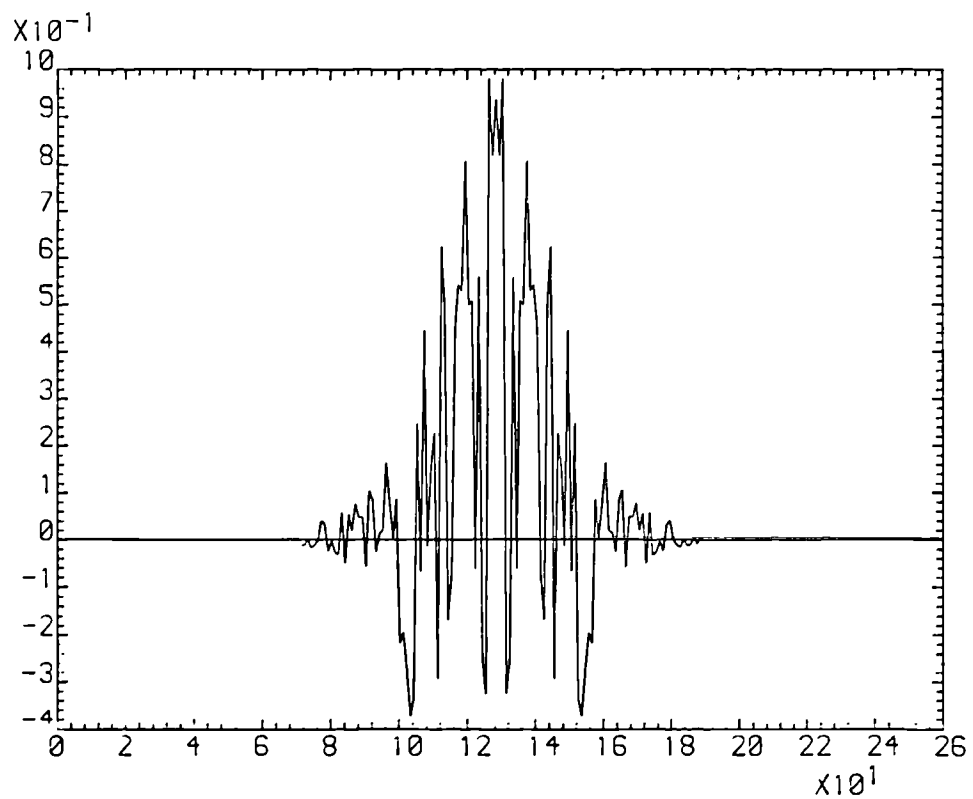


Figure 3.37(a) Cross-section at $y = 0$ of real field components directly behind unmodified E180 filter with E191 as input

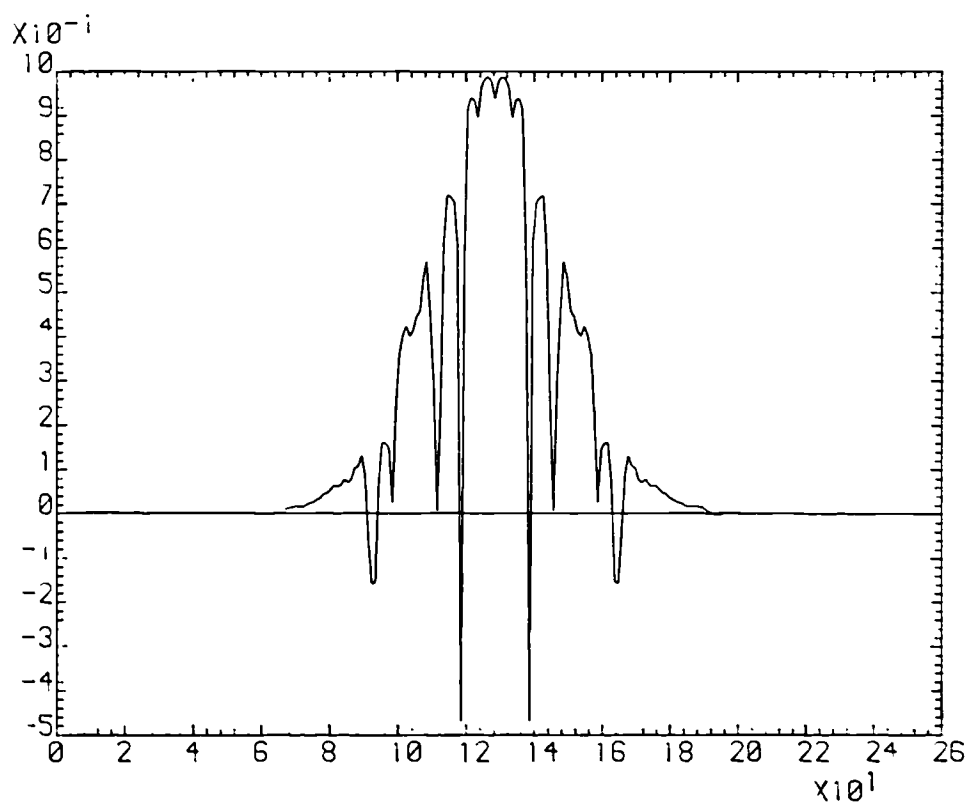


Figure 3.37(b) Cross-section at $x = 0$ of real field components directly behind unmodified E180 filter with E191 as input

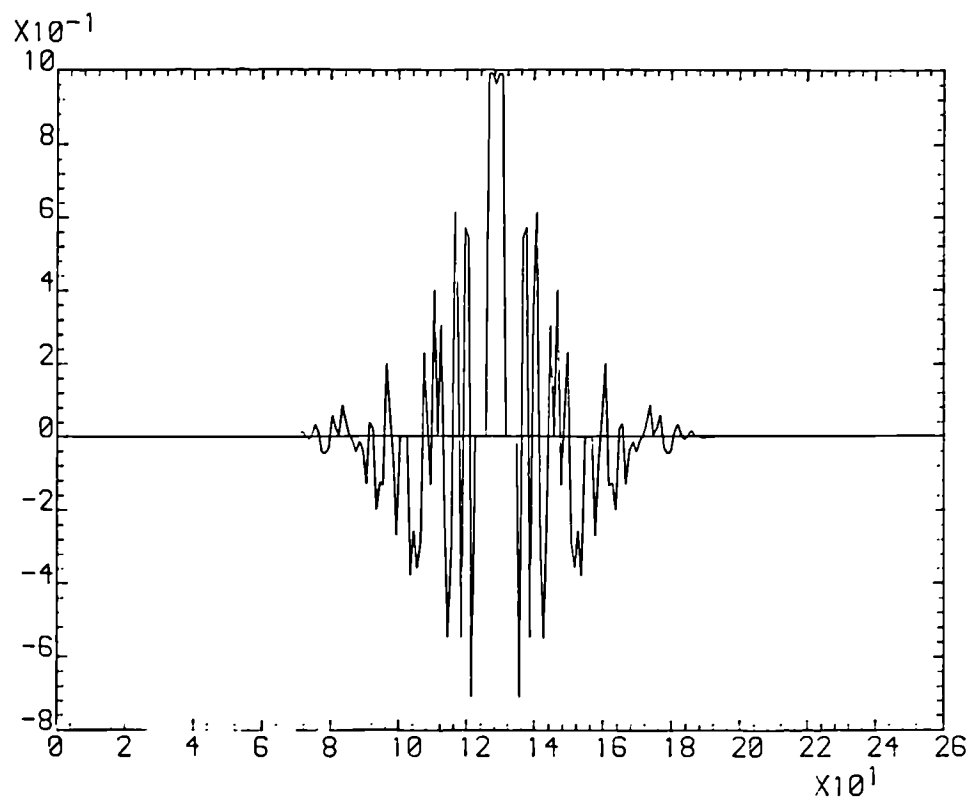


Figure 3.38(a) Cross-section at $y = 0$ of real field components directly behind modified E180/E191/M180 filter with M180 as input

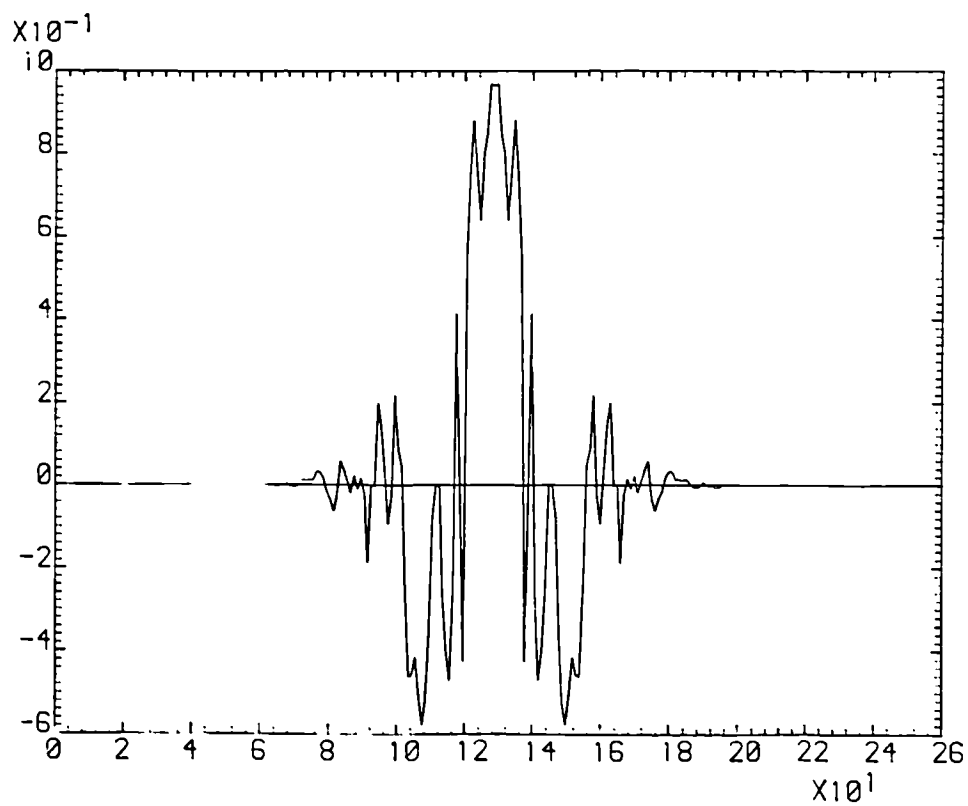


Figure 3.38(b) Cross-section at $x = 0$ of real field components directly behind modified E180/E191/M180 filter with M180 as input

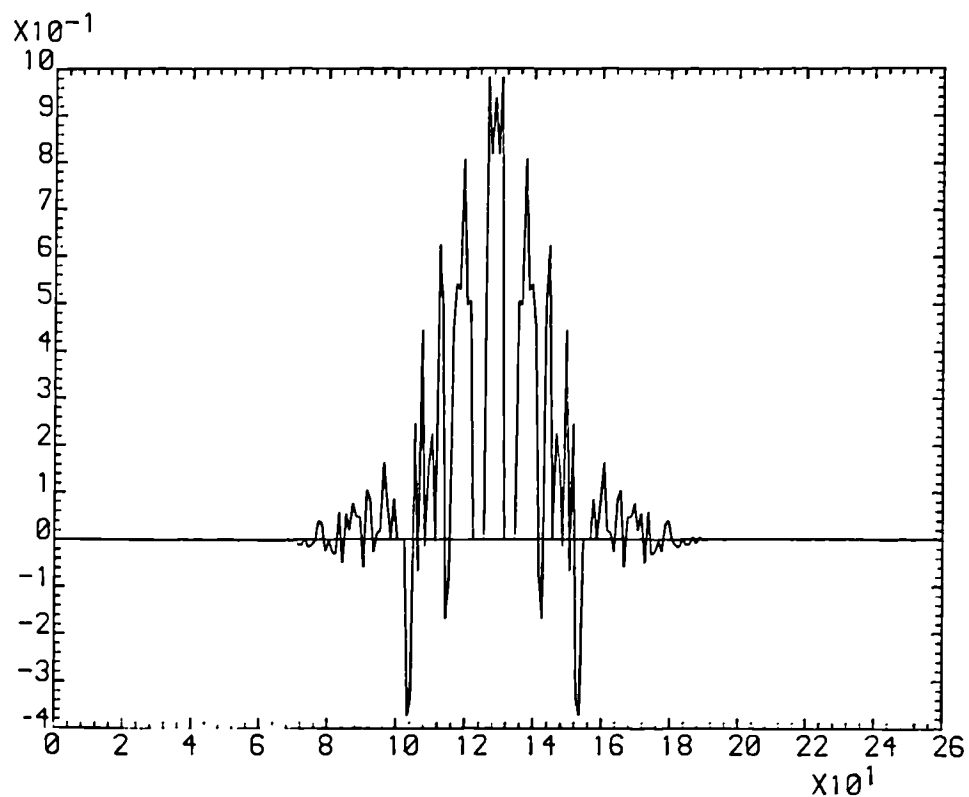


Figure 3.39(a) Cross-section at $y = 0$ of real field components directly behind modified E180/E191/M180 filter with E191 as input

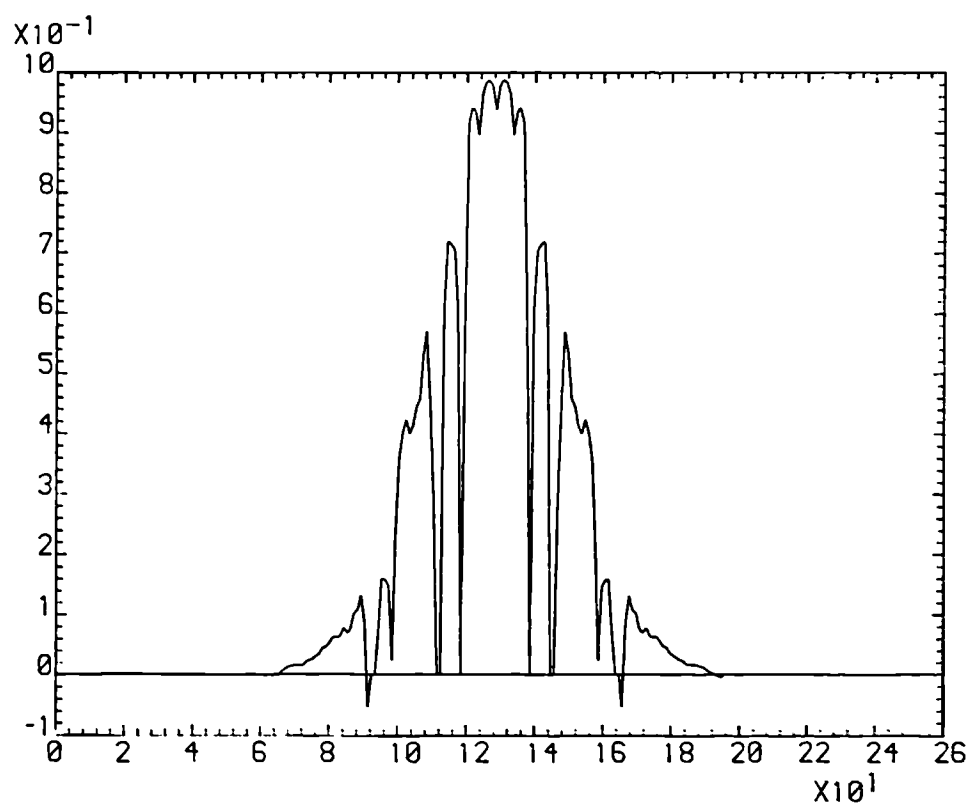


Figure 3.39(b) Cross-section at $x = 0$ of real field components directly behind modified E180/E191/M180 filter with E191 as input

3.4 Conclusions

In the foregoing examples the filter was modified by selective frequency removal in order to maximise the intra-class response to a change of orientation of the target car while maintaining discrimination against a similarly orientated non-target car. Only a single complex spectrum of the target car in its original orientation was used to do this. The intra-class distortion was quite large corresponding to a $11\frac{1}{4}^\circ$ angle change of the target car from a front or rear view (from which orientation the shape changes most rapidly with viewing angle). This degree of orientation change would correspond to the maximum accommodated by that single filter; full 360° orientation coverage (or possibly 2π , or even 4π , steradian coverage) being achieved by separate filters which the correlator would rapidly sequentially access from a filter bank.

An issue requiring further simulation is the behaviour of the correlation response to views of the target car intermediate to the two views used in generating the modified filter. It has been noted in the previous section that the ACF will always be larger than the cross-correlation response (in the cases considered by more than three times). Since there is a smooth alteration of the target car's shape as the orientation is changed between the two extreme views used to generate the modified filter view, the spectrum of the car will alter smoothly between these two extremes also, although the rate of change of shape will not be constant with change of angle, being more rapid, for instance, with alterations from a front or rear view than a side view. Thus it may be expected that the correlation signal will approximately reduce linearly in height between the auto-correlation response and that resulting from the $11\frac{1}{4}^\circ$ orientation change. For the filter to remain effective over this swathe of angles the inter-class response should remain well below the lower boundary of the intra-class response. This is known to be the case when the non-target car view orientation is the same as that of the target car since the filter is designed to give a zero on-axis response for this orientation. When the orientation is changed, the on-axis intra-class cross-correlation response will fall due to the alteration in appearance of the input. However, the modified filter will no longer necessarily have the correct frequencies removed to give maximum discrimination between the intra- and inter-class views since these may well alter with the orientation of the inputs. Thus the intrinsic loss of correlation due to alteration of the non-target car's orientation must be largely relied on to provide a minimal inter-class response (from the simulation results this has been seen to be rapid even for different views of the same car). The possibility does arise, however, that secondary peaks will be produced off-axis which will have to be significantly smaller than the lower bound of the intra-class response. Further simulation studies are

required to see if the filter remains effective for recognition of target and non-target views intermediate to those used to generate the filter.

A disadvantage of the filter described is the large reduction of the correlation peak when the car view alters from that used to generate the filter. As pointed out above, there remained over a three times reduction in the correlation response when the orientation was changed by $11\frac{1}{4}^\circ$, even after the filter was modified by the frequency removal technique to prevent the cross-correlation collapsing completely. A more uniform filter response is desirable although, provided the discrimination capability of the filter is adequate, the large ACF response could be ignored by simply setting the output detector such that it saturates well below this maximum (this is easy to do with a CCD detector although a problem could be incurred by spillage of the photo-generated charge between the CCD pixels resulting in blooming of the detector, although modern CCDs are designed to alleviate this effect).

The SDF technique has been developed, of course, to overcome the problem of intra-class correlation peak height variability by constraining the on-axis correlation plane pixel to a value set in the design, as described in detail in Section 2.3. However, as discussed in this section the specified value will only be maintained for intra-class views explicitly contained in the training set. If the SDF is of the high-pass variety (eg. a MACE filter) many training set views are required to maintain the intra-class non-training set response of the filter. If the filter transfer function is in a bandpass the intra-class degradation to non-training set data will be reduced; however, it may be necessary to include inter-class views in the SDF with an on-axis correlation plane value specified to be zero, in order to suppress otherwise high inter-class correlation responses. This will lead to negative weighting coefficients in the SDF filter which, if it is to be introduced into the correlator as a space domain function on an SLM, will require a bias to be added to ensure the SDF image is all positive. The SDF filter will then produce the correct correlation plane weightings (since the bias is removed by the null zero frequency response of the bandpass filter) but higher demands are placed on the dynamic range requirement of the SLM. It must be emphasised that the SDF technique is a dynamic range multiplexing technique and, as shown in Section 2.4, results in a reduction of the correlation peak in proportion to the number of images multiplexed within a single filter (as a worst case when there is no cross-correlation between the stored images).

Combination of an SDF multiplexed filter with the frequency removal technique would alleviate the demands on the dynamic range of the input SLM by reducing the

level of multiplexing while still maintaining a much more constant intra-class correlation response than can be achieved with a single image filter of the type described above. It should be possible to obtain a minimal inter-class response, without placing further demands on the degree of multiplexing, by the selective removal of frequency components from the spectrum of the multiplexed input image.

The method of producing the filter may be broken down into the following steps :

1. Multiplex together intra-class images with uniform weightings covering the swathe of orientation angles that the filter is to accommodate. The increments in angle should be such that intermediate non-training set target orientations do not have significantly degraded cross-correlation. The inter-class object is similarly multiplexed into a separate image, again with uniform weightings over the same range of angles.
2. The frequency removal procedure is applied to the spectra of intra-class and inter-class object views that are of intermediate orientation, within the swathe of angles for which the filters are multiplexed.
3. The SDF design procedure is applied to the intra-class multiplexed images as summarised in Section 2.3 but using the MTF generated by the frequency removal technique. Thus weighting factors are calculated to apply to the intra-class training images in order to ensure they give a constant correlation response and that intermediate views give a more nearly constant response within the range of angles covered by the filter.

It can be appreciated that this method is not guaranteed to produce a zero on-axis response to all views of the inter-class object since specific intermediate views of the target and non-target objects have been used in the frequency removal process. An alternative would be to use a multiplexed inter-class filter and suppress the cross-correlation of this with an intermediate intra-class object view by removing specific frequencies. It can be seen that this method is similar to the suggestion made by Yaroslavsky [6] for de-correlating target and non-target objects. The filter transfer function he suggests is:

$$H(\mathbf{k}) = \frac{S^*(\mathbf{k})}{AV_{sn} |S_{\eta}(\mathbf{k})|^2} \quad \{3.24\}$$

where $S^*(\mathbf{k})$ is the complex conjugate spectrum of the object to be detected and $S_{\eta}(\mathbf{k})$ is the spectrum of the object(s) to be rejected. AV_{sn} represents an averaging operation on the individual non-target spectra. This filter is the same as the matched spatial filter (see Equation {1.7}) where the average of the non-target spectra is considered as the noise spectrum, $|N(\mathbf{k})|^2$. Also, a similar technique is a modification of the binary phase only filter, suggested by Flannery *et al.* [18], by the addition of a binary amplitude modulation which depends on the local ratios of intra-class and inter-class averaged power spectra, as mentioned previously in Section 1.2.3.

All these methods seek to increase the discrimination of the filter against the inter-class object by the removal of selective spatial frequencies rather than by applying a uniform negative bias to the spectrum as is effectively done by including the inter-class images directly into an SDF multiplexed filter. By considering field values rather than the power spectrum, the technique described in this section has the advantage of being able to increase the intra-class correlation, since portions of the intra-class spectrum that have negative real field components may be identified and removed. Further simulation work is required to determine if the method is effective when used with multiplexed filters and over what range of orientations of the object it remains effective. The latter, of course, will be strongly data specific being dependent, for instance, on the rate of distortion of a given object with aspect changes.

Since intra- and inter-class object views are in general highly correlated, the filter should ideally have the form of the complete Wiener filter as discussed in Section 1.2.4. If the filter is to be multiplexed, the noise spectrum (required in full complex form and not as a power spectral density) should be some average of the inter-class object taken over the same range of angles as those of the target object used in the generation of the filter; it is this data that is most likely to generate strong cross-correlations. The assumption must be made that if the filter is able to discriminate well against the similar inter-class data, its performance against background clutter will be good also. Whether this complete form of the transfer function is indeed required, and what performance increase it gives over simply using the inter-class averaged power spectrum in a matched filter, requires investigation.

References to Chapter 3

- [1] E. Brigham, "The fast Fourier transform and its applications," Prentice Hall, (1988)
- [2] S. Lipson, H. Lipson, "Optical physics," Cambridge, (1981)
- [3] J. Goodman, "Introduction to Fourier optics," McGraw-Hill, (1968)
- [4] R. Easton, H. Barret, "Tomographic transformations in optical signal processing," Section 5.2 in *Optical signal processing*, J. Horner Ed., Academic press, (1987)
- [5] V. Leavers, J. Boyce, "The Radon transform and its application to shape parametrisation in machine vision," *Image and Vision Computing*, Vol. 5, No.2, p.161, (1987)
- [6] A Mahalanobis, B. Vijaya Kumar, D. Casasent, "Minimum average correlation energy filters," *Appl. Opt.*, Vol. 26, No. 17, p.3633 (1987)
- [7] L. Yaroslavsky, "Is the phase-only filter and its modifications optimal in terms of the discrimination capability in pattern recognition?" *Appl. Opt.*, Vol. 31, No. 11, p.1677, (1992)
- [8] D. Flannery, J. Loomis, M. Milkovich, "Transform-ratio ternary phase-amplitude filter formulation for improved correlation discrimination," Vol. 27, No.19, p.4079, (1988)

Chapter 4

Implementation and experimental assessment of bandpass filters

4.1 Introduction

The discussion of previous Chapters has emphasised that in application of the correlation technique to the recognition of a three dimensional object there is a requirement for some form of bandpass filtering in order to maintain inter-class discrimination while accommodating a degree of intra-class tolerance. In this Chapter, the experimental implementation of bandpass filters in a photorefractive based joint transform correlator is described. The performance improvements over all-pass correlation is assessed with input imagery consisting of a target and non-target car in a range of orientations at zero elevation angle with fixed scale and in-plane rotation.

A function particularly convenient for the implementation of a bandpass filter is the difference of Gaussian (DOG) approximation to the Laplacian of a Gaussian. This function may be generated directly by exploiting the TEM_{00} mode of the Helium Neon laser employed as the read-out beam of the photorefractive Bismuth Silicon Oxide crystal in the non-degenerate four wave mixing implementation of the up-dateable correlator. After briefly reviewing the theory behind the DOG filter, its experimental implementation in the joint transform correlator is described and the resulting correlation performance assessed with the car test imagery.

4.2 Joint transform correlator experimental implementation

A schematic of the experimental joint transform correlator is shown in Figure 4.1 and a photograph of it in operation in Figure 4.2. It was constructed from standard optical components on a vibration isolated optical table. Some details of the components and equipment used are given in Table 4.1. The main difficulty encountered was the requirement for absolute interferometric stability in order that stable volume holographic gratings could be written into the photorefractive material. To accomplish this, rigid

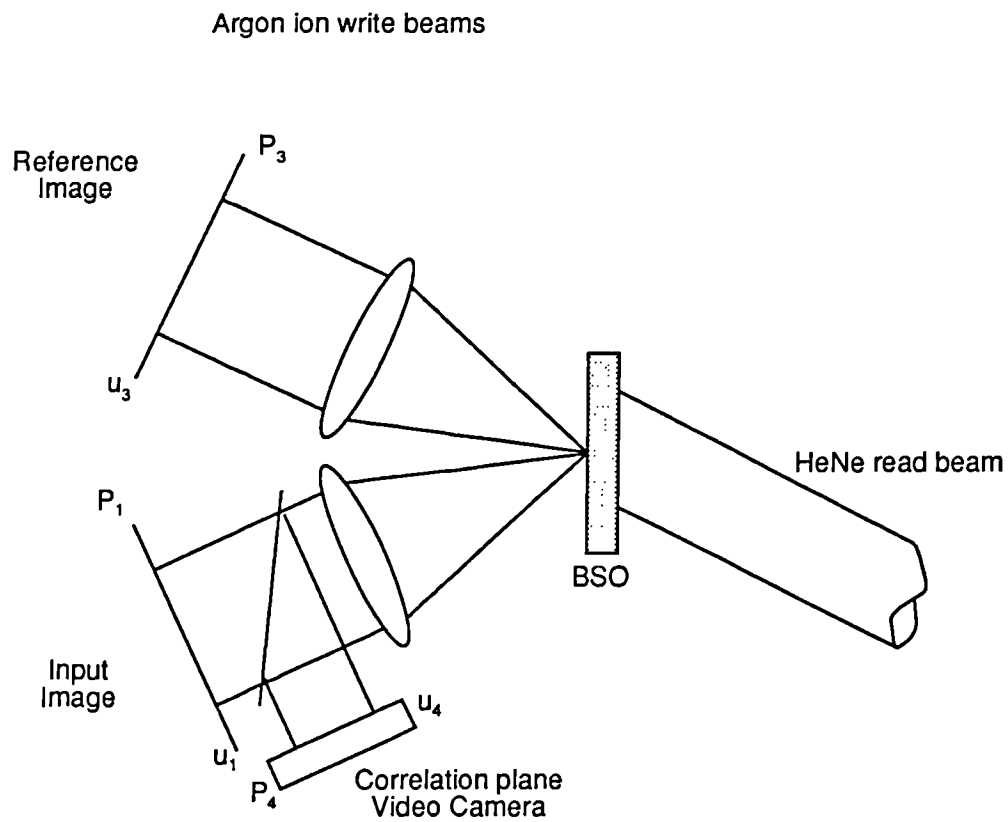


Figure 4.1 Schematic of photorefractive based joint transform up-dateable correlator.

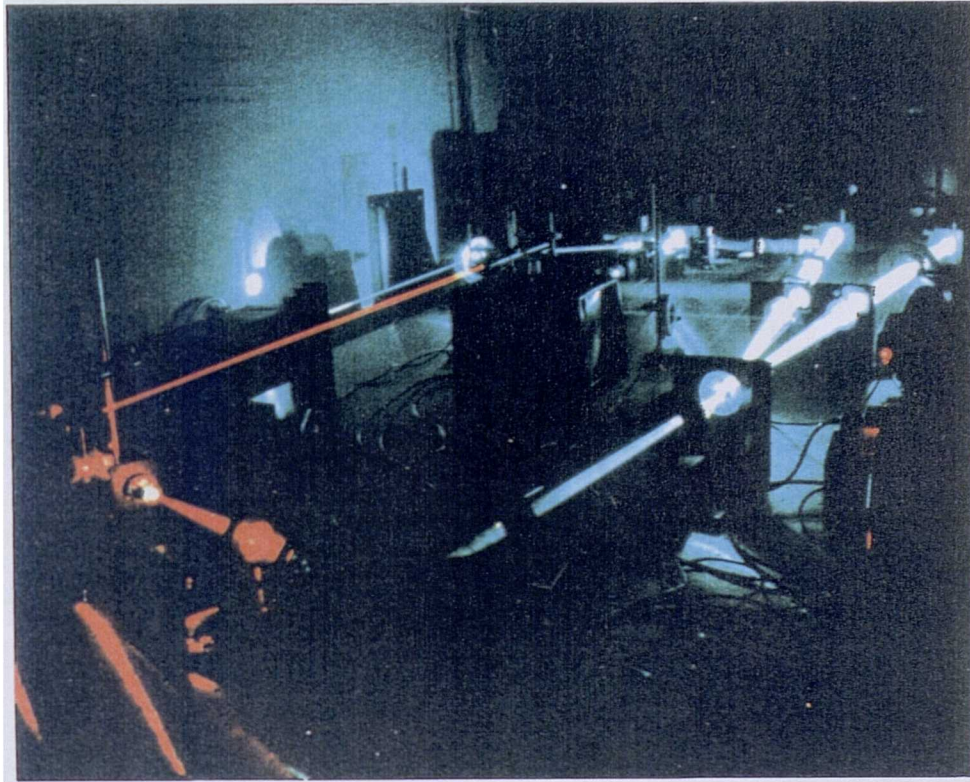


Figure 4.2 Experimental implementation of joint transform correlator

Equipment Item	Manufacturer	Specifications and Comments
ARGON ION LASER	COHERENT	<p>INOVA 90-2 Power: 1.5 W at 514.5 nm 1.2 W at 488 nm Coherence length approx. 2 cm Linearly polarized (Intra-cavity Fabry-Perot etalon removed, prism wavelength selector)</p> <p>Beam diameter: 1.2 mm to $\frac{1}{e^2}$ points Beam divergence: 0.6 mrad Long term power stability: $\pm 0.5\%$</p>
HELIUM NEON	SPECTRA-PHYSICS	<p>MODEL 105 Power: 5 mW Linearly polarized</p> <p>Beam diameter: 0.9 mm to $\frac{1}{e^2}$ points Beam divergence: 1 mrad</p>
FOURIER TRANSFORM LENSES	MELLESGRIT	<p>TYPE: O1LAO 349 OPTIMISED ACHROMATIC DOUBLET Focal length: 600 mm Diameter: 63 mm</p>
BISMUTH SILICON OXIDE CRYSTAL ($\text{Bi}_{12}\text{SiO}_{20}$)	GEC-HIRST	<p>Dimension: 10 mm x 10 mm x 1.4 mm (Symmetry Group $\bar{4}3\text{M}$) Electric field applied parallel to (110) Crystallographic axis Gold electrodes for electrical HT contact Orientation exploits $r_{41} (=5.0 \times 10^{-12} \text{ m/V})$ linear electro-optic tensor (Pockels' coefficient)</p>
HT POWER SUPPLY FOR CRYSTAL	WALLIS TYPE PM100CP	<p>Supply: 24 volts DC Output: 0 - 10 kV</p>
Auxillary Optics		
MIRRORS	EALING-BECK	<p>Flatness: $\lambda/10$ Diameter: 50 mm Coating: Aluminium with SiO_2 protective layer Material: Pyrex</p>
NEUTRAL DENSITY FILTERS	<p>MELLES GRIT</p> <p>EALING BECK</p>	<p>REFLECTIVE TYPE PRECISION METALLIC Size: 50 mm x 50 mm Flatness: 4λ Material: BK7 Glass N.D. accuracy: $\pm 5\%$</p> <p>ABSORBATIVE TYPE Size: 50 mm x 50 mm Flatness: 5λ N.D. accuracy: $\pm 5\%$</p>

Equipment Item		Specifications and Comments
BEAM-SPLITTERS		CUBE TYPE: 50:50 Split: $\pm 5\%$ Flatness: $\lambda/2$ Material: Crown glass
COLOUR GLASS FILTER	EALING BECK (SCHOTT)	SCHOTT No. 06-570
HALF WAVE PLATE	MELLES GRIOT	Material: Mica Flatness: $\lambda/4$ Retardation tolerance: $\lambda/20$
POLARIZER	MELLES GRIOT	DICHROIC POLARIZING SHEET Extinction ratio: 10^{-4}
Beam Expander		
MICROSCOPE OBJECTIVES	EALING BECK	x24 N.A. = 0.54 x5 N.A. = 0.12
PINHOLES (PRECISION)	EALING BECK	Diameter: 20 μm
COLLIMATING LENS	EALING BECK	TYPE: AIR SPACED DOUBLET Focal length: 160 mm Diameter: 40 mm
OPTICAL TABLE	ANASPEC	ZERO G ^(R) Size: 3 m x 1.5 m Low pressure gas isolation system Steel sheet top
LASER VIBRATION ISOLATION	PHOTON CONTROL	ZORBOPHANE ^(R) RUBBER (isolation cups) Isolation of Argon ion laser vibrations caused by flow of cooling water
OPTICAL MOUNTS	DESIGNED AND FABRICATED IN-HOUSE	Solid Aluminium mounts to achieve interferometric stability
MICROMETER MIRROR MOUNTS	EALING-BECK	Accuracy: 10 arc seconds Range: 3° Accurate beam pointing to achieve Bragg matching in the volume hologram

Equipment Item	Manufacturer	Specifications and Comments
VIDEO CAMERA	VICON	ULTRACON (silicon) faceplate Faceplate: $\frac{3}{4}$ " CCIR Standard: (625 lines, interlace, 25 frames/sec) High Sensitivity in red/near infra-red region of spectrum Used to detect correlation plane
PHOTODIODE	EALING BECK	SILICON PHOTODIODE Rise Time (10% to 90%): 50 nsec Peak Wavelength: 900 nm Responsivity: 0.2 A/W at 450 nm 0.35 A/W at 633 nm Area: 100 mm ² (used with 20 μ m - 100 μ m pinhole covers) Used with variable gain amplifier
OSCILLOSCOPE	TETRONIX 468	Digital storage, dual time base Bandwidth: 50 MHz
ISOMETRIC DISPLAY DEVICE	DESIGNED AND BUILT IN-HOUSE	Input: CCIR standard composite video signal Output: X, Y, Z, channels of oscilloscope

Table 4.1 Summary of equipment used in the experimental implementation of the joint transform correlator

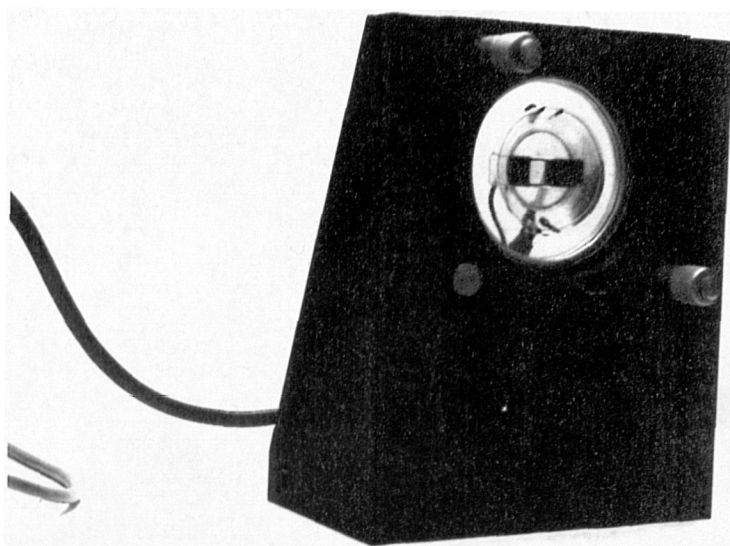


Figure 4.3(a) Photorefractive crystal mount with insulating support and HT electrodes

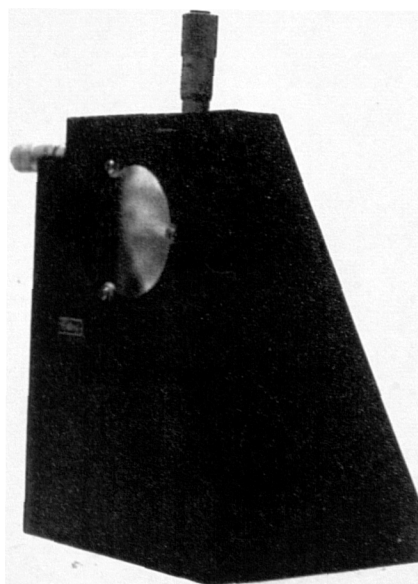


Figure 4.3(b) Mount supporting micrometer adjustable mirror holder

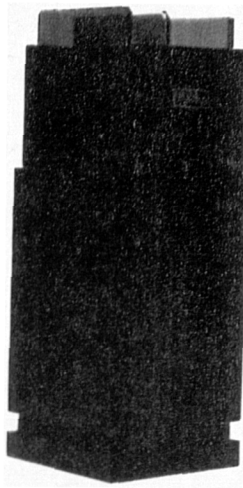


Figure 4.3(c) Filter holder

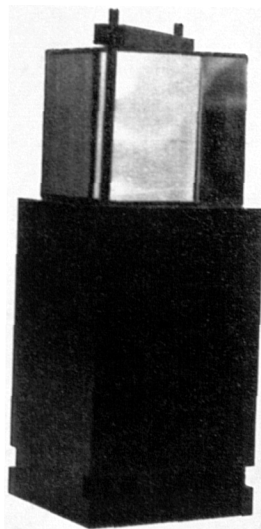


Figure 4.3(d) Cube beamsplitter mount

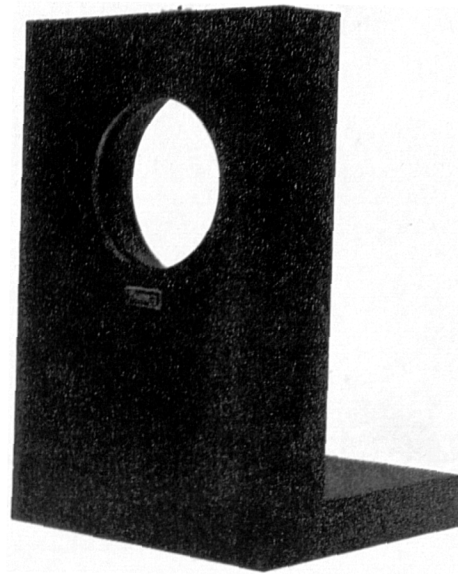


Figure 4.3(e) Lens mount

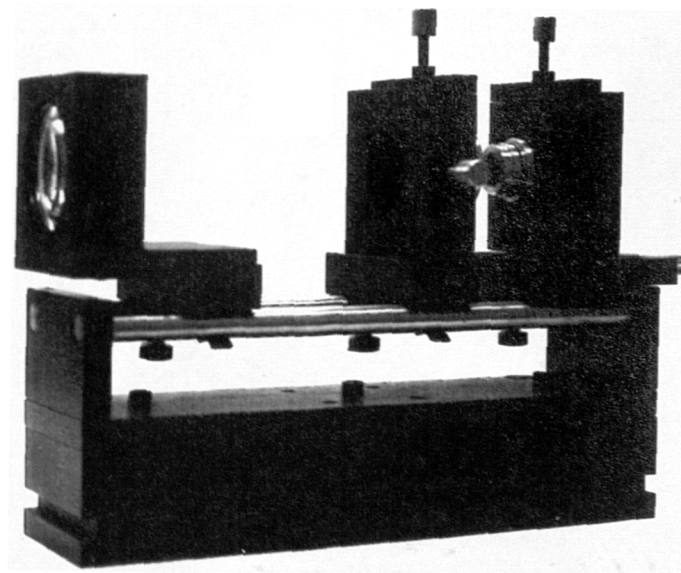


Figure 4.3(f) Beam expander/spatial filter assembly

optical mounts were designed and constructed, examples of which are shown in Figure 4.3, and the water cooled Argon ion laser mounted on vibration isolation supports to prevent on-table mechanical disturbance. The frequency plane complex multiplication between the optically formed Fourier transforms of the input and reference functions is accomplished by the non-linear interaction of the mutually coherent Argon ion wavefronts within the photorefractive crystal Bismuth Silicon Oxide ($\text{Bi}_{12}\text{SiO}_{20}$ - abbreviated to BSO).

The recorded grating is read-out with a plane coherent beam generated by a Helium Neon (HeNe) laser which has a negligible non-linear interaction with BSO due to its longer wavelength and low intensity. This may, therefore, be considered to simply diffract from the grating written to the BSO by the Argon ion beams. Formally, the process is described as a non-degenerate four wave mixing interaction (NDFWM), the fourth wave being that diffracted from the grating, although it is not the purpose of this Chapter to describe the non-linear interaction any further, full accounts of the linear electro-optic (Pockels') effect and photorefractive effect being available from the literature [1,2]. For an examination of the filtering effects that can be implemented using NDFWM in the photorefractive material it is adequate to describe the interaction by the transfer function :

$$m(\mathbf{k}) = \frac{2U_1^*(\mathbf{k}) U_3(\mathbf{k})}{|U_1(\mathbf{k})|^2 + |U_3(\mathbf{k})|^2} \quad \{4.1\}$$

where the upper case letters represent the Fourier transforms of the input functions, $u_1(\mathbf{r})$ and $u_3(\mathbf{r})$, and $m(\mathbf{k})$ is the modulation of the holographic grating formed in the BSO by the non-linear interaction of $U_1(\mathbf{k})$ and $U_3(\mathbf{k})$. In the steady state this transfer function adequately represents the modulation of the HeNe addressing beam, simulations giving accurate predictions of the effect of alterations of $m(\mathbf{k})$ on the correlation plane light distribution. However, a further factor of critical importance to the experimental implementation are the Bragg matching effects due to the volume nature of the holographic gratings formed in the BSO and the limiting effect this has on the attainable space bandwidth product (SBP) of a photorefractive based correlator of given physical dimensions.

Predictions of this effect can be made using Kogelink's coupled wave theory [3]. The modulation of the diffraction efficiency from a lossless dielectric volume transmission hologram takes the form of a $|\text{sinc}(x)|^2$ function. In the case of a dynamic grating in a photorefractive, the $|\text{sinc}(x)|^2$ function takes on an asymmetric form due to the grating tilt

and change in spatial frequency as the input beam, $U_1(\mathbf{k})$, is changed in direction [4]. In DFWM this has the effect of maintaining the Bragg condition over large excursions of the input beam angle. However, in NDFWM the wavelength shift of the readout beam leads to a loss of perfect Bragg matching with change of angle of $U_1(\mathbf{k})$, although the field of view reduction is not very severe due to the compensating effect of the combined tilt and change of spatial frequency of the grating. Thus an acceptably large field of view is obtainable in beam $U_1(\mathbf{k})$. However, as beam $U_3(\mathbf{k})$ changes angle away from the Bragg condition the holographic fringes change spatial frequency but tilt in the *wrong* direction. Thus the field of view in this beam is considerably reduced as compared to that from a static grating.

Kogelink's theory in its original form does not give an accurate estimate of the field of view from a grating formed in BSO [3] since it does not take account of the electric field induced birefringence and strong optical activity (39°/mm at 514.5 nm) present in this material. The experimentally measured values for the field of view correspond in functional form to that predicted by coupled wave theory but suggest a grating thickness approximately half that of the crystal thickness [5]. For the experimental parameters outlined in Table 4.1 the fields of view at the FWHM points on the $|\text{sinc}(x)|^2$ modulation were 18 mm and 2.3 mm for beams 1 and 3 respectively, clearly indicating the asymmetry expected from the effect described above. To maximize the SBP of the correlator the input function is, therefore, used to modulate beam 1 and the reference function beam 3. In this way the maximum amount of shift-invariance can be achieved for a given set of parameters of the correlator, allowing the reference object to be identified and its location determined from within a larger background scene.

Nevertheless, for the BSO crystal thickness employed the field of view in beam 3 is quite restrictive, requiring the reference function imagery to be recorded on high resolution film (Kodak orthochromatic copy film). If an SLM of more limited resolution had been employed (such as the Seiko-Epson LCTV) a demagnifying lens system would have been required prior to input of the image to the correlator. However, it should be noted that there is a fundamental limit on the SBP achievable in a correlator with a given set of parameters, arising from a combination of a finite space domain aperture (due to the Bragg matching restrictions) and a finite frequency domain aperture (due to the finite size of the BSO crystal). Thus the input image cannot be demagnified indefinitely (apart from any practical imaging limitations) since the mid-band spatial frequencies of interest will then fall outside the aperture of the BSO crystal. If the half-width of the crystal is $x_{f(\text{max})}$, the maximum spatial frequency will be $k_{x(\text{max})} = x_{f(\text{max})} / \lambda f$ or 14.6 cycles/mm with $\lambda = 514.5$ nm and $f = 600$ mm. Parallel to the grating vector, the maximum number of

resolution elements will be: $2 \times k_{x(\max)} \times \text{FWHM} = 68$ data points. Thus it can be seen that the Bragg matching effects are quite restrictive even when long focal length lenses are employed in the correlator.

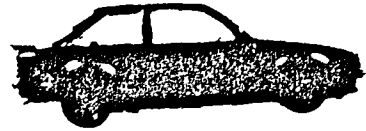
As mentioned above the imagery used in the correlator was recorded on high resolution copy film in order that apodisation of the reference image due to the Bragg matching effect in beam 3 could be minimized. The images could not be reduced beyond a certain size, however, since the limited MTF of the film and frequency domain apodisation would then generate unwanted low-pass filtering effects. The scale reduction used was a 46.8 times reduction of the $1/25^{\text{th}}$ scale models of the Escort and Mazda cars employed as test imagery. The largest image was that of the Escort sideview which measured 2.03 mm between wheel bases. This was the only image in which there may have been a slight degree of apodisation due to the Bragg matching constraint. It was found necessary to further apodise the frequency spectrum of the reference and input images to realise a filter with a useful intra-class distortion tolerance. Thus the available area of the BSO crystal is more than adequate for imagery of this size. The spectra of the inputs (examples of which as shown in Figure 4.5) extended well beyond the limits of the crystal aperture indicating the images recorded on the film had a higher resolution than required. This was confirmed by examination under a microscope at low power, the recorded images showing crisp detail with a good grey tone representation and minimal grain size, which was only visible under a high power. (The high quality achromatic doublets employed as the Fourier transform lenses had a 50% MTF of over 200 lines/mm and so were far removed from being a system limitation.)

Some example views of the Escort and Mazda car models used as test imagery to the joint transform correlator are shown in Figure 4.4. The photographs in this figure were recorded on a Kodak PAN F film for reproduction onto photographic paper. This film has too large a grain size to be used for recording the high resolution imagery required for direct modulation of the coherent wavefronts into the correlator. In addition it is only possible to achieve an optical density of the order of unity on a PAN F film. This leads to a non-zero background illumination being imposed on the car image which in turn generates a $|\text{sinc}|^2$ spectral intensity function resulting from the Fourier transform of the square aperture (although the phase will be complex due to the car image also modulating the space domain function). In contrast, using the orthochromatic film it is possible to obtain an optical density greater than four when the car image is photographed against a uniform white background. The negative can then be used directly to modulate the coherent processing wavefront with the high resolution image of the car recorded within the uniformly opaque background. The car models were made in

red plastic allowing long exposures to be made to obtain saturation of the background while simultaneously ensuring the fine detail of the models is maintained, the orthochromatic film having a very low sensitivity at the red end of the spectrum. Views of the Escort and Mazda were recorded at $11\frac{1}{4}^\circ$ increments of orientation, with zero elevation angle and constant scale. Some example spectra of the car views are shown in Figure 4.5. Note the complicated structure of the intensity distributions in comparison to the simple $|\text{sinc}|^2$ structure that results from inputs with a non-zero background. As remarked in Chapter 3 the spectral energy is more dispersed in the vertical direction due to the fine detail of the car imagery being composed of largely horizontal structures. Figure 4.6 shows the same spectrum recorded with different exposures. The white horizontal lines shown in the maximum exposure locate the position of the edges of the photorefractive crystal in the joint transform correlator. The dynamic range of all the car spectra vastly exceeds that which it is possible to record on the high γ orthochromatic film used to photograph the spectra, the longer exposures revealing the weaker high spatial frequency components but resulting in saturation of the much more intense low frequency terms. The limited dynamic range of high resolution holographic plate can, of course, be used to good effect in order to generate bandpass matched filters as described in Chapter 1.



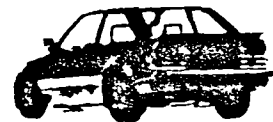
E000



E270

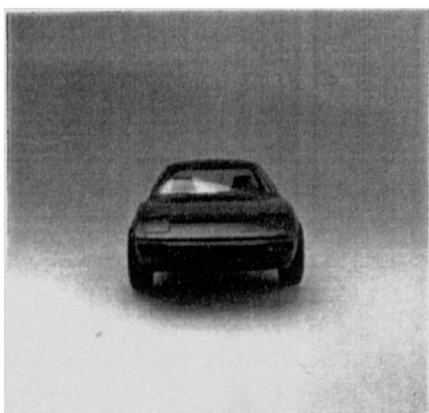


E315

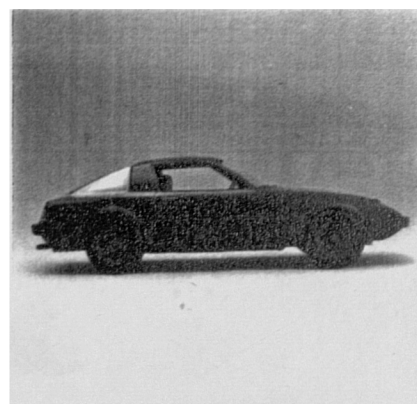


E135

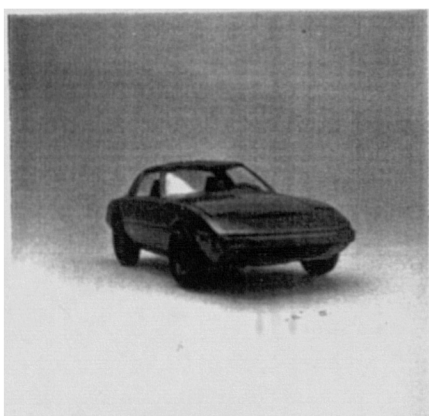
Figure 4.4(a) Example views of the Escort car test imagery



M000



M270



M338



M158

Figure 4.4(b) Example views of the Mazda car test imagery

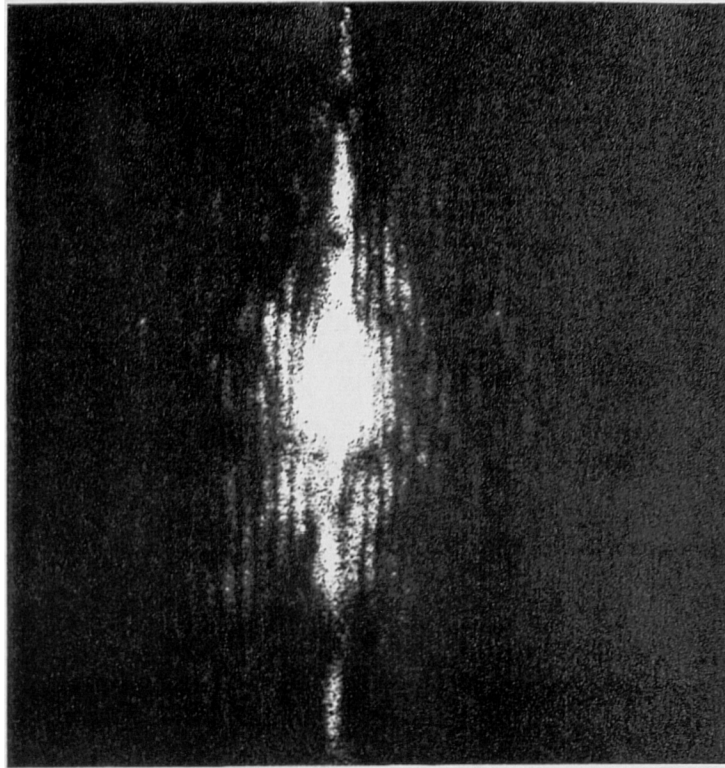


Figure 4.5(a) Power spectrum of Escort car - view E090

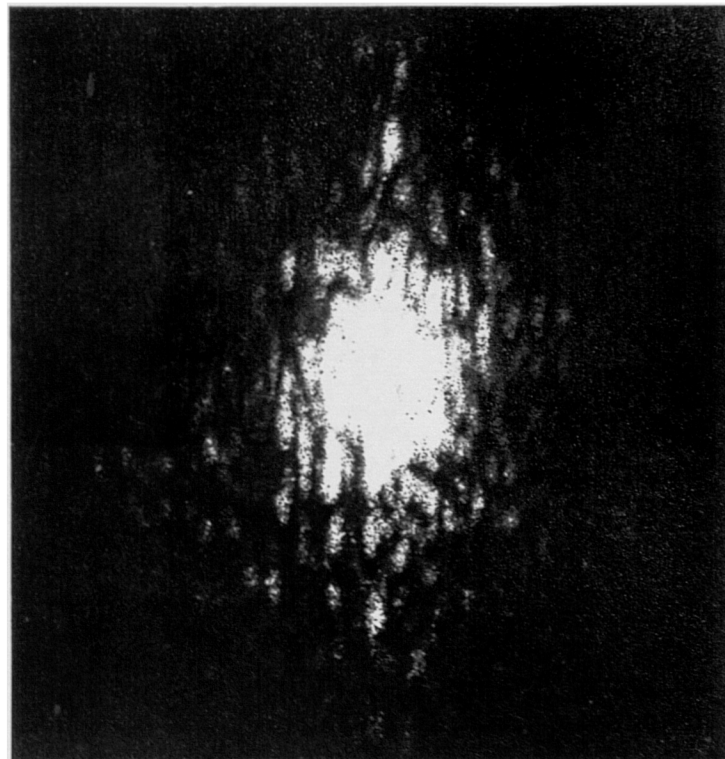


Figure 4.5(b) Power spectrum of Escort car - view E135

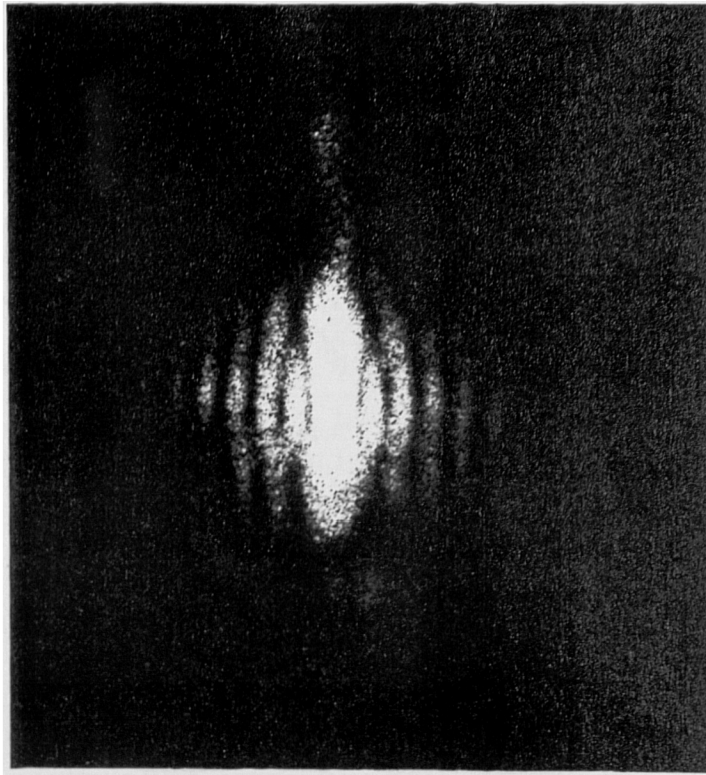


Figure 4.5(c) Power spectrum of Mazda car - view M180

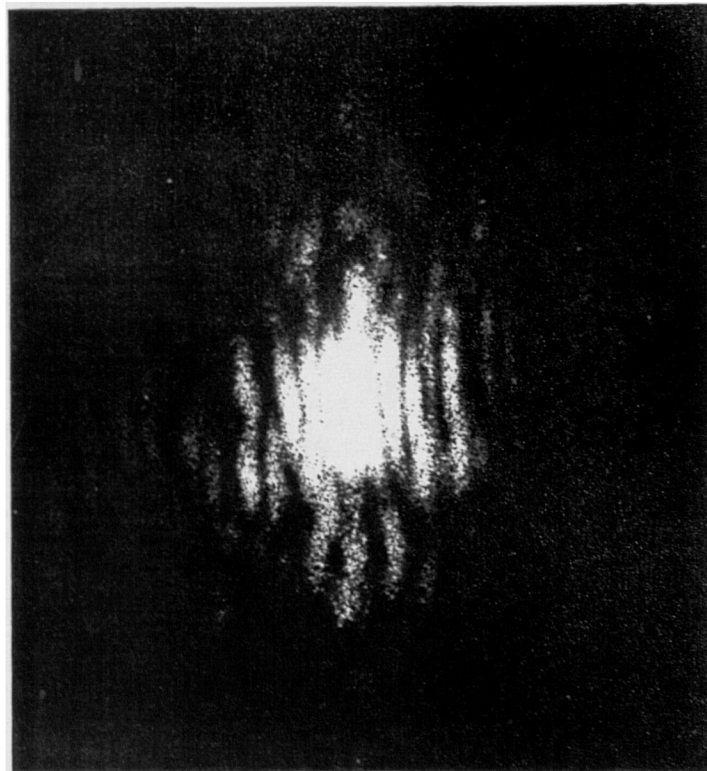


Figure 4.5(d) Power spectrum of Mazda car - view M158

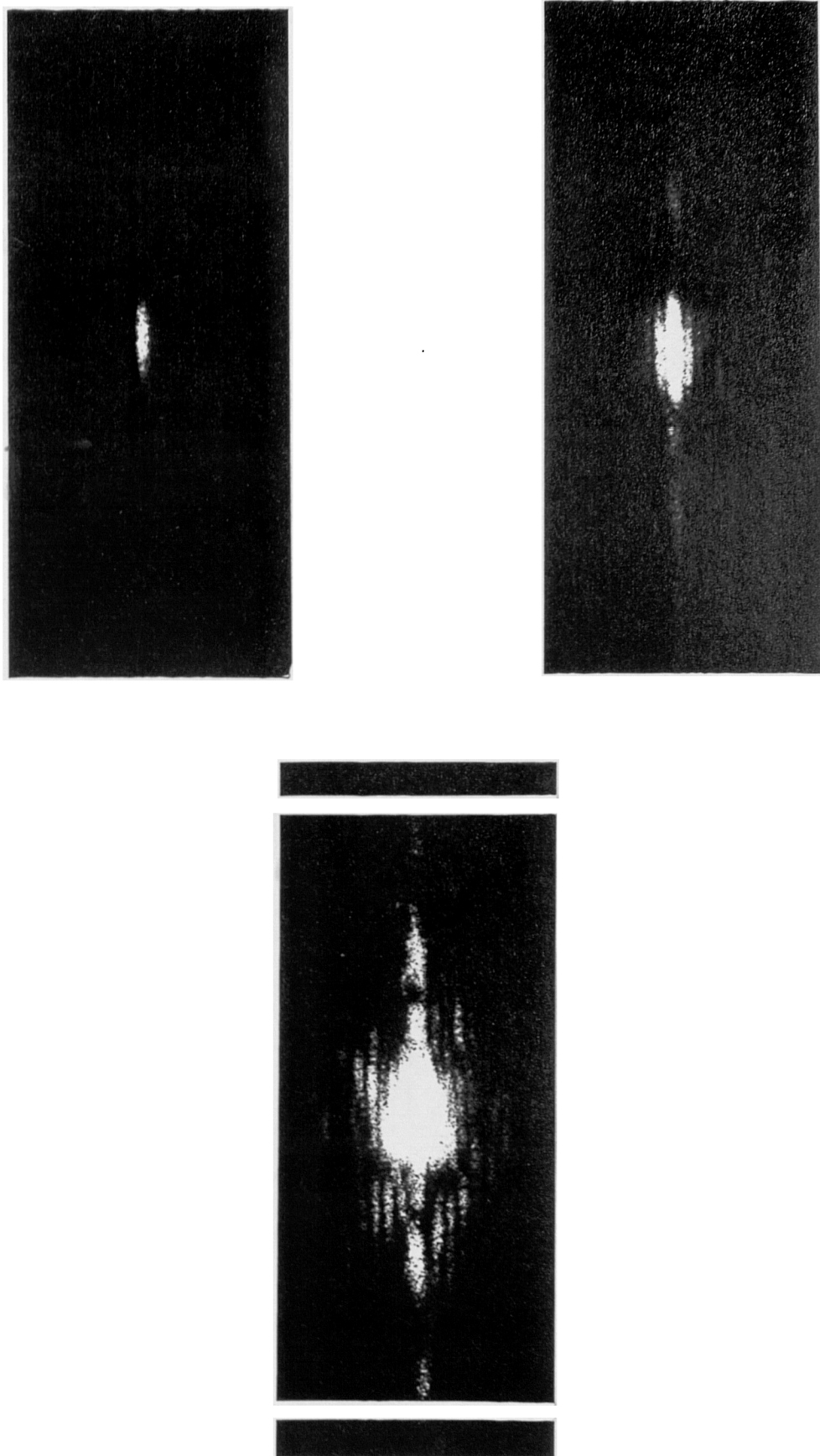


Figure 4.6 Power spectrum of Escort car (view E090) at varying exposures

4.3 Joint transform correlator full bandwidth performance

As described in Chapter 3, the auto-correlation response of a joint transform correlator is independent of the shape and extent of the input image since Equation {4.1} reduces to a constant when $U_1(\mathbf{k})$ and $U_3(\mathbf{k})$ are identical. Thus the ACF corresponds to the Fourier transform of the HeNe readout beam profile. Initial experiments were performed with the HeNe beam expanded sufficiently that it could be considered as being essentially uniform over the 9 mm x 9 mm clear aperture of the BSO crystal. In this configuration the correlator auto-correlation function, therefore, had the functional form :

$$\text{ACF}(x, y) = \frac{l_x l_y}{\lambda f} \text{sinc}\left(\frac{l_x x}{\lambda f}\right) \text{sinc}\left(\frac{l_y y}{\lambda f}\right) \quad \{4.2\}$$

(l_x, l_y being the dimensions of the frequency plane aperture). The measured values were, of course, the correlation plane intensity distribution corresponding to the square modulus of Equation {4.2}. The FWHM width of this function is given by $\Delta x_{1/2} = 0.886 \lambda f / l_x$ and the first minima are separated by $\Delta x_0 = 2\pi \lambda f / l_x$, which, for the parameters of the correlator, should result in the ACF being 0.265 mm at its base. The first local maxima have an intensity only 0.0471 of the peak at $1.43\pi \lambda f / l_x$ from the origin. In the simulations of Chapter 3 this sidelobe structure was barely visible even in the ideal noise free digitally calculated ACF response shown in Figure 3.13.

Figure 4.7(a) shows an isometric plot of the full bandwidth auto-correlation response of the joint transform correlator. This was obtained by imaging the output plane of the correlator onto the faceplate of a video camera (CCIR standard) and passing the video signal through an in-house developed isometric display module driving the x,y electrodes of an oscilloscope. In this way a useful qualitative two dimensional view of the whole correlation plane could be obtained. Figure 4.7(b) shows an oscilloscope linescan trace through the correlation plane. This is accomplished by triggering the slow-scan A trace of a dual trace oscilloscope on the video field sync. and the fast (B delayed) trace on the line sync. and then adjusting the trigger delay so as to display the video line passing through the correlation peak. The storage facility of the oscilloscope is employed to generate a time averaged response in order to remove high (temporal) frequency noise arising from the video camera. In addition to the correlation peak height, the width could be measured by relating the time base to the dimension of the correlation plane. In this way FWHMs of between 0.098 mm and 0.110 mm were measured for the full bandwidth auto-correlation response of the correlator which are in agreement with that expected, as determined from Equation {4.2}. The noise floor in the correlation plane,

resulting from coherent scatter from the various components of the system, is minimal but obscures the visibility of the weak sidelobe structure of the $|\text{sinc}(x)|^2$ ACF. As noted in Chapter 3 hard clipping in the frequency domain has, in practice, a negligibly deleterious effect on the sidelobe structure of the correlation function. Figure 4.7(c) and (d) show two further examples of the auto-correlation response of the correlator for the indicated car views. It can be seen from these line scans that the ACF remains unaltered in functional form with inputs of significantly different shape and size (particularly, note the FWHM remains constant) as predicted by consideration of the grating function (see Equation {4.1}) when the two input images are identical.

Figure 4.8(a) and (b) show the inter- and intra-class response of the full bandwidth correlation for the car views indicated, the lower oscilloscope line scan displaying the ACF for direct visual comparison with the cross-correlation response. These results should also be compared to the simulations shown in Figure 3.15 which suggest that the correlation response will be very sensitive to alterations of the input image from that of the reference, a complete collapse of the signal being predicted for both the inter- and intra-class test imagery. In the experimental case, however, a residual response does remain to both the inter- and intra-class inputs. In both cases the cross-correlation response is 0.15 that of the ACF peak. Thus although good discrimination between the rear views of the Escort and Mazda is possible, an $11\frac{1}{4}^\circ$ alteration in orientation of the Escort results in a loss of the correlation response. Figure 4.8(c) and (d) show the inter- and intra-class response of a front three-quarters view of the Escort and Mazda. Again discrimination between the cars is good, the cross-correlation peak being 0.09 of the ACF response. The intra-class peak is higher in this example, however, being 0.186 the ACF peak height. As has been noted previously in Chapter 3, the car alters most abruptly in shape for rotations away from the front and rear views. Thus the correlation response is expected to deteriorate more rapidly for alterations of the Escort from view E180 and E000 than for other orientations. The slight improvement in the intra-class response shown in Figure 4.8(d) as compared to that shown in Figure 4.8(b) may result from this difference, since a constant change in orientation angle was employed in each example to assess the intra-class response.

Table 4.2 shows the full bandwidth correlator response to in-plane rotations of the test input with respect to the reference image for the views of the Escort indicated. As expected there is a high degree of sensitivity to in-plane rotation, comparable to that obtained with the Vander Lugt correlator configuration when this is tuned to the highest frequency response with a beam ratio of 32 (see Table 2.2). The sensitivity depends on the input shape to some extent, the fall-off being less rapid for the uniform aspect ratio

rear view as compared to the more elongated sideview. The correlation peaks decrease smoothly in all cases, however, maintaining a constant FWHM and without fragmentation into multiple peaks. Figure 4.9 shows a typical in-plane rotation correlation response demonstrating this behaviour.

Thus, as suggested by the simulation work described in Chapter 3, the full bandwidth correlation is too sensitive to variations of the input from the reference function to be useful for the recognition of a 3-D object unconstrained in orientation. The response is similar to that which would be obtained with an inverse filter since the read-out HeNe beam is unattenuated with frequency, the spectrum being given equal weighting up to the cut-off frequency of 14.6 cycles/mm caused by the 9 mm aperture of the BSO crystal.

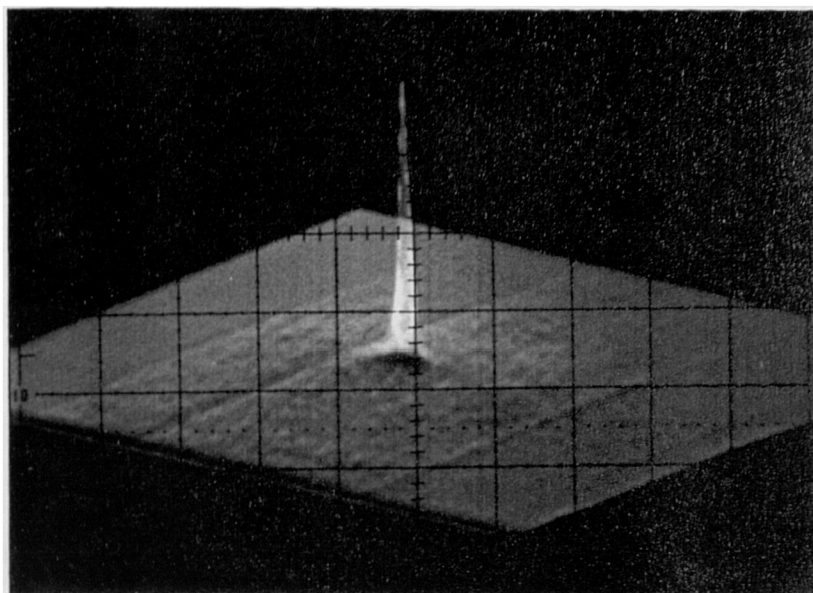


Figure 4.7(a) Isometric of full bandwidth auto-correlation response of the joint transform correlator to Escort view E090

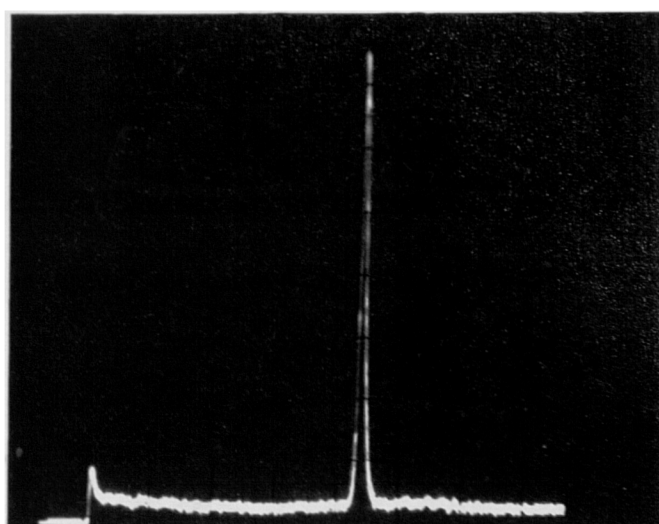


Figure 4.7(b) Linescan through full bandwidth auto-correlation of E090

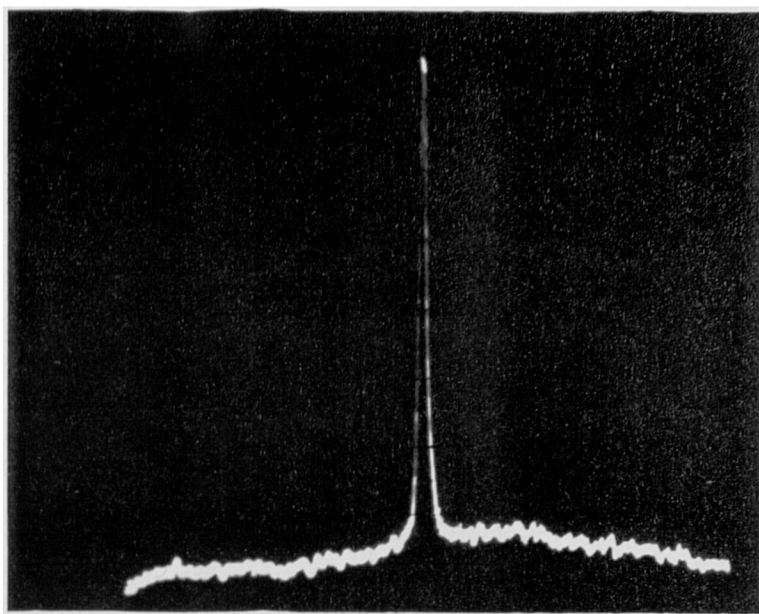


Figure 4.7(c) Linescan through full bandwidth auto-correlation of E000

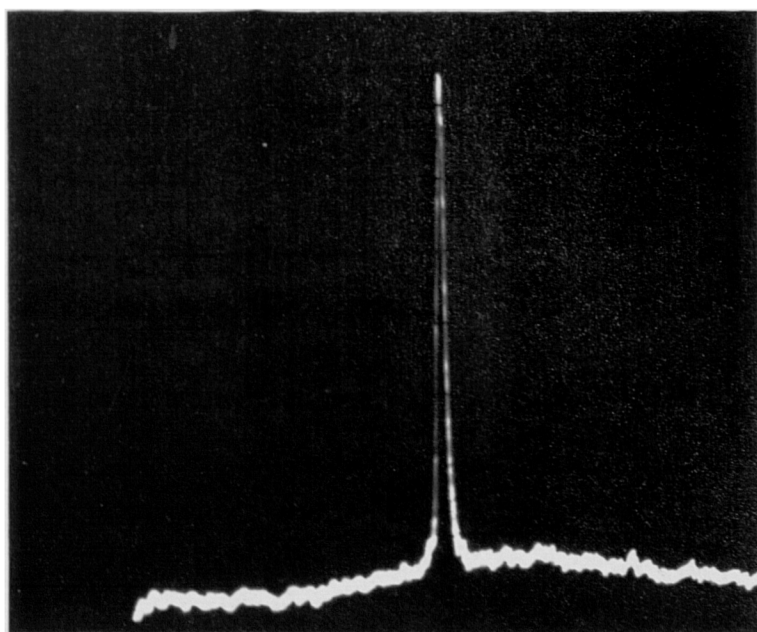


Figure 4.7(d) Linescan through full bandwidth auto-correlation of E045

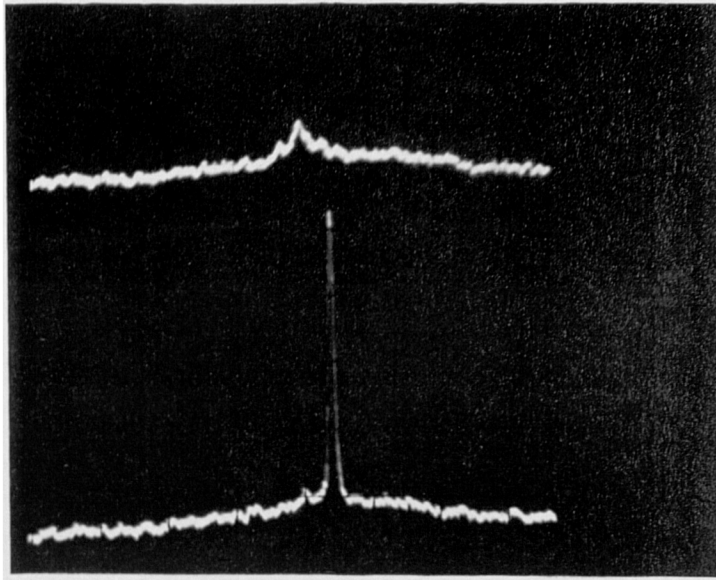


Figure 4.8(a) Full bandwidth inter-class response of the joint transform correlator - E180 cross-correlated with M180

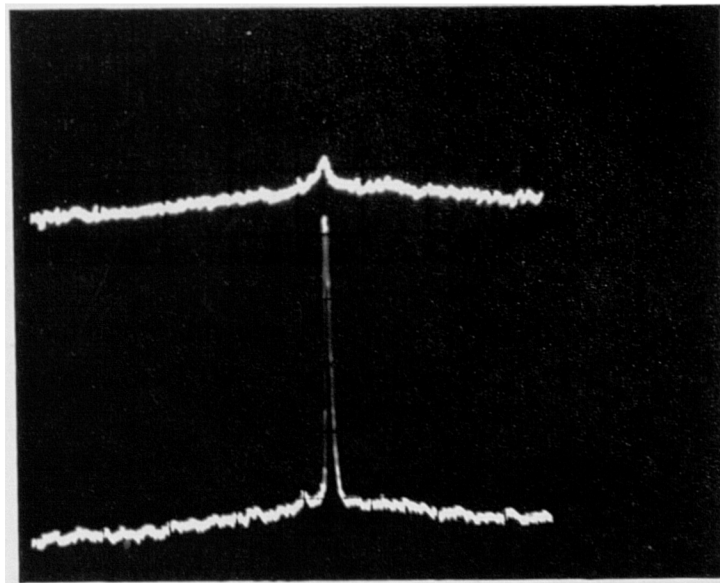


Figure 4.8(b) Full bandwidth intra-class response of the joint transform correlator - E180 cross-correlated with E011

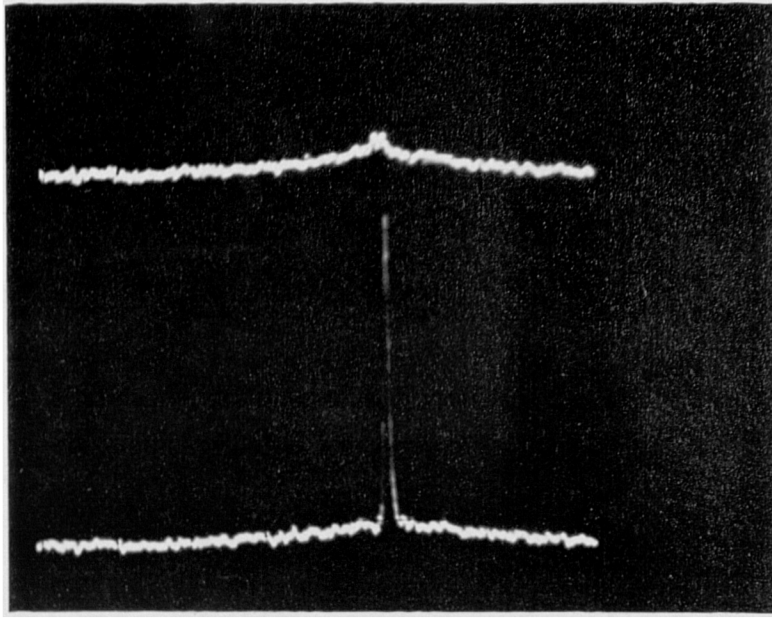


Figure 4.8(c) Full bandwidth inter-class response of the joint transform correlator - E045 cross-correlated with M045

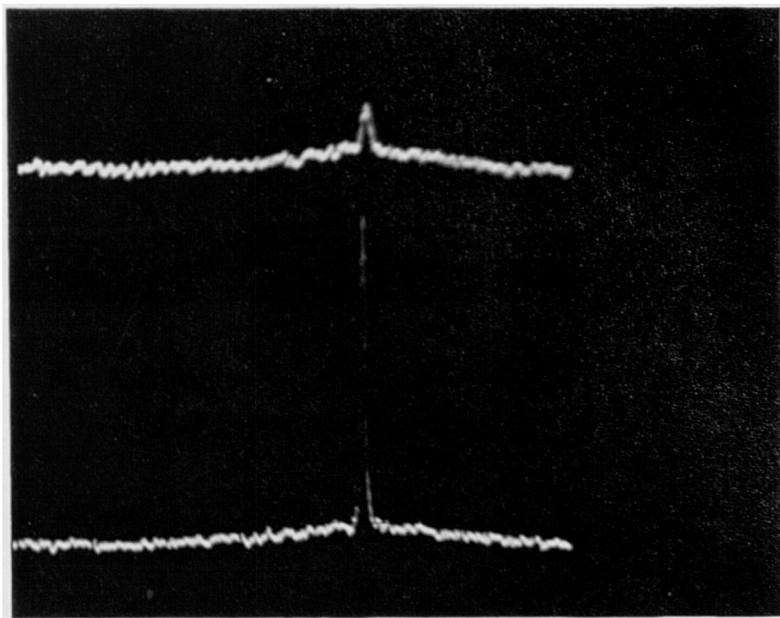


Figure 4.8(d) Full bandwidth intra-class response of the joint transform correlator - E045 cross-correlated with E056

CAR ORIENTATION	E090		E180		E000		E045		E135	
IN-PLANE ROTATION	RATIO	dB	RATIO	dB	RATIO	dB	RATIO	dB	RATIO	dB
1°	0.88	-0.56	0.90	-0.46	0.87	-0.60	0.92	-0.36	0.89	-0.51
2°	0.68	-1.67	0.87	-0.60	0.76	-1.19	0.80	-0.97	0.73	-1.37
4°	0.45	-3.47	0.67	-1.74	0.59	-2.29	0.73	-1.37	0.56	-2.52
8°	0.38	-4.20	0.46	-3.37	0.41	-3.87	0.47	-3.28	0.43	-3.67

Table 4.2 Full bandwidth joint transform correlator response to in-plane rotation of car test imagery

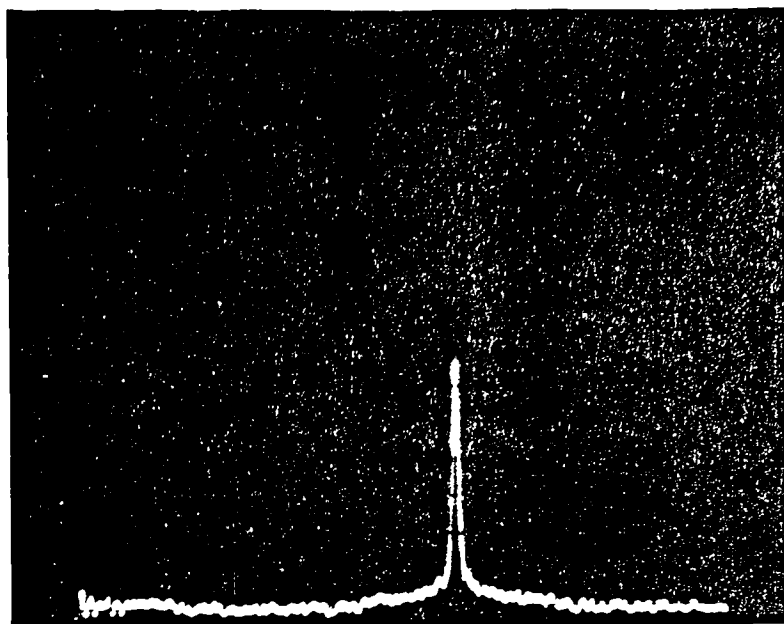


Figure 4.9 Fall in full bandwidth correlation with in-plane rotation
E090 at 4° rotation

4.4 Joint transform correlator bandpass performance

4.4.1 The difference of Gaussian filter

To improve the intra-class tolerance, while simultaneously maintaining inter-class discrimination, the frequency response of the correlator must be modified to form a bandpass. This is conveniently accomplished in the joint transform correlator configuration by exploiting the Gaussian amplitude profile of the processing wavefronts to form a difference of Gaussian (DOG) diffraction envelope from the BSO crystal. The theory of the DOG filter and the means used to implement it in the joint transform correlator are described below.

The difference of Gaussian function is a form of wavelet filter first introduced by Gabor [6]. The recent revival of interest in wavelet filtering has centred on the simultaneous space and frequency domain localisation property of these functions and applications resulting from this [7]. However, prior to this more recent work, physiologists had identified an on-centre/off-surround (and off-centre/on-surround) receptive field response within the vertebrate retina. The receptive fields were shown to have a difference of Gaussian response profile, the size of the receptive field becoming larger with eccentricity from the centre of the fovea, ie. a space variant structure with progressively lower resolution towards the periphery of the visual field. These observations were further analysed and their signal processing role in primary vision identified by Marr and Hildreth [8]. These workers demonstrated that the DOG filter is a close approximation to the Laplacian of a Gaussian. Since the Laplacian is an isotropic second differential operator, when convolved with an image it will produce zero-crossings in regions of the image in which the intensity is changing (see Figure 4.10). The location of the zero-crossings thus correspond to edge detail in the image; drawing a line through the zero-crossings results in an edge map, each element of the line being orthogonal to the direction of maximum slope of the local intensity gradient. However, direct application of the Laplacian to an image gives unsatisfactory results due to the high sensitivity to disruption of the resulting edge map by high frequency noise contained within the image [9]. The application of the Laplacian of a Gaussian to an image, $f_g(\mathbf{r})$, may be described as :

$$f_g(\mathbf{r}) = \nabla^2(\text{Gauss}(\mathbf{r}) * f(\mathbf{r})) \quad \{4.3a\}$$

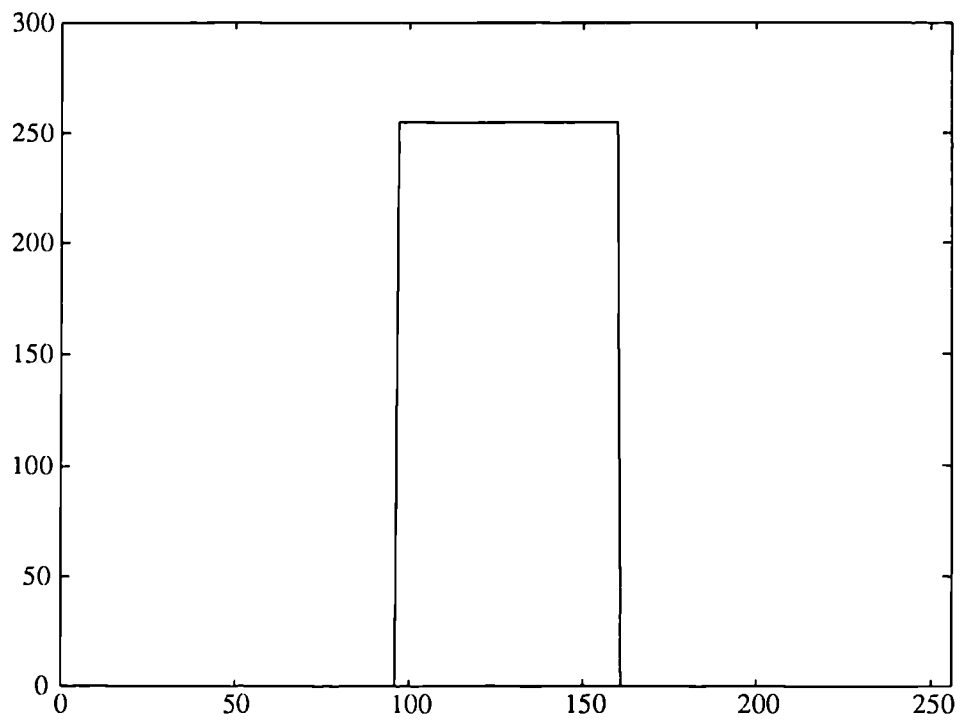


Figure 4.10(a) Positive and negative step changes in intensity

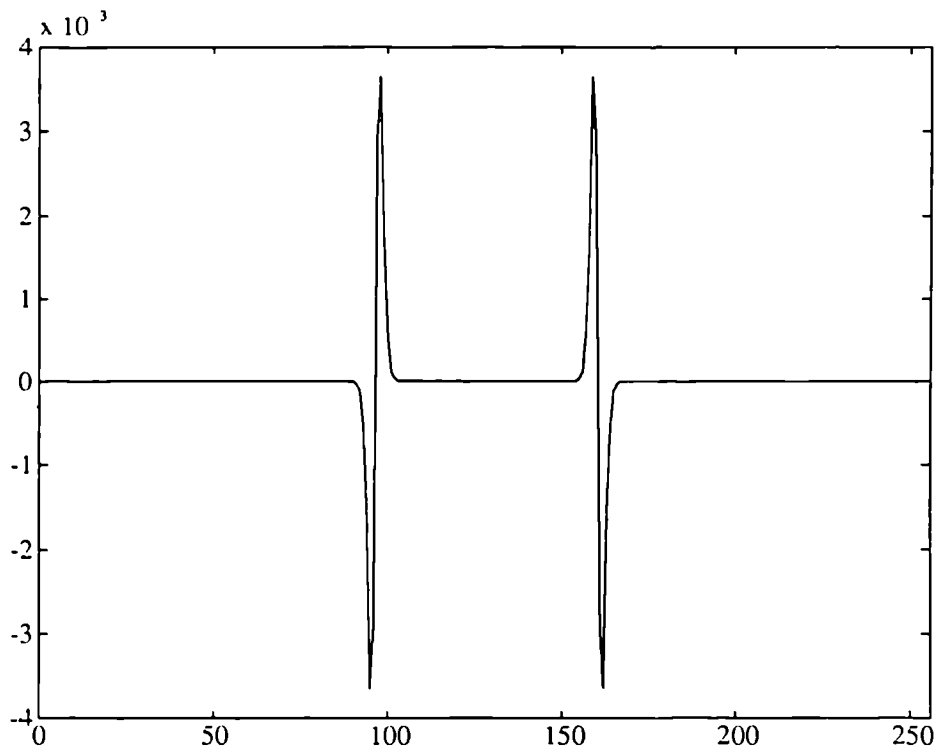


Figure 4.10(b) Zero-crossing response of the difference of Gaussian filter to step changes in intensity (maximum frequency response of filter at 46 pixel point)

which, since the derivative and convolution operations are linear, may equivalently be written :

$$f_g(\mathbf{r}) = \nabla^2(\text{Gauss}(\mathbf{r})) * f(\mathbf{r}) \quad \{4.3b\}$$

Thus the operation of LOG filtering an image is equivalent to first convolving it with a Gaussian blurring function, which acts as a low pass filter, followed by zero crossing location at that reduced resolution. By changing the size of the Gaussian function the image resolution can be altered and thus zero-crossings detected at a hierarchy of spatial resolutions within the image.

Marr and Hildreth [8] show that the ∇^2 of a Gaussian may be approximated by a difference of two Gaussian functions, the closest match occurring when the standard deviations of the inhibitory and excitatory Gaussians, σ_i and σ_e , are in the ratio of 1.6:1. The DOG filter may be written in the space domain as :

$$\text{DOG}(r) = d_e^2 \exp(-d_e^2 r^2) - d_i^2 \exp(-d_i^2 r^2) \quad \{4.4\}$$

or in the frequency domain as :

$$\text{DOG}(k) = \exp\left(-\frac{k^2}{d_e^2}\right) - \exp\left(-\frac{k^2}{d_i^2}\right) \quad \{4.5\}$$

where r and k are radial co-ordinates in the space and spatial frequency domains and:

$$d_{e,i} = \frac{1}{\sqrt{2\pi}\sigma_{e,i}}$$

The space and frequency domain forms of the function are shown in Figure 4.11(a) and (b) respectively. The frequency domain function may be thought of as a Gaussian lowpass filter multiplied by ω^2 . This results in a bandpass filter which, when $d_e/d_i = 1.6$, has a half power bandwidth of 1.3 octaves.

The effect of changing the Gaussian kernel size of the ∇^2 filter on the response to a step change (ie. infinitely steep edge) is shown in Figure 4.12. It can be seen from these plots that the localisation of the positive and negative swings either side of the zero-crossing become progressively more pronounced with reduction of the kernel size which is equivalent to a correspondingly higher bandpass filtering operation on the image. These

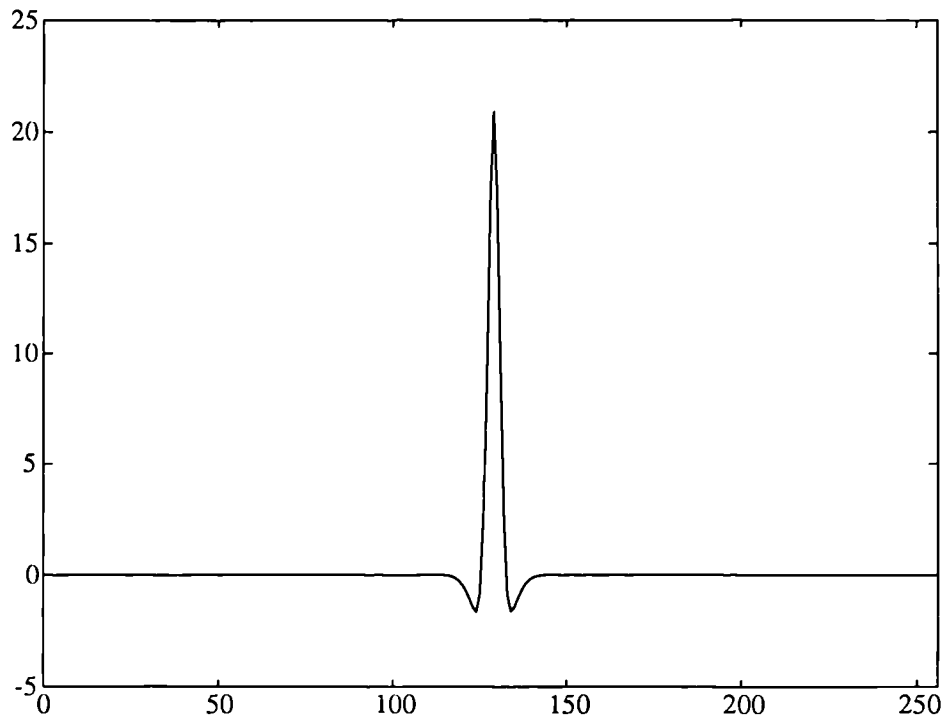


Figure 4.11(a) The difference of Gaussian function (space domain convolving function)

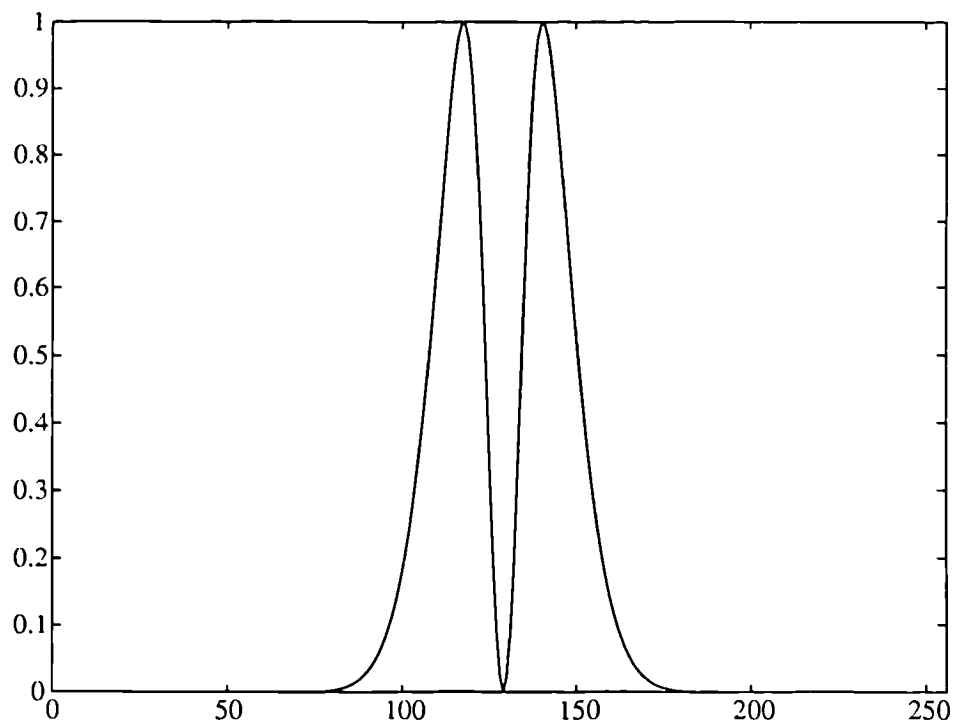


Figure 4.11(b) The difference of Gaussian function (frequency domain bandpass filter)

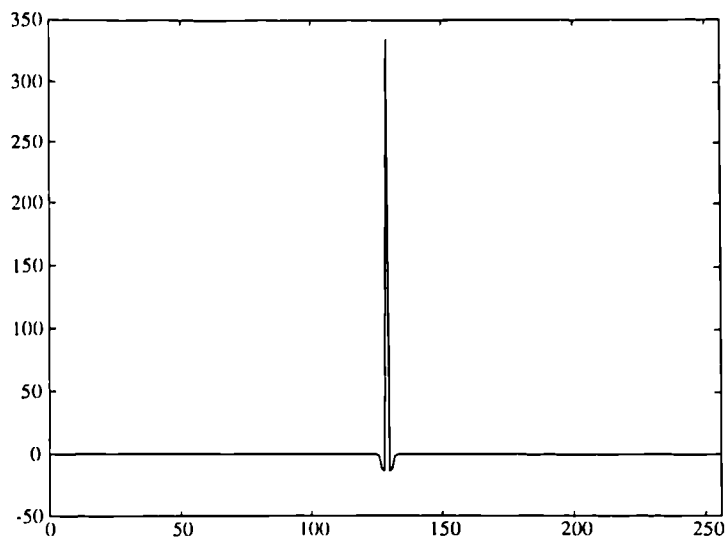


Figure 4.12(a) Space domain convolving function

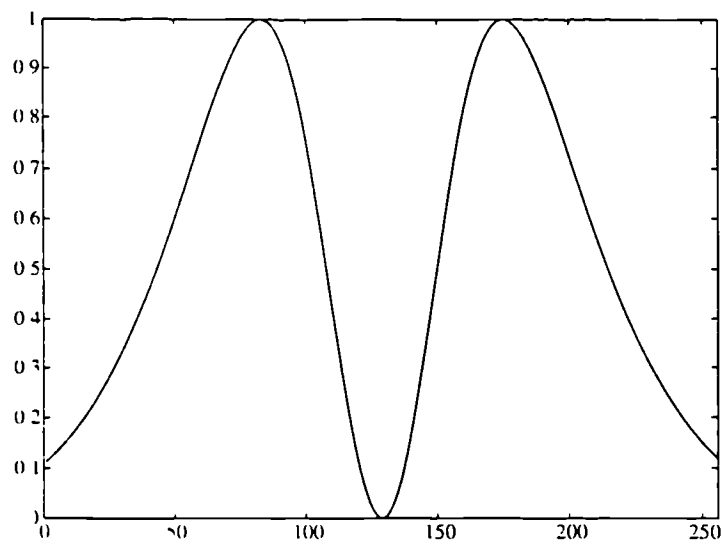


Figure 4.12(b) Frequency domain bandpass filter (max. response at 46 pixel point)

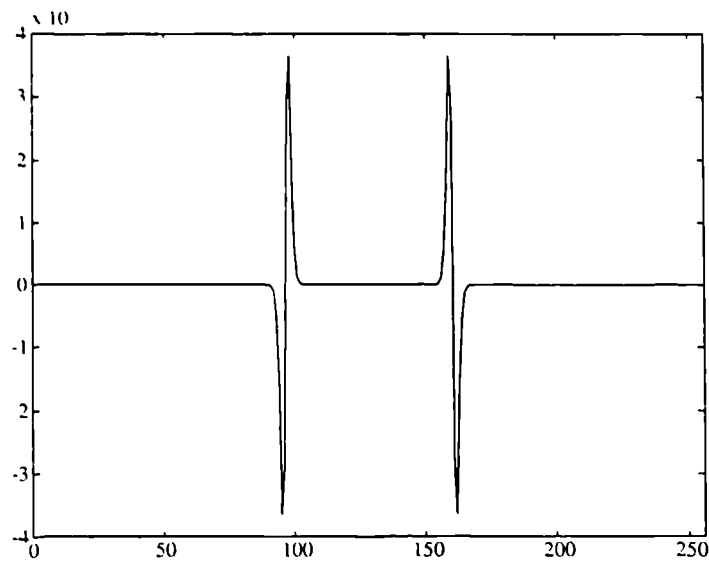


Figure 4.12(c) Response of this filter to step change in intensity

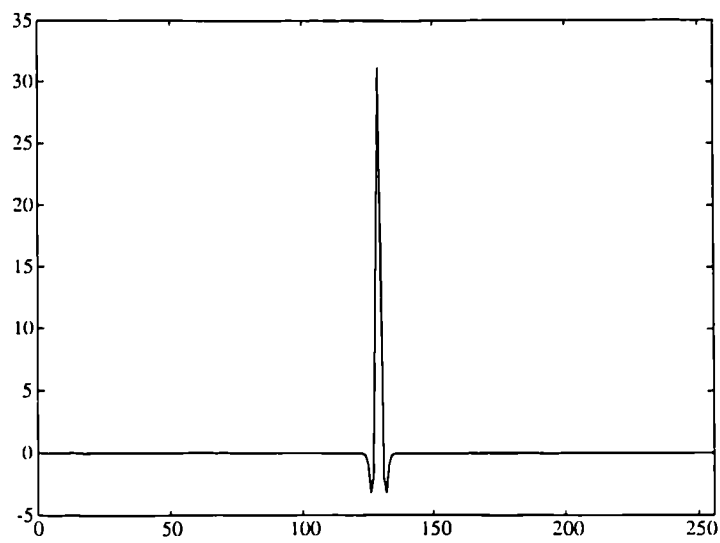


Figure 4.12(d) Space domain convolving function

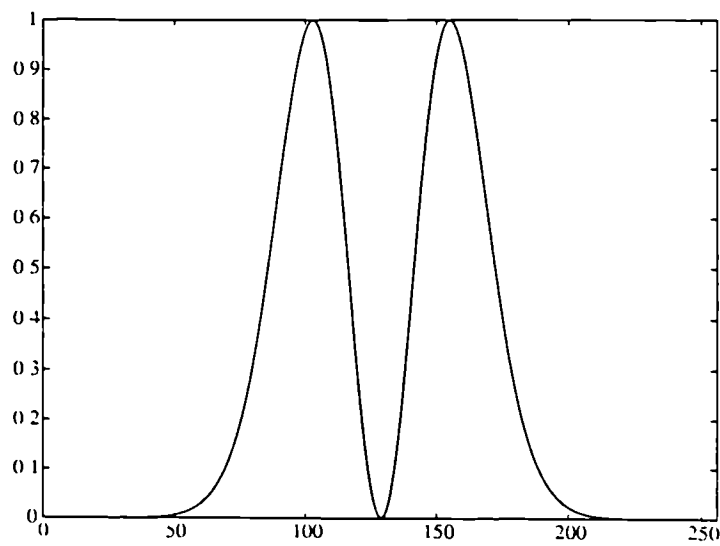


Figure 4.12(e) Frequency domain bandpass filter (max. response at 26 pixel point)

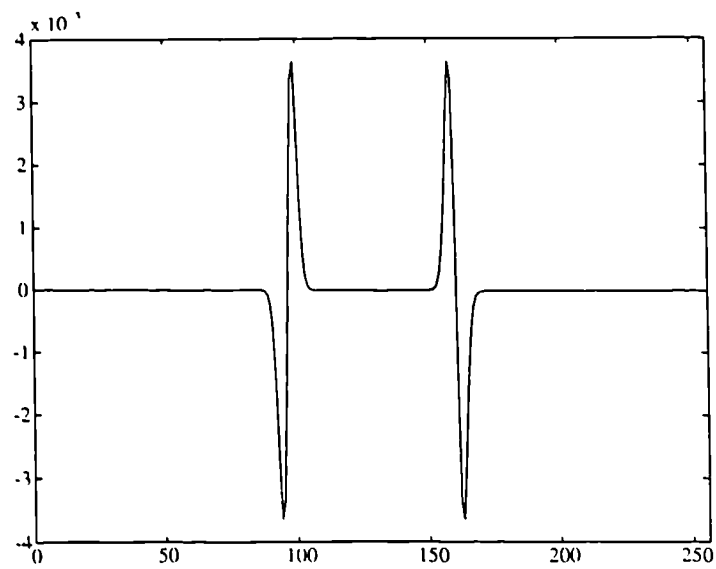


Figure 4.12(f) Response of this filter to step change in intensity

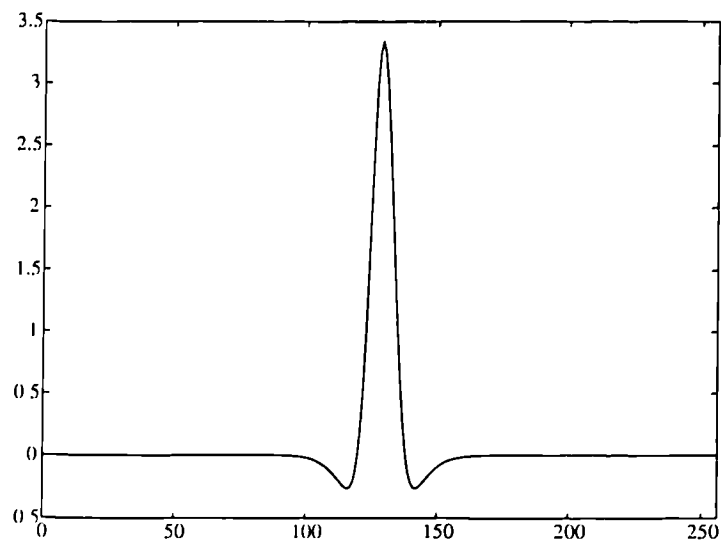


Figure 4.12(g) Space domain convolving function

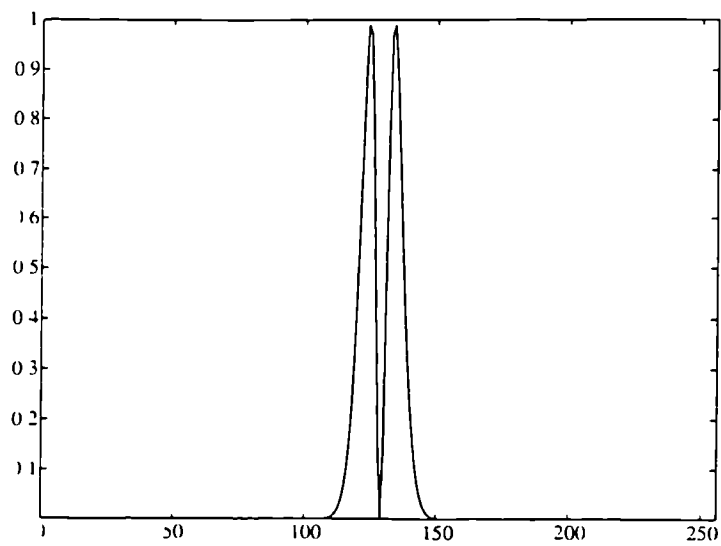


Figure 4.12(h) Frequency domain bandpass filter (max. response at 5 pixel point)

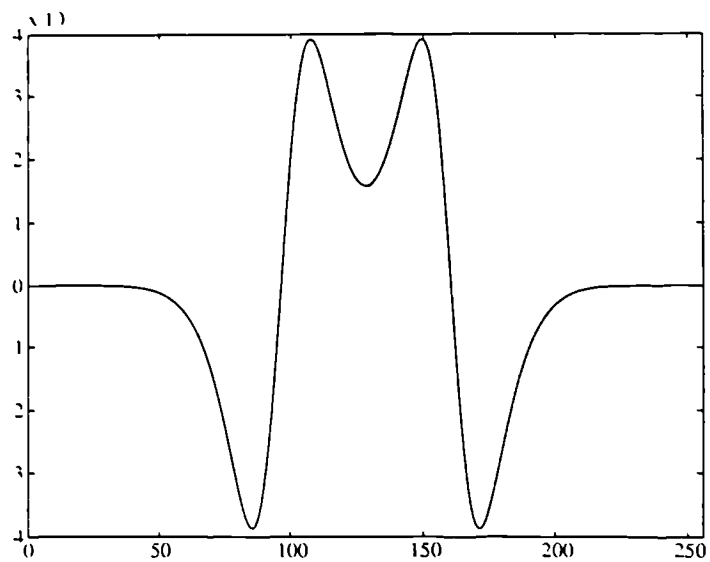


Figure 4.12(i) Response of this filter to step change in intensity

results should be compared to those shown in Figure 2.2 which result from tuning the Vander Lugt correlator to different spatial frequency bands by adjusting the reference beam strength. The similarity due to the bandpass nature of both filters is readily apparent.

Marr's proposal [8] was that the vertebrate visual system employs several spatial frequency channels each tuned to a separate band of frequencies by utilising DOG filters at different centre frequencies. The suggested purpose is to build a low resolution edge map (and from this a scene segmentation) and progressively fill in the detail obtained by the higher frequency channels. Thus the low frequency channel output would be a sparse edge map of the major regions of the image, the finer detail having been convolved out by the large kernel Gaussian of this channel. The process requires, of course, an edge map to be generated from the zero-crossing locations provided by the DOG (or LOG) filter. It must be emphasised at this point that the reason for incorporating the DOG filter bandpass in a correlator is somewhat different. The purpose in this case is to achieve good inter-class discrimination while maintaining intra-class tolerance to, in the specific case under consideration, orientation changes of a solid object. The discrimination is achieved by some degree of edge enhancement which can be perhaps more appropriately described as removal of zero and low gradient regions of the input images which do not vary significantly between intra- and inter-class data. In the frequency domain this is equivalent to the removal of the zero and low spatial frequencies. In contrast, the tolerance to intra-class distortion is achieved by spreading the remaining edge information about its original locations within the input image (obliterating altogether very fine detail) so that some degree of tolerance can be achieved to variations in the edge locations and hence to the shape of the object giving rise to them. In the frequency domain this may be thought of as a band-limiting effect which, taken together with the low frequency removal, results in the bandpass structure of the filter.

Thus, the $\nabla^2 G$ (or DOG approximation) filter may be expected to improve the inter-class discrimination as compared to a matched spatial filter and also improve the intra-class tolerance as compared to the inverse filter. Further, by adjusting the kernel size of the convolving Gaussian (or equivalently the peak frequency response of the bandpass filter) the trade-off between discrimination and tolerance may be adjusted to be optimum for a given application.

4.4.2 Experimental implementation of the difference of Gaussian filtering in the joint transform correlator

As mentioned in the introduction to this Chapter the difference of Gaussian bandpass filter may be experimentally implemented within the joint transform correlator configuration by exploiting the inherent TEM₀₀ mode of the HeNe and Ar⁺ laser beams illuminating the BSO crystal. Three methods were considered for the generation of the DOG profile :

1. The beam from the HeNe laser is divided by a beamsplitter, differentially expanded such that the Gaussian profiles have space constants d_e and d_i in the ratio 1.6 : 1, and then recombined at the output plane of an interferometer. A wedged glass plate is placed in the smaller beam to introduce a π phase shift so that it coherently subtracts from the larger beam. The DOG profile thus generated is used to illuminate the BSO crystal as the read-out beam. This technique was not pursued due to its relative complexity as compared to the methods described in (2) and (3).
2. The HeNe beam read-out beam addressing the BSO is expanded so as to be the correct size for the excitatory component of the DOG filter. The inhibitory component is realised by using a third Ar⁺ beam of Gaussian profile (which may be coherent *or* incoherent with respect to beams $U_1(\mathbf{k})$ and $U_3(\mathbf{k})$) expanded such that its space constant, d_i , is $1/1.6$ that of the HeNe beam. This beam reduces the modulation of the grating formed by beams $U_1(\mathbf{k})$ and $U_3(\mathbf{k})$ in the BSO in proportion to its local intensity provided that the maximum intensity of the beam is adjusted such that saturation of the photorefractive effect does not occur. Thus the HeNe beam diffracted from the grating will have an approximately DOG amplitude profile.
3. A HeNe beam is used to expose a photographic emulsion which is then developed as a negative. This is then placed in a HeNe beam of larger size so generating the DOG profile in the transmitted beam.

The implementation of method (3) requires a brief consideration of the exposure properties of photographic emulsions. For coherent illumination where field amplitudes are of importance the *amplitude* transmittance versus exposure curve is a more convenient representation than the logarithmic intensity transmittance versus logarithmic exposure (Hurter-Driffeld) curve more commonly used in conventional photography.

Figure 4.13(a) shows a typical t-E curve for a low contrast (low γ) emulsion such as HP 5. It can be seen from this that after an initial bias exposure, E , a longer exposure results in an approximately linear fall in amplitude transmittance with exposure until at high exposures the transmittance falls off exponentially to a small residual value. Figure 4.13(b) shows the t-E curve resulting from a very high contrast emulsion such as the Kodak orthochromatic film used to record the input imagery. It can be seen that this is linear over a much narrower range of exposures but achieves a very low residual transmittance rapidly once beyond this range (as is used to advantage in generating a high density background for the input imagery). Pan F is a fine grain emulsion striking a compromise between these two extremes, giving a linear recording over a range of exposures while achieving a moderate optical density (approximately unity for high exposures). For these reasons this emulsion was chosen to record the inhibitory Gaussian in the generation of the DOG filter. With reference to the notation shown in Figure 4.13(c) the amplitude transmittance of the exposed and developed negative may be written :

$$t_n = t_b + \beta (E - E_b) = t_b + \beta' (I - I_b) \quad \{4.6\}$$

where:

$$\beta = \frac{t_{n_{\min}} - t_b}{E_{\max} - E_b}$$

and hence:

$$\beta' = \frac{t_{n_{\min}} - t_b}{I_{\max} - I_b}$$

If $t_b \approx 1$ and $\beta \approx -1$, Equation {4.6} may be more simply written:

$$t_n \approx 1 - |\Delta U|^2 \quad \{4.7\}$$

where ΔU is the difference between the maximum and minimum field values used to expose the emulsion. Thus the film acts as a square law detector when used in the linear region of the t-E curve.

Since a Gaussian remains unchanged in functional form when squared, a film with this transfer characteristic may be used to record the negative Gaussian comprising the DOG

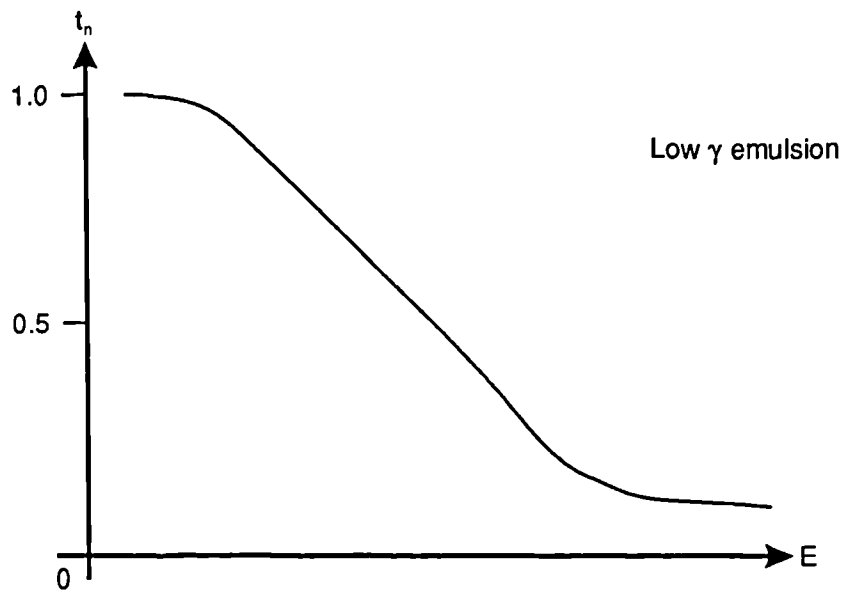


Figure 4.13(a) The amplitude transmittance-exposure curve for a low γ emulsion

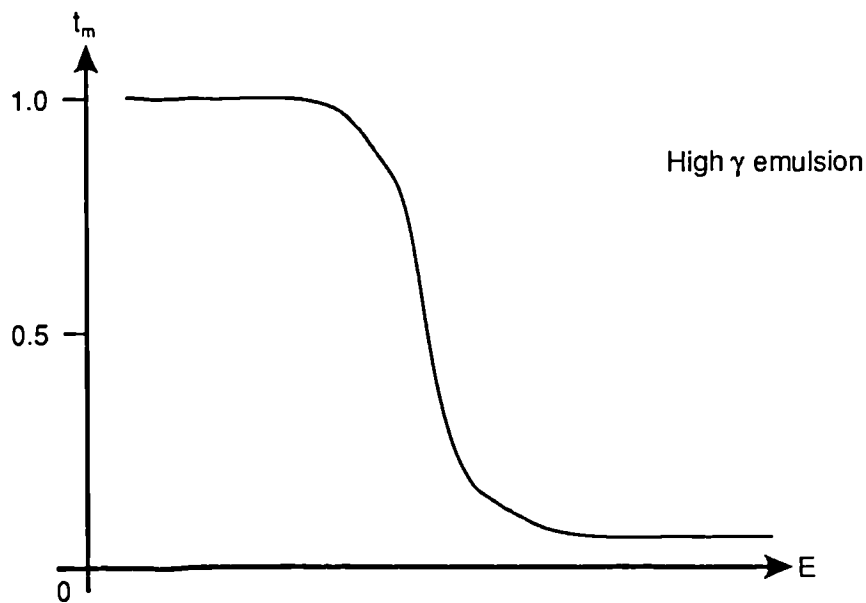


Figure 4.13(b) The amplitude transmittance-exposure curve for a high γ emulsion

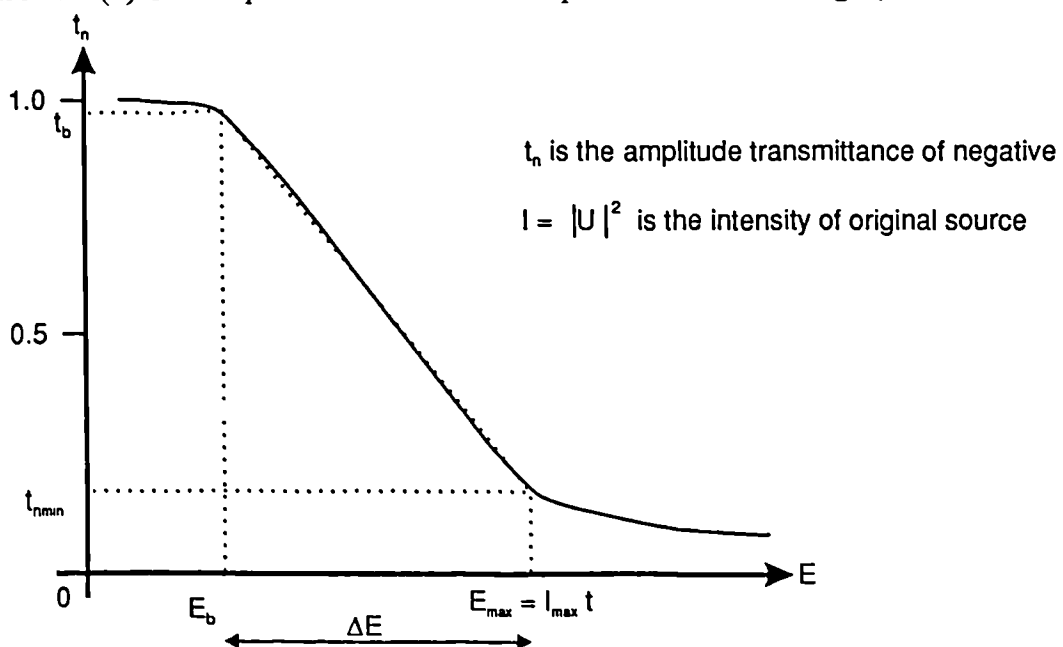


Figure 4.13(c) Notation for amplitude transmittance-exposure curve

function, provided account is taken of the square law action of the film on the resulting transmitted field amplitude profile. The beam used to initially expose the film should thus be expanded such that its $1/e$ power point is the same distance from the central intensity peak of the Gaussian as the point at which the intensity transmitted by the film when developed and uniformly illuminated has fallen by $1/e^2$ of its maximum value. When illuminated by a Gaussian beam with a $1/e^2$ power point 1.6 times this distance, the transmitted beam will have the required difference of Gaussian profile. This distribution may then be imaged onto the BSO crystal at Bragg incidence to illuminate within the required bandpass the grating formed by the non-linear interaction of the Ar^+ wavefronts, $U_1(\mathbf{k})$ and $U_3(\mathbf{k})$, within the photorefractive material.

Figure 4.14(a) shows an isometric display of the intensity distribution directly behind a DOG filter generated in this manner. Figure 4.14(b) shows a oscilloscope linescan through a radial cross-section of this distribution. It must be remembered that this measured profile is necessarily an intensity distribution and so appears to have a narrower bandwidth than the DOG field distribution (a half power bandwidth of 1.3 octaves as compared to a half amplitude bandwidth of 1.8 octaves). Examination of Figure 4.14 indicates the limitations of using photographic film to generate a DOG filter in this way. As described above, the film must have a moderate γ value so that the t-E curve is linear over a sufficiently wide dynamic range to record the Gaussian fall-off without hard-clipping the distribution. However, such a film is not capable of achieving a sufficiently high optical density to provide adequate attenuation in the central zero frequency region of the filter which should ideally have a zero transmittance. Since the maximum achievable optical density with the Pan F film employed was of the order of unity, a residual transmittance remained in the zero and low frequency regions of the filter as can be clearly seen from the cross-section shown in Figure 4.14(b). Thus the inter-class discrimination performance may be expected to be poorer than that achievable with an ideal DOG distribution. In addition, the isometric shows that the transmittance of the filter is somewhat non-uniform around the maximum within the bandpass. This occurs because this region of the filter corresponds in location to the extremity of the attenuating Gaussian function recorded on the film. In this very low exposure region the film is beyond the linear region of the t-E curve and the development of the emulsion tends to become erratic, leading to a ragged appearance of the transmitted beam. However, the robustness of the frequency plane distribution to random amplitude perturbation as described and analysed in Chapter 3 should result in this degradation not severely affecting the localisation of the correlation peak generated from this filter. The possibility exists, of course, that a high frequency phase perturbation of the transmitted wavefront is introduced by the film which for its prevention would necessitate the filter

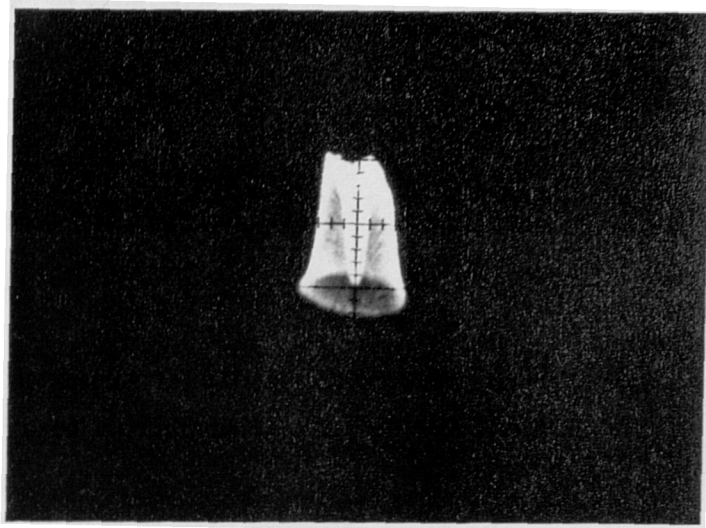


Figure 4.14(a) Isometric display of intensity distribution directly behind photographic film exposed and illuminated with Gaussian beams

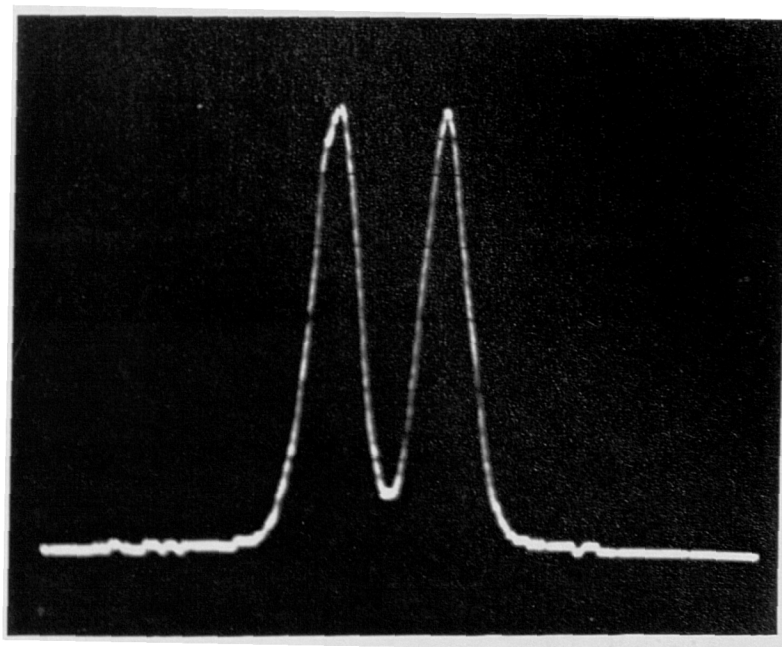


Figure 14(b) Oscilloscope linescan through intensity distribution

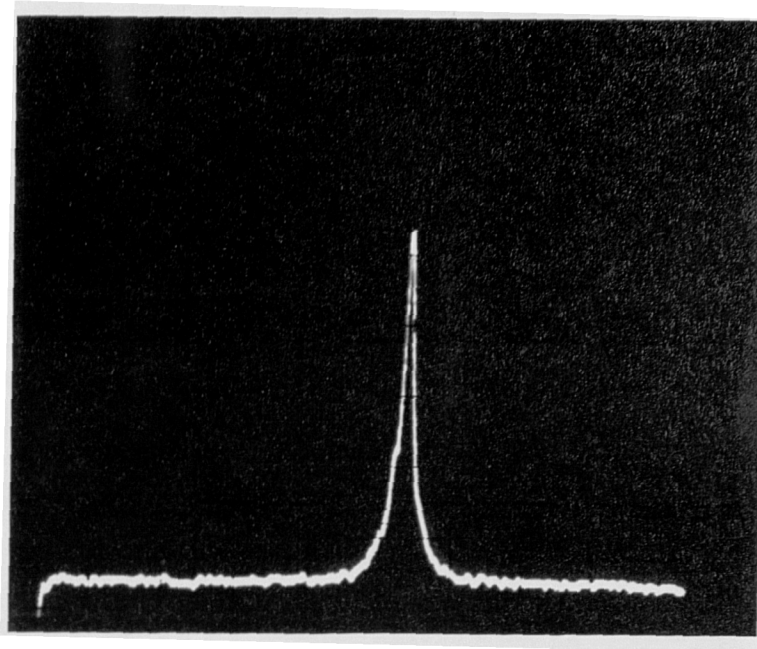


Figure 4.15(a) Impulse response of difference of Gaussian filter as implemented in the joint transform correlator

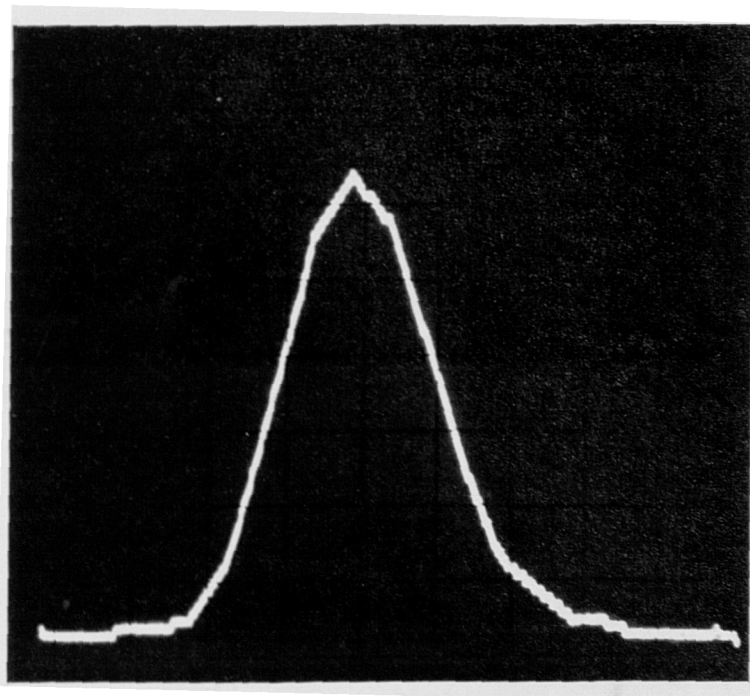


Figure 4.15(b) Expanded time-base of difference of Gaussian filter impulse response

to be placed in an index matching fluid. In practice, however, this effect does not appear to have been serious, Figure 4.15(a) showing the relatively localised impulse response of the filter when it is used in the joint transform correlator system. Figure 4.15(b) shows the oscilloscope time-base expanded to demonstrate that the impulse response has a smooth Gaussian profile (the very low sidelobes that should result from the Fourier transform of the DOG distribution being too faint to be detectable). The results presented in the next section were generated with a DOG filter implemented by the method just described.

A suitable location of the frequency bandpass generated by the difference of Gaussian filter must be determined. Previous work, most notably by Guenther *et al.* [10], in addition to simulation work and experimental trials carried out with the photorefractive based joint transform correlator constructed for the current work, suggest that the bandpass should be set to relatively low spatial frequencies. In the correlator constructed by Guenther *et al.*, (work, which although performed almost fifteen years ago, achieved some of the best results to date for an unconstrained vehicle recognition task) the reference filters were operated over the spatial frequency range of 0.4 - 4.6 lines/mm. In their experimental configuration this corresponded to the car input image being resolved to between 2 and 24 video lines on the TV monitor used to input the test scene to the correlator via the intermediacy of a Hughes liquid crystal light valve.

In the experimental joint transform correlator under examination the HeNe Gaussian output beam forming the excitatory portion of the DOG filter was adjusted to the desired size by exploiting the inherent divergence of the beam due to diffraction from the laser output window. By moving the laser on a sliding mount and adjusting its distance to the photographic film holder, Gaussian beam profiles of the correct size could be conveniently generated. The DOG field distribution at the plane of the film was then imaged onto the face of the BSO crystal at Bragg incidence to the central spatial frequency grating formed in the material by interference of the Ar⁺ input beams.

The position of the HeNe laser was adjusted such that the $1/e^2$ intensity point of the beam (corresponding to the $1/e$ amplitude point on the excitatory Gaussian forming the DOG filter) was 0.95 mm from the beam axis. For the reason discussed above, the $1/e$ intensity point of the beam used to expose the photographic film should, therefore, be $0.95/1.6 = 0.594$ mm, as this will lead to the inhibitory Gaussian having a $1/e$ amplitude point this distance from its centre. Thus, Equation {4.5} for the frequency domain representation of the DOG filter may be written:

$$\text{DOG}(r_f) = \exp\left(-\frac{r_f^2}{\lambda^2 f^2 d_e^2}\right) - \exp\left(-\frac{r_f^2}{\lambda^2 f^2 d_i^2}\right) \quad \{4.8\}$$

in which:

$$r_f = k\lambda f,$$

$$d_e = \frac{0.95}{x_f} = 3.077 \text{ mm}^{-1} \text{ and } d_i = 1.923 \text{ mm}^{-1}$$

At radial distance from the axis $r_f = \lambda f d_e$, the amplitude of the excitatory Gaussian is only $1/e^4$, ie. 1.83%, of the Gaussian maximum and only 3.34% of the DOG peak response. If this point is, therefore, taken as the frequency cut-off point for the filter, the maximum spatial frequency represented will be: $k_{\max} = r_{f_{\max}} / \lambda f$ ie. 6.15 cycles/mm (as compared to 14.58 cycles/mm when the full 9 mm aperture of the crystal is employed).

The image of the Escort sideview, which is typical of the test input imagery, is such that the distance between the car wheel base on the photographic transparency at the correlator input plane is 2.03 mm. Thus, when filtered with the DOG bandpass used in the correlator, approximately 12 cycles at the maximum represented spatial frequency will be needed in this distance or, equivalently, at least 24 data points when sampled, thus defining the required spatial bandwidth product for the reference template. The location of the peak frequency response of the DOG filter may be determined by differentiation of Equation {4.8} with respect to r_f and setting the result to zero. This yields :

$$r_{f_p} = \left[\frac{2 \log_e \left(\frac{d_e}{d_i} \right)}{\frac{d_e^2 - d_i^2}{(d_e d_i)^2}} \right]^{\frac{1}{2}} \quad \{4.9\}$$

which when $d_e/d_i = 1.6$ results in $r_{f_p} = 0.776 \lambda f d_e$ ie. 0.737 mm in the example under consideration (the peak response of the DOG filter being 0.548 of the central maximum of the excitatory Gaussian). This frequency corresponds to 2.39 cycles/mm ie. only 5 cycles within the 2.03 mm image or 10 data points when discretised.

The impulse response of the DOG bandpass filter was obtained by modulating beam 1 with a slide of the reference function and making $u_3(\mathbf{r})$ a point source (or equivalently removing the Fourier transform lens from this beam which is experimentally more

convenient) such that $U_3(\mathbf{k})$ becomes a plane wave and interferes with $U_1(\mathbf{k})$ to form a Fourier transform hologram within the BSO. The amplitude of $U_3(\mathbf{k})$ is adjusted so as to be greater than the zero order peak of the spectrum $U_1(\mathbf{k})$ in order that no biasing towards high frequencies occurs – illumination of this hologram with a uniform amplitude HeNe beam leads to a faithful reconstruction of the grey level distribution of the reference image without any edge enhancement. However, Figure 4.16 shows some example reconstructions obtained when a filter recorded in this manner in the BSO crystal is illuminated by a HeNe reconstruction beam with a DOG amplitude profile (adjusted to the frequency response discussed above ie. $r_{f_0} = 0.737\text{mm}$ at the plane of the BSO filter); these reconstructions demonstrate the effect of the DOG filter.

It will be remembered from the discussion of the DOG filter in Section 4.4.1 that its effect on the image may be considered as convolution with a Gaussian blurring function (the size of this corresponding to the Fourier transform of the excitatory Gaussian forming the bandpass filter, ie. $\exp(-d_c^2 k^2)$) followed by a second differential of the reduced resolution image. As can be seen from Figure 4.12, this results in a distribution of the remaining edge information about the zero crossing locations. The image reconstructions shown in Figure 4.16, although they are disrupted by speckle effects arising from the coherency of the HeNe read-out beam, demonstrate the effect of the DOG filter; the fine detail present in the input images has been blurred out but the lower resolution image has been edge enhanced. However, the edge information is diffused away from the zero crossing locations. The intensity should fall to zero at the zero crossing locations giving a double edge effect. The resolution of the coherently imaged reconstructions is, however, insufficient to reveal such fine detail, the double edge coalescing into a single broad line.

A problem that can arise in a DOG filtered image is that, in regions of the original image that contain acute angles or relatively small regions enclosed by a boundary, a "locking in" of the available energy occurs which results in high intensity localised peaks in these locations of the filtered image. This effect can be seen in the reconstructions shown in Figure 4.17, particularly around the wheels and wing mirrors in the front and rear views of the Escort which have been underexposed so that these high intensity regions do not saturate the film and their magnitude relative to the other regions of the image can be emphasised. This effect has the undesirable tendency to enhance small features in the input object which, although improving the discriminatory ability of the filter, will have a deleterious effect on tolerance to orientation changes of the object. If the filtered reference was input to the system as an image, a thresholding operation could be applied to the DOG filter output prior to its cross-correlation in order to overcome this effect,

although it is not apparent how this could be achieved with the direct frequency domain implementation being considered here.

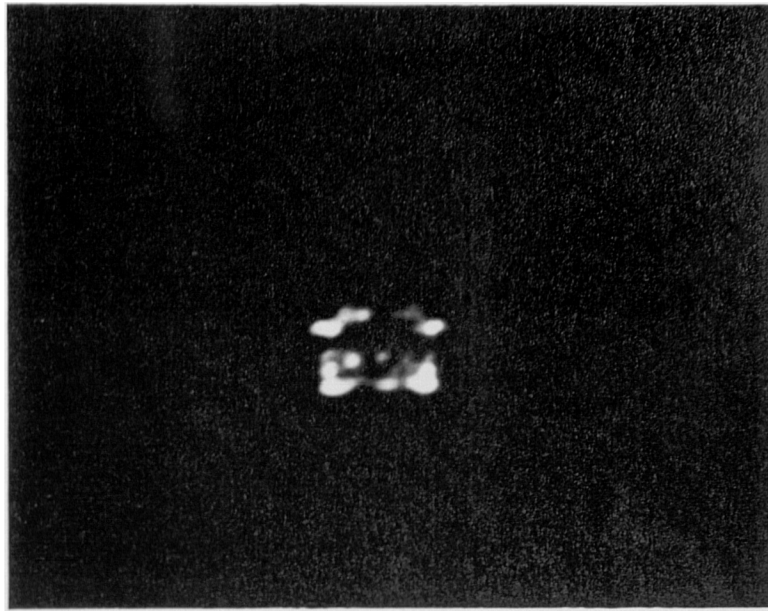


Figure 4.16(a) Image reconstruction of difference of Gaussian filtered car test imagery
Escort view E000

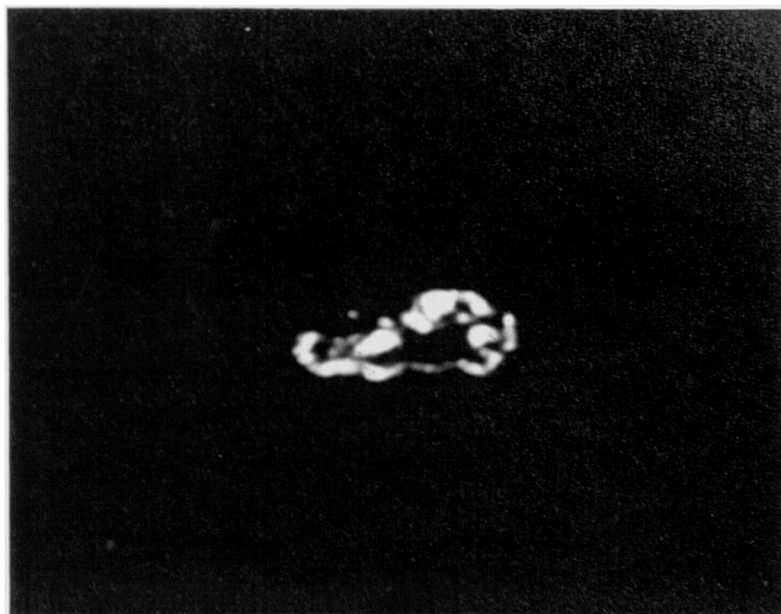


Figure 4.16(b) Image reconstruction of difference of Gaussian filtered car test imagery
Escort view E045

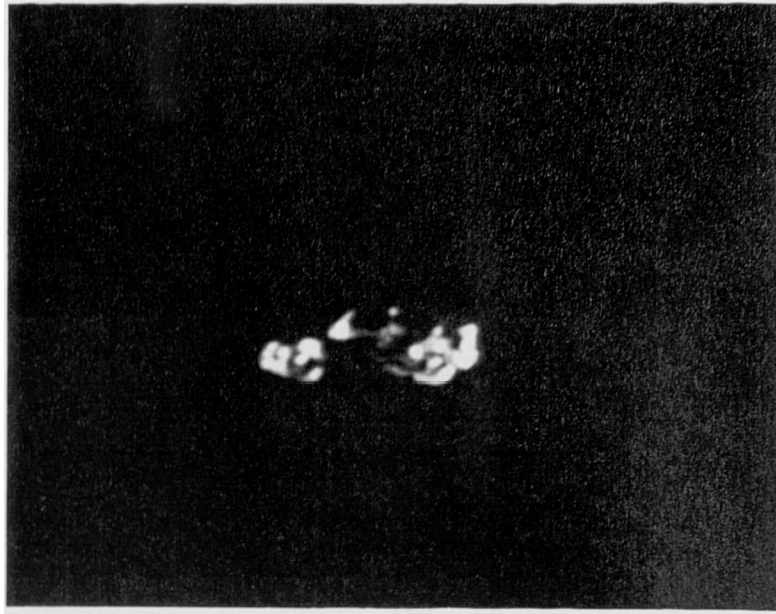


Figure 4.16(c) Image reconstruction of difference of Gaussian filtered car test imagery
Escort view E090

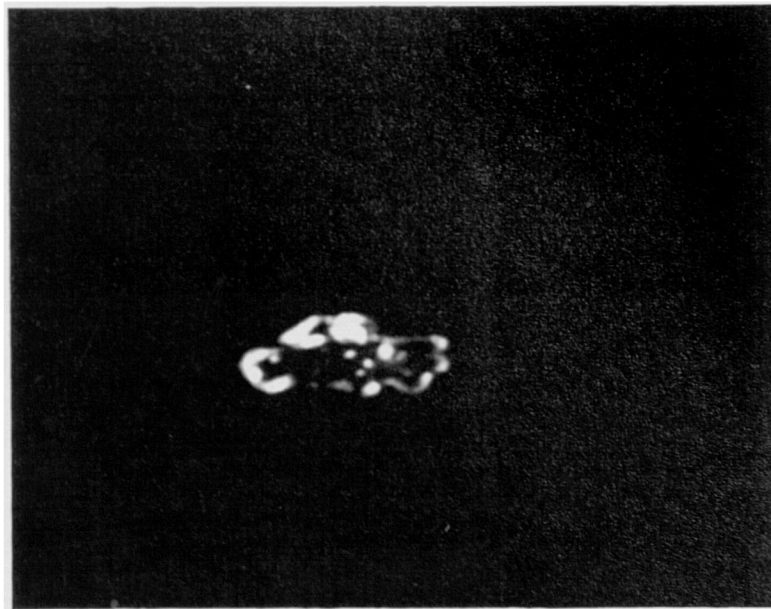


Figure 4.16(d) Image reconstruction of difference of Gaussian filtered car test imagery
Escort view E135

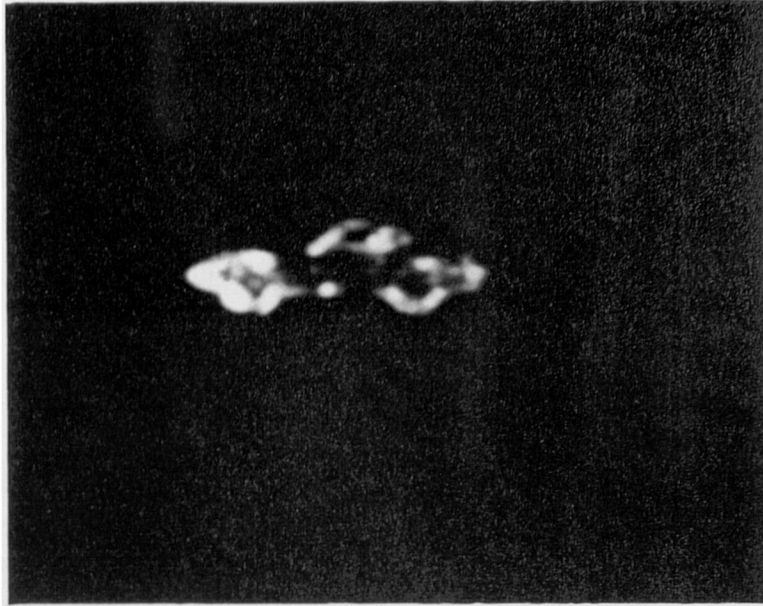


Figure 4.16(e) Image reconstruction of difference of Gaussian filtered car test imagery
Mazda view M090

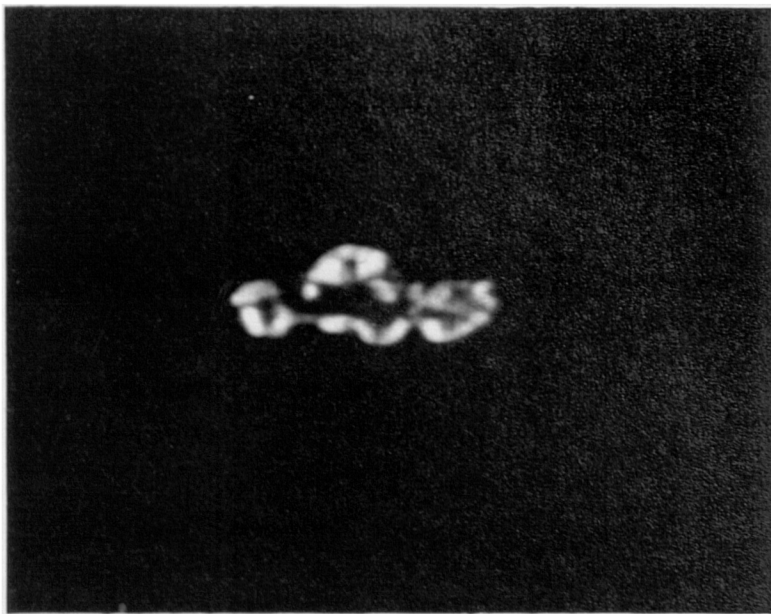


Figure 4.16(f) Image reconstruction of difference of Gaussian filtered car test imagery
Mazda view M135

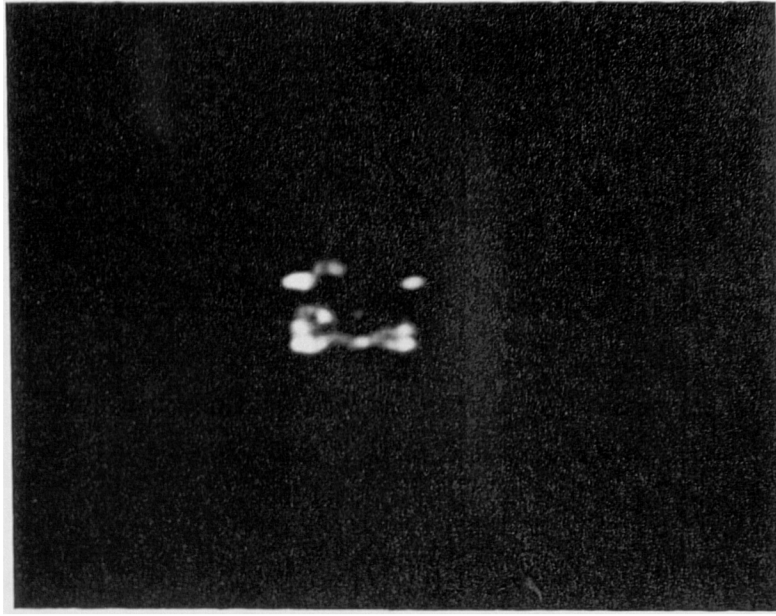


Figure 4.17(a) Under-exposed image reconstruction to demonstrate energy localisation effect of difference of Gaussian filter. Escort view E000.

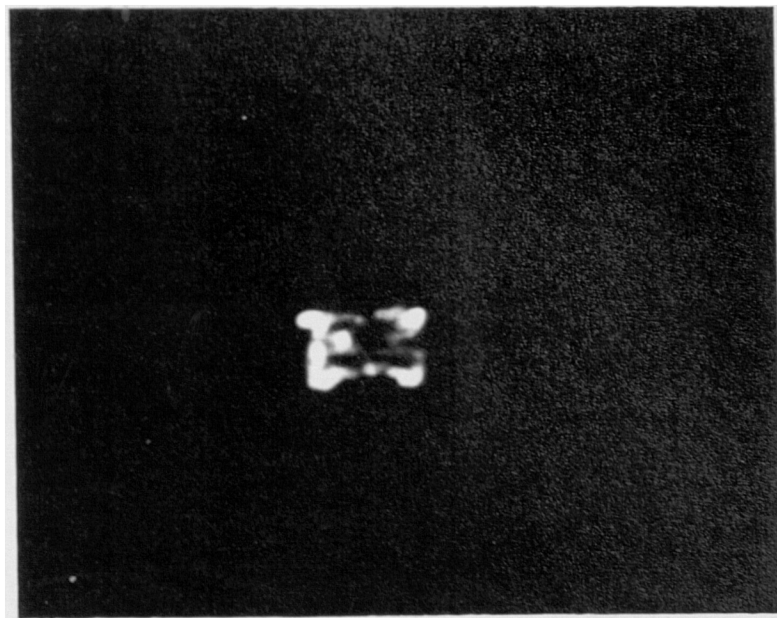


Figure 4.17(b) Under-exposed image reconstruction to demonstrate energy localisation effect of difference of Gaussian filter. Escort view E180.

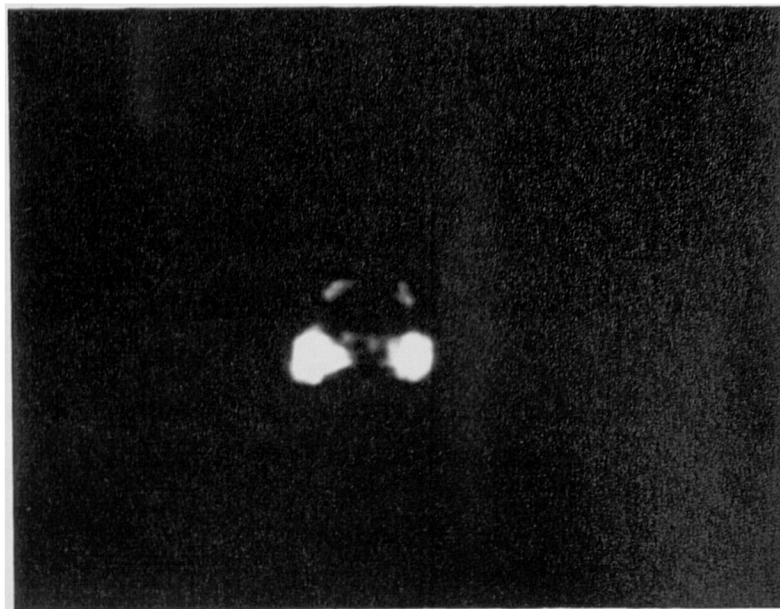


Figure 4.17(c) Under-exposed image reconstruction to demonstrate energy localisation effect of difference of Gaussian filter. Escort view M000.



Figure 4.17(d) Under-exposed image reconstruction to demonstrate energy localisation effect of difference of Gaussian filter. Madza view M045.

4.5 Intra- and inter-class performance of the bandpass joint transform correlator

Figure 4.18 shows an isometric of the auto-correlation response of the joint transform correlator when operated in the bandpass generated by DOG filtering the HeNe read-out beam in the manner detailed in the previous Section, the peak frequency response of the filter being set to 2.39 cycles/mm. Figure 4.19(a) shows an oscilloscope linescan through this auto-correlation function. The following parameters of the modulus squared of the space domain DOG function may be readily deduced from Equation {4.4} :

1. The full width at half maximum (FWHM) of the central peak can be obtained by an iterative solution to the equation :

$$\left(\frac{d_e}{d_i}\right)^2 \exp(-d_e^2 r_h^2) - \exp(-d_i^2 r_h^2) = \left[\frac{1}{2} + \frac{1}{2} \left(\frac{d_e}{d_i}\right)^4 - \left(\frac{d_e}{d_i}\right)^2 \right]^{\frac{1}{2}} \quad \{4.10\}$$

where:

$$\Delta r_{FWHM} = 2r_h$$

2. The full width at the base of the central peak ie. the distance between the first zeroes of the function is :

$$\Delta r_o = 2 \left[\frac{2 \log_e \left(\frac{d_e}{d_i} \right)}{(d_e^2 - d_i^2)} \right]^{\frac{1}{2}} \quad \{4.11\}$$

3. The full width between the maxima of the sidelobes is :

$$\Delta r_{SL} = 2 \left[\frac{4 \log_e \left(\frac{d_e}{d_i} \right)}{(d_e^2 - d_i^2)} \right]^{\frac{1}{2}} \quad \{4.12\}$$

4. The ratio of the sidelobe maximum to the central peak height is :

$$SLPR = \left[\frac{d_e^2 \exp(-d_e^2 r_{SL}^2) - d_i^2 \exp(-d_i^2 r_{SL}^2)}{(d_e^2 - d_i^2)} \right] \quad \{4.13\}$$

Due to the wavelength shift of the readout HeNe beam from that of the Ar⁺ input beams, d_e and d_i change by a proportionate amount, leading to a larger space domain DOG function; the modified values are $d_e = 2.502 \text{ mm}^{-1}$ and $d_i = 1.564 \text{ mm}^{-1}$. The detected ACF will have the functional form which is the modulus squared of a space domain DOG function, the size of the parameters of which may be calculated from Equations {4.9}-{4.11}. These predict : FWHM=0.32 mm, $\Delta r_o = 0.99 \text{ mm}$ and $\Delta r_{SL} = 1.40 \text{ mm}$. For the ratio $d_e : d_i = 1.6 : 1$, Equation {4.12} indicates the sidelobe will only be 1.37% of the peak intensity. Examination of the ACF cross-section shown in Figure 4.19(a) does not reveal a sidelobe structure but the base does spread noticeably before reaching the noise floor. Since the sidelobe is a relatively weak structure it is not distinguishable as a separate maximum in the experimental trace but probably does lead to the relatively large "footprint" of the ACF peak. This is confirmed by measurement of the base width which is 1.91 mm, ie. wider than Δr_{SL} , suggesting that the sidelobe has been coalesced with the main peak. The measured FWHM is 0.38 mm, or 19% larger than the predicted value; notice also the slight disruption to the smooth contour of the peak on its left hand side. Considering the relatively approximate way in which the frequency domain DOG filter was produced, the localised ACF peak is further evidence of the robustness of the 2-D Fourier transform to random frequency plane amplitude perturbations. In addition, the phase noise introduced by the photographic negative must have been relatively slight since this certainly would have led to a strong disruption of the ACF had it been severe.

Before discussion of the performance of the bandpass correlator on the car test imagery, a reminder of the notation used to designate the orientation of the car views is given. The letter E refers to the Escort and the letter M to the Mazda. The three numerals following the letter indicate the rotation of the car in degrees about an axis perpendicular to the viewing direction ie. the viewing angle is at zero elevation (resulting in the greatest change of the car appearance for a given change of angle). Thus a front view is designated E/M000, a sideview E/M090 and a rear view E/M180 etc. A 180° orientation change is sufficient to assess the correlator performance since the cars are bilaterally symmetric.

The auto-correlation response of the bandpass correlator to view E090 is shown in Figure 4.18 and Figure 4.19(a). The broadening of the peak, as compared to the full bandwidth response, as has been demonstrated above to be consistent with the frequency plane apodisation. Figure 4.19(b) shows the cross-correlation response to the Mazda

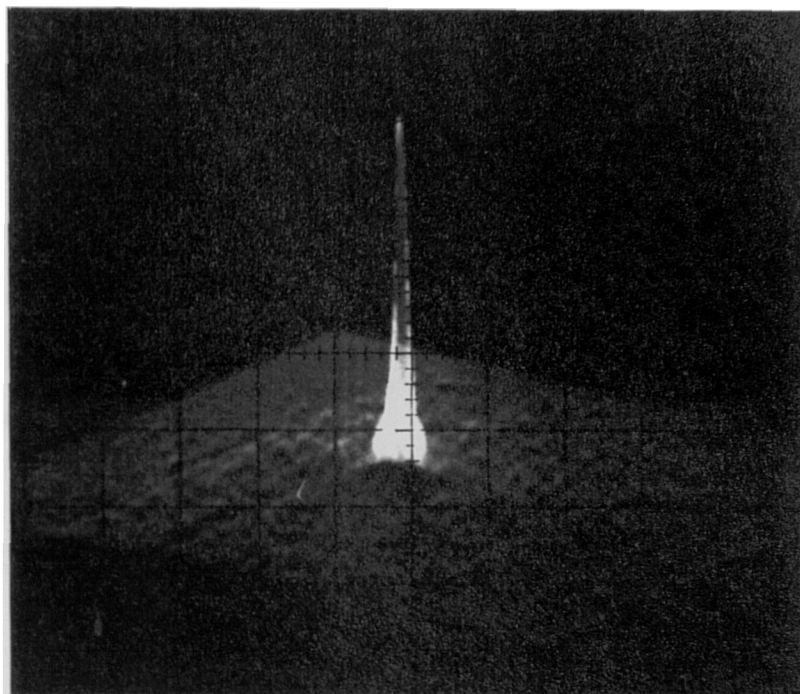


Figure 4.18 Auto-correlation response of bandpass joint transform correlator to Escort view E090

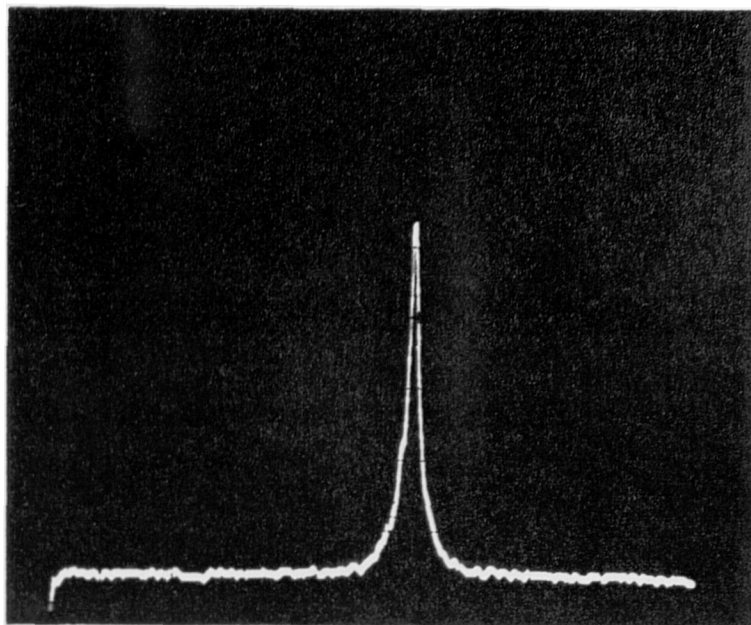


Figure 4.19(a) Auto-correlation response of bandpass joint transform correlator to Escort view E090

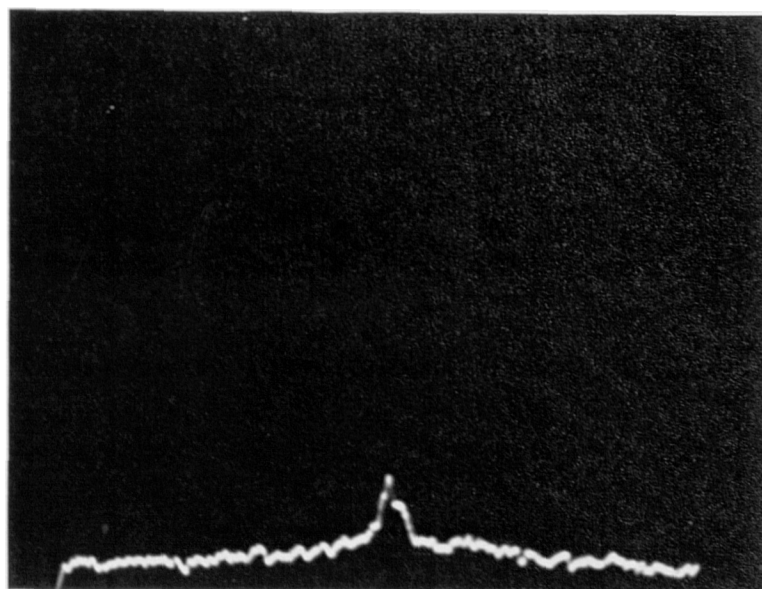


Figure 4.19(b) Cross-correlation response of bandpass joint transform correlator between Escort view E090 and Mazda view M090

side view (M090). Although a residual correlation peak is present, it is 0.2 (ie. -7dB) of the auto-correlation peak. This should be compared to the full bandwidth inter-class response of between 0.009 and 0.15 of the auto-correlation peak (see Figures 4.8(a) and 4.8(c) which show very little improvement in discrimination capability despite the higher bandwidth of the filter in these examples).

In Figures 4.20 to 4.27 the inter-class discrimination performance of the bandpass correlator is assessed on a sample of intermediate orientations of the Escort and Mazda cars over the 180° range, from front to rear views. In these figures (excepting Figure 4.23 which shows isometric views of the auto- and cross-correlation responses to inputs E/M045) the (a) and (b) parts show line scans of the auto- and cross-correlation without the inhibitory Gaussian in place. The (c) part indicates the improvement to the inter-class discrimination that can be achieved by removing the low frequency response with a difference of Gaussian profiled read-out beam. Although this results in a decrease of the available read-out beam energy by four times (and hence a corresponding reduction the absolute peak intensity of the auto-correlation function), this is unimportant provided sufficient beam power is available to saturate the output detector; it is the relative intensities of the ACF and inter-class response that is of overriding interest. The inter-class discrimination ability of the correlator when employing a DOG bandpass filter is quantified in Table 4.3 where comparisons are also made with the response without low frequency removal. In two of the example orientations, E180 and E045, there is more than a six-fold increase in the ratios of the auto- and cross-correlation signals when low frequencies are removed. As can be seen from Figures 4.20 and 4.22, in both these cases an initially very high inter-class correlation signal has been almost totally suppressed by the bandpass filter, leading to the large improvement in discrimination shown in Table 4.3. In four of the remaining examples an improvement in discrimination of approximately two and a half times results when the inhibitory Gaussian is placed in the read-out beam. In three of these cases, as can be seen in Figures 4.21, 4.24 and 4.25, an initially high inter-class response has been substantially suppressed but a residual broad and fractionated cross-correlation signal remains which, combined with the reduced auto-correlation strength, leads to a smaller increase in discrimination than achieved in the first two examples. The remaining orientation showing this degree of improvement is view E124. In this case, as can be seen from Figure 4.27, the inter-class response was initially much better than in the preceding examples. However, the relatively small and localised cross-correlation has been almost totally suppressed by low frequency removal, leading to a four fold increase in discrimination ratio for this view. In the remaining two example orientations, E090 and E079, the discrimination prior to low frequency was good, with a very low cross-correlation response, resulting in very little further improvement being

possible. It is interesting to note that these two views correspond to a sideview and its nearest angular increment. The Escort and Mazda cars are very different in shape when viewed from the side and so an intrinsically low cross-correlation may have been anticipated. In contrast, in other orientations, particularly from a direct front or rear view, the low spatial frequency profiles of the two cars are quite similar. Thus, at the reduced bandwidth being employed, the high cross-correlations initially obtained are to be expected. It is thus necessary in these cases to remove the very low frequencies so that some degree of edge enhancement occurs, resulting in an enhancement of the discriminating features at the reduced resolution being employed. As has been detailed above, the DOG bandpass filter has this effect on the image, and the examples shown in Figures 4.20-4.27 and Table 4.3 suggest that it is capable of increasing the inter-class discrimination performance of a low bandwidth correlator for input images of the type considered.

Figures 4.28-4.32 show the intra-class cross-correlation responses of the bandpass joint transform correlator with views changed by $11\frac{1}{4}^\circ$ and $22\frac{1}{2}^\circ$ on either side of the indicated Escort orientation (except for the front and rear views in which an orientation change in one direction only is required due to the bilateral symmetry of the car). These results are summarised in Table 4.4 from which it can be seen that the intra-class cross-correlation peak height remains greater than 60% of the ACF peak for an $11\frac{1}{4}^\circ$ orientation increment between the reference image and the input for all the examples examined (which cover the whole range of possible orientations of the car at zero elevation angle). As pointed out in Chapter 3 the ACF will always be higher than the intra-class cross-correlation due to the perfect frequency plane phase matching that occurs for this condition. If the intensity of the correlation peak is arranged so as to saturate the output CCD at 60% of the ACF maximum, this variation can be negated prior to digitisation of the signal, thus alleviating the dynamic range requirements of the output detector.

In contrast to the smaller increments, at a $22\frac{1}{2}^\circ$ orientation differential, the cross-correlation response has in most cases halved again, the residual peak being between 28% and 40% of the ACF. As can be seen from the oscilloscope line scans, the cross-correlation peak remains localised without break-up into multiple peaks (with the single exception of the cross-correlation of E045 with E067 shown in Figure 4.31(c)). Comparison with the inter-class results reveals that the worst case orientation is the front view, E000. However, the intra-class cross-correlation is 2.3 times the inter-class response even for this case (the highest ratio being that for E045 which is 5.6). Despite the relatively low spatial frequencies being employed, a variation in orientation sensitivity

occurs between the front/rear views and to the other orientations due to the more rapid change in appearance of the cars with alteration of angle in these two cases.

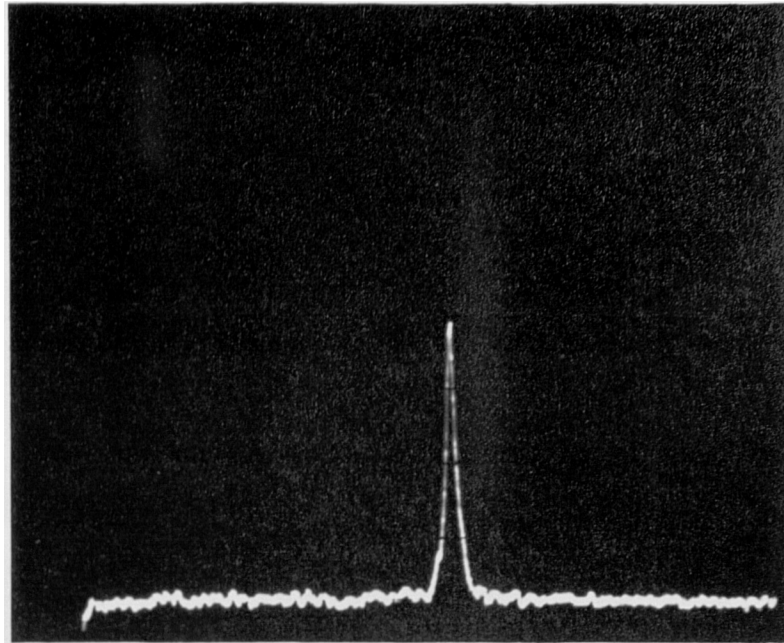


Figure 4.20(a) Auto-correlation response of bandlimited joint transform correlator to Escort view E180

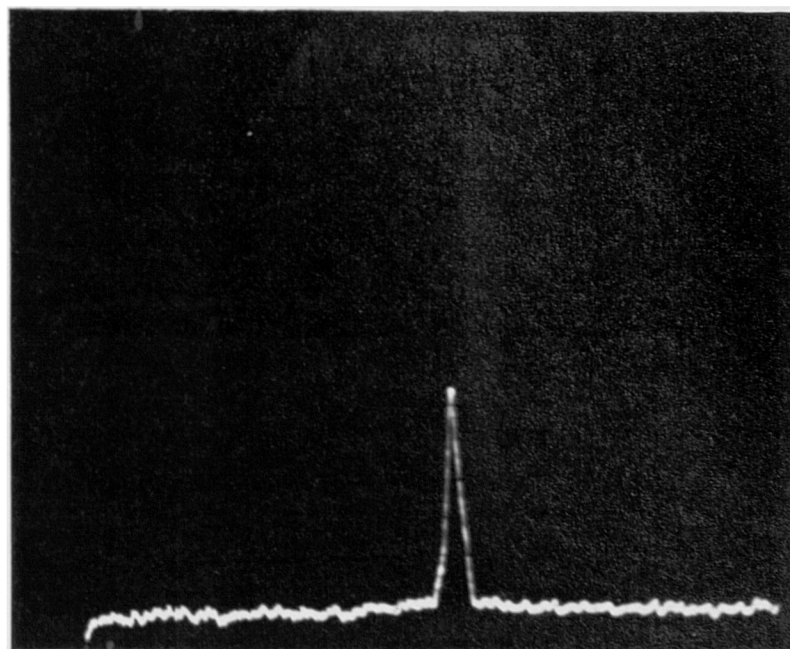


Figure 4.20(b) Cross-correlation response of bandlimited joint transform correlator between Escort view E180 and Mazda view M180

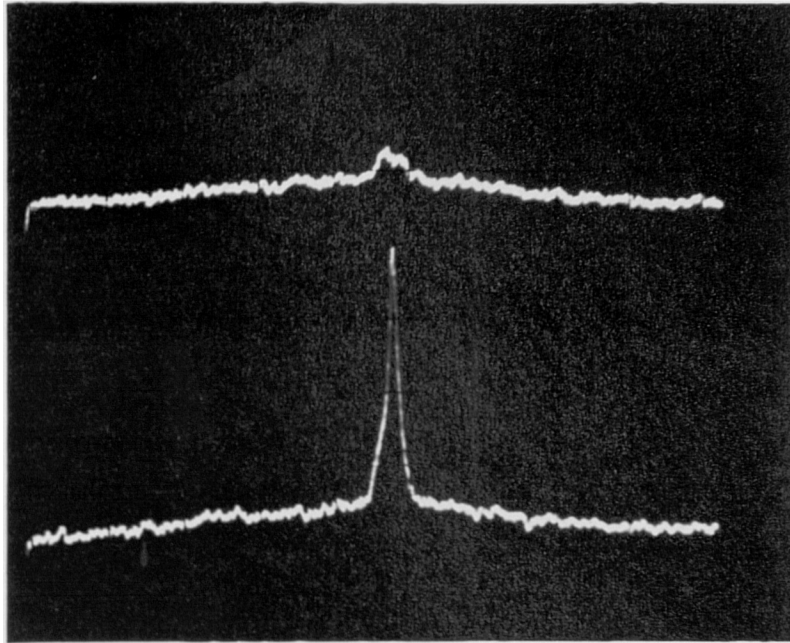


Figure 4.20(c) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E180 and Mazda view M180

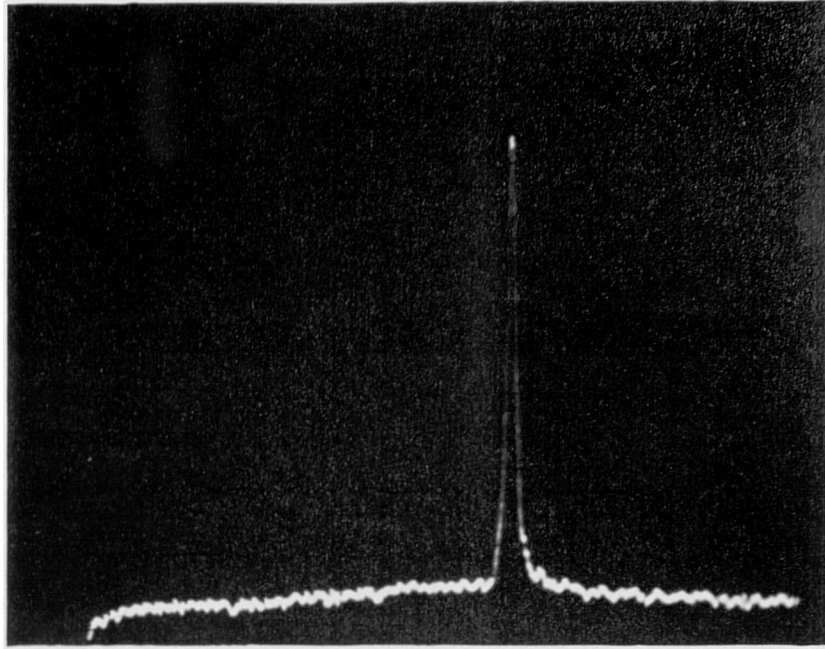


Figure 4.21(a) Auto-correlation response of bandlimited joint transform correlator to Escort view E000

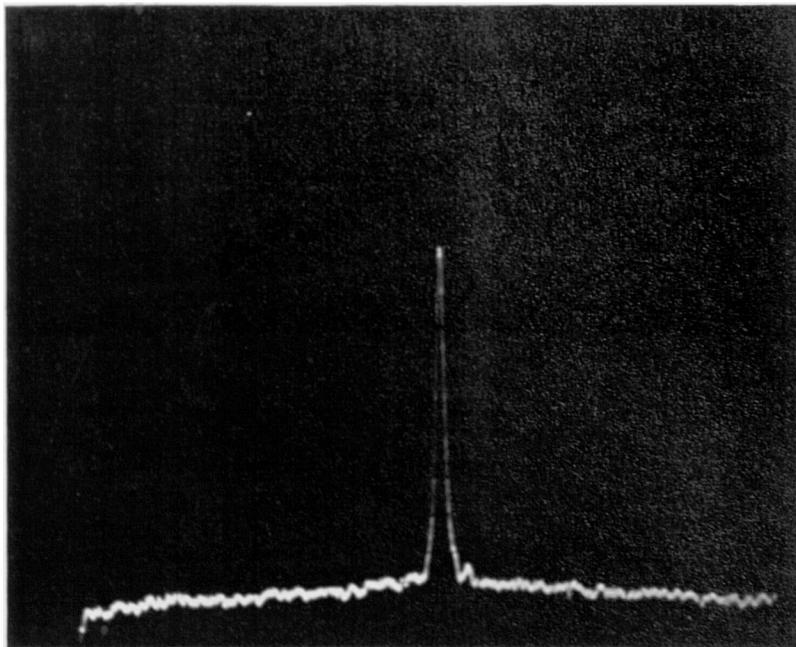


Figure 4.21(b) Cross-correlation response of bandlimited joint transform correlator between Escort view E000 and Mazda view M000

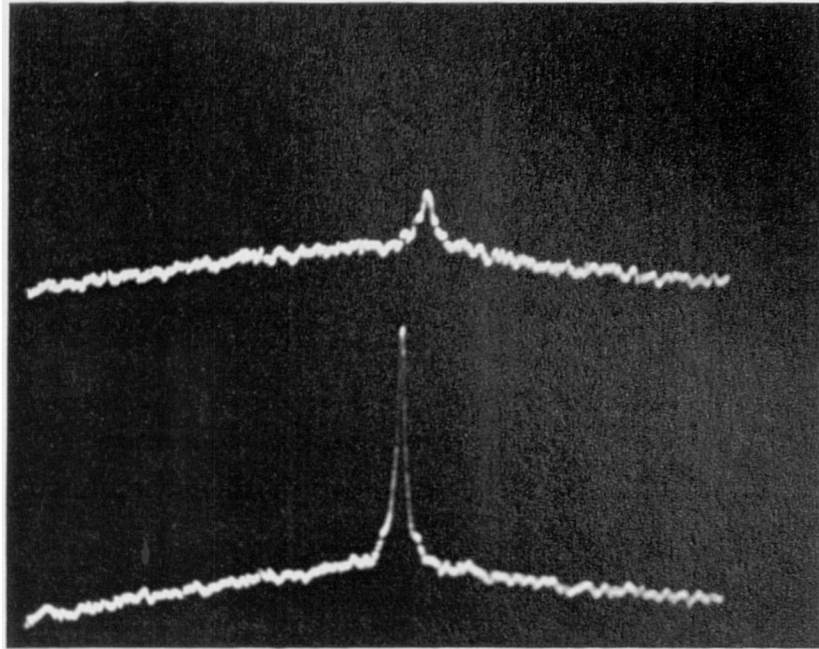


Figure 4.21(c) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E000 and Mazda view M000

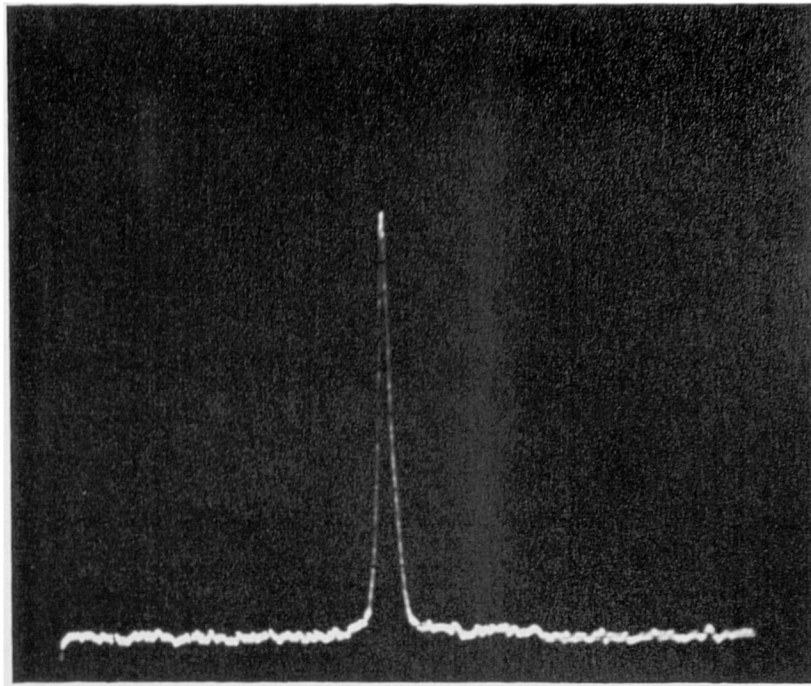


Figure 4.22(a) Auto-correlation response of bandlimited joint transform correlator to Escort view E045

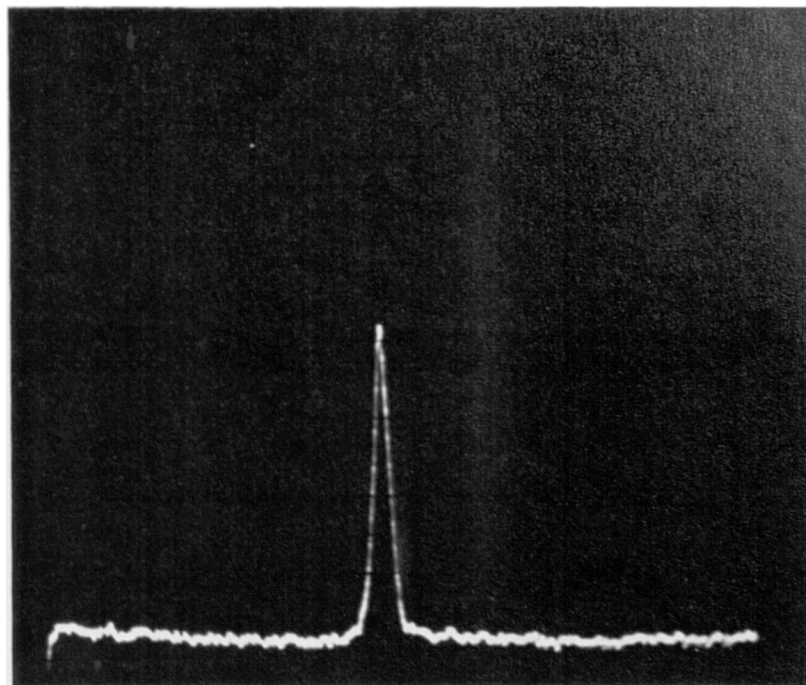


Figure 4.22(b) Cross-correlation response of bandlimited joint transform correlator between Escort view E045 and Mazda view M045

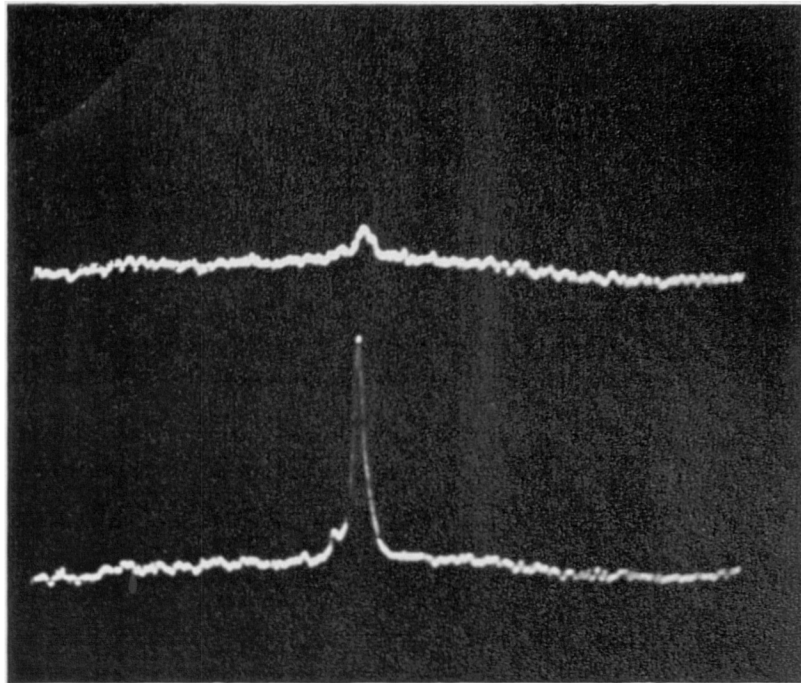


Figure 4.22(c) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E045 and Mazda view M045

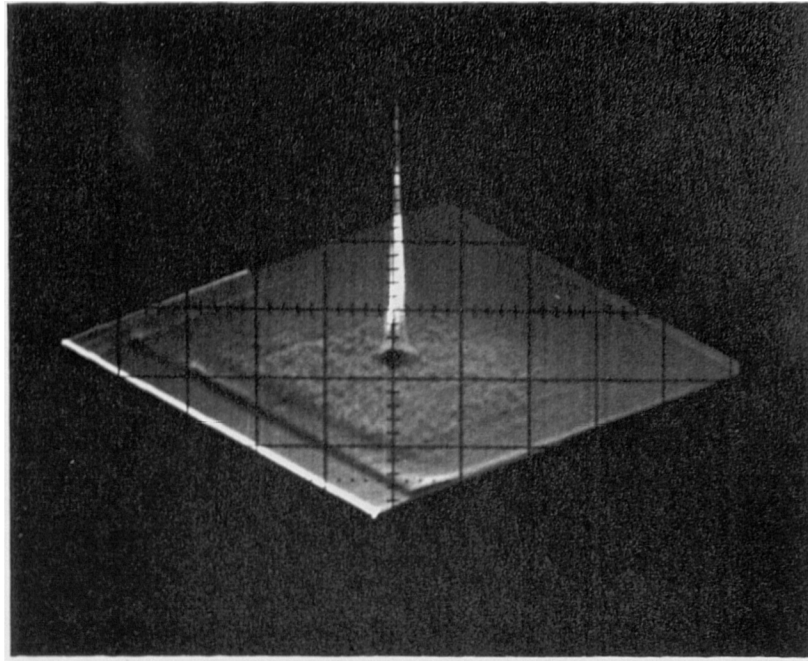


Figure 4.23(a) Isometric of auto-correlation response of bandlimited joint transform correlator to Escort view E045

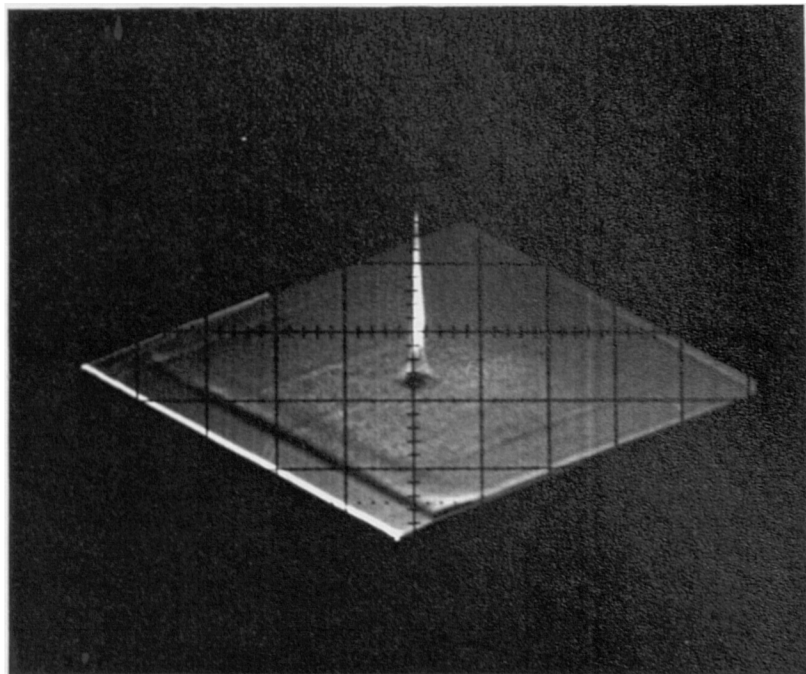


Figure 4.23(b) Isometric of cross-correlation response of bandlimited joint transform correlator between Escort view E045 and Mazda view M045

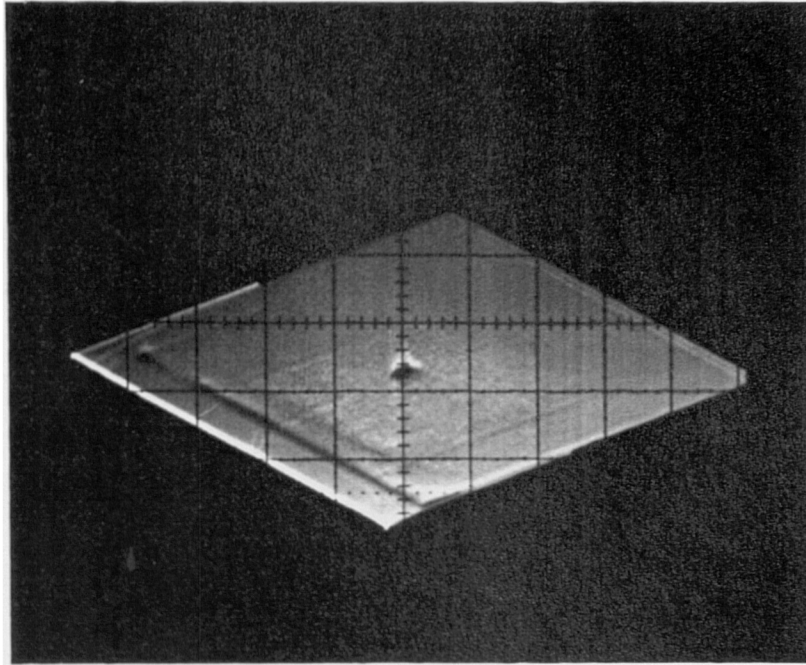


Figure 4.23(c) Isometric of cross-correlation response of bandpass joint transform correlator between Escort view E045 and Mazda view M045

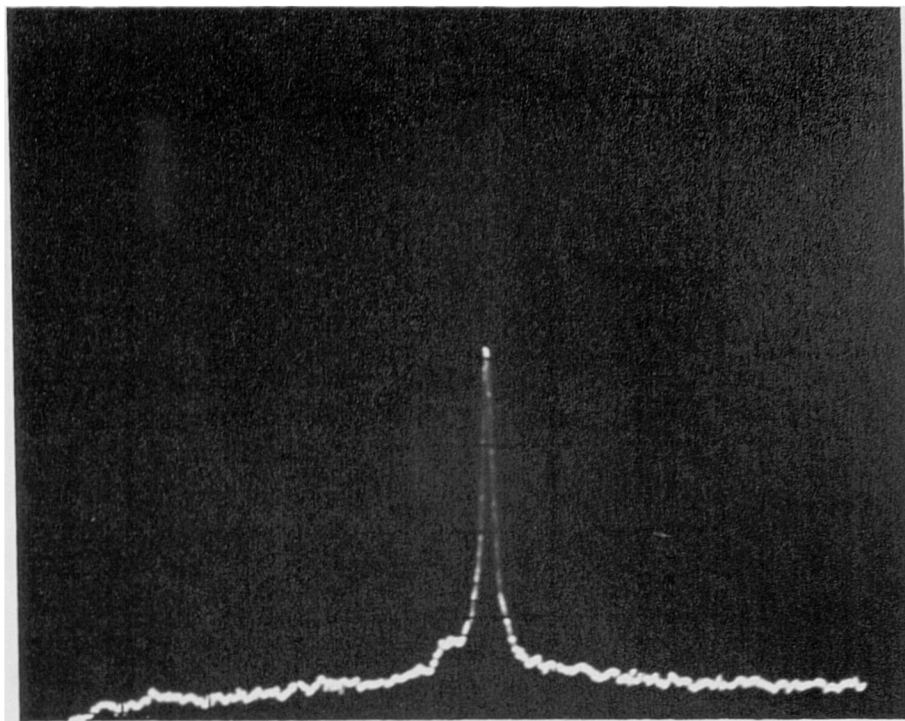


Figure 4.24(a) Auto-correlation response of bandlimited joint transform correlator to Escort view E135

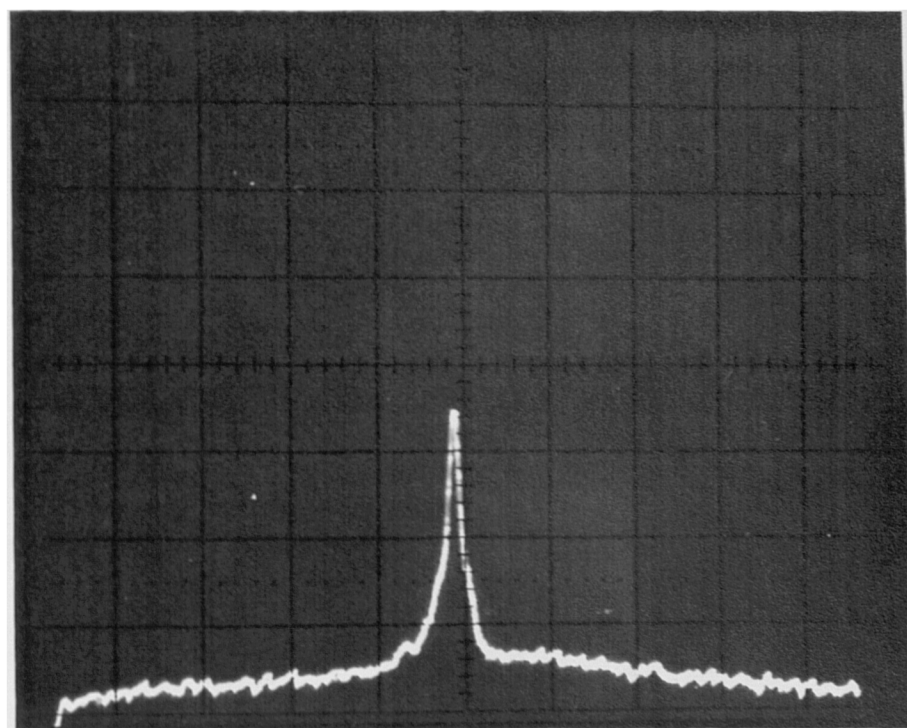


Figure 4.24(b) Cross-correlation response of bandlimited joint transform correlator between Escort view E135 and Mazda view M135

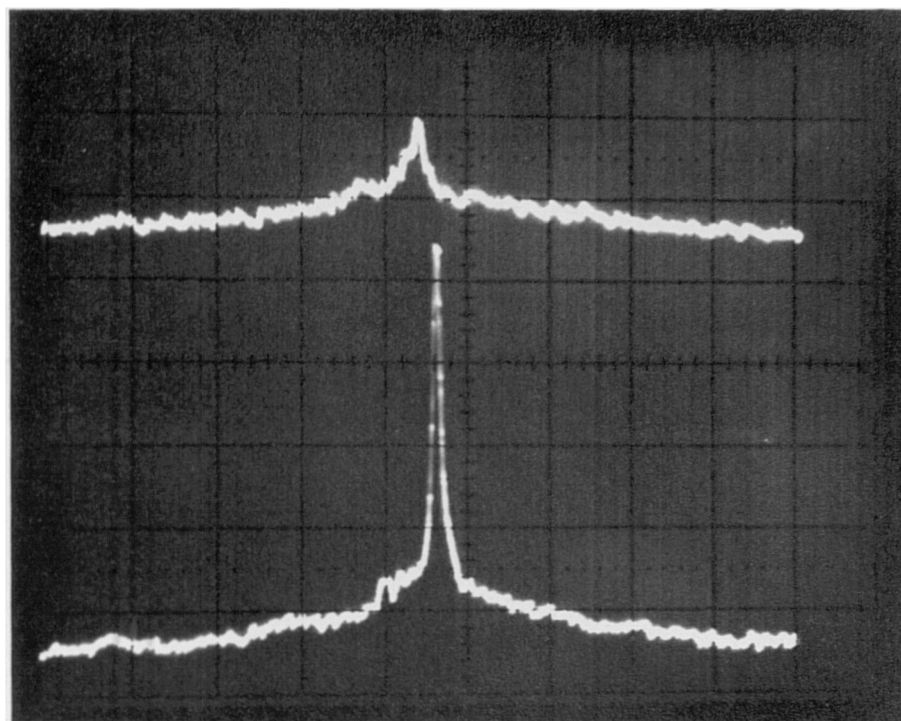


Figure 4.24(c) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E135 and Mazda view M135

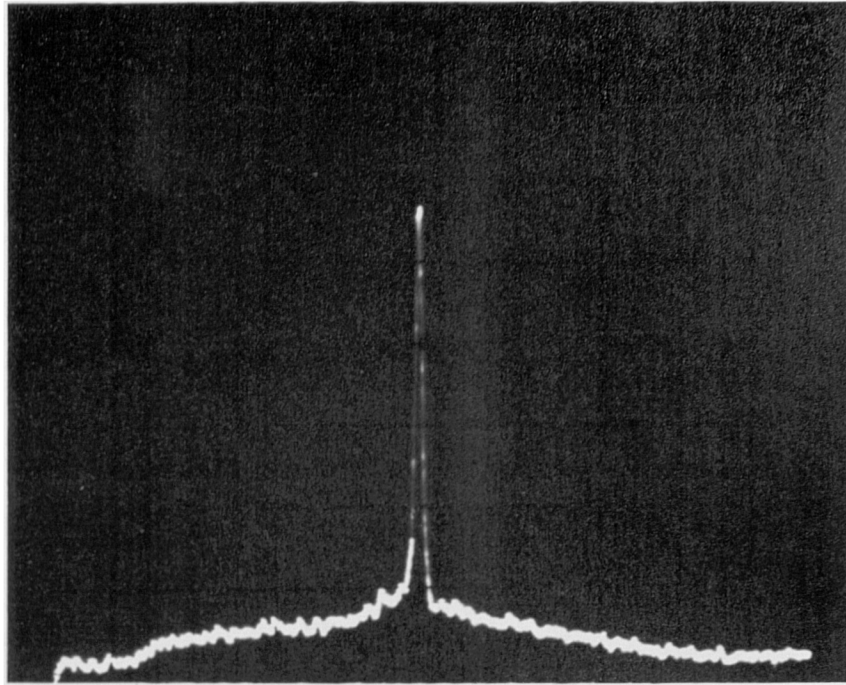


Figure 4.25(a) Auto-correlation response of bandlimited joint transform correlator to Escort view E011

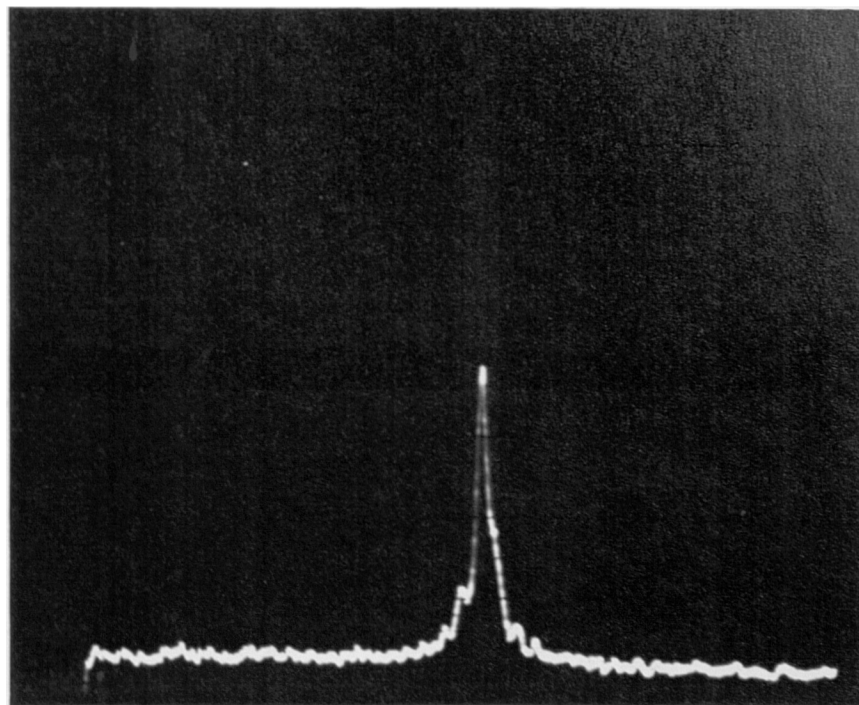


Figure 4.25(b) Cross-correlation response of bandlimited joint transform correlator between Escort view E011 and Mazda view M011

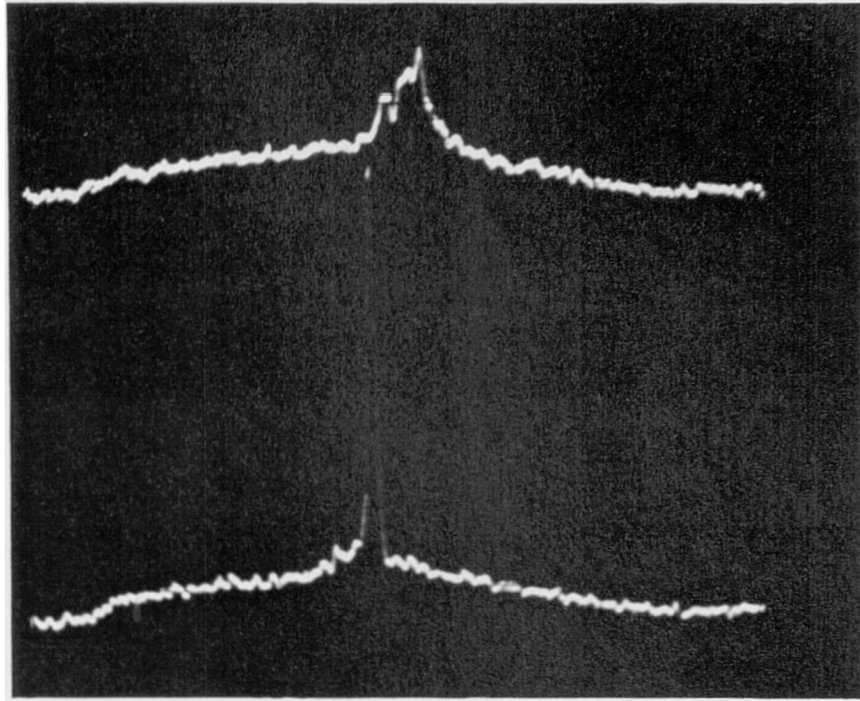


Figure 4.25(c) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E011 and Mazda view M011

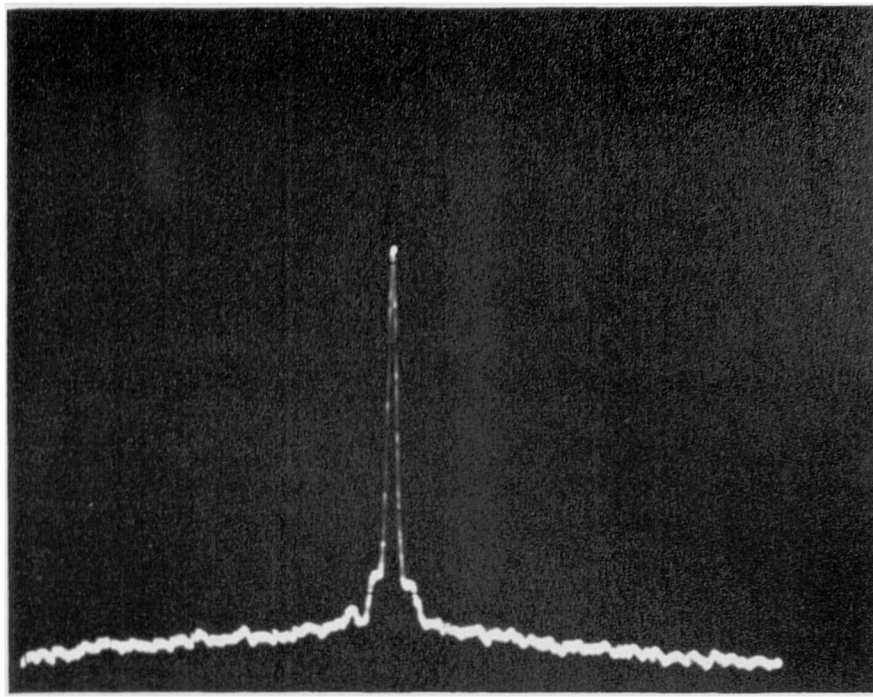


Figure 4.26(a) Auto-correlation response of bandlimited joint transform correlator to Escort view E079

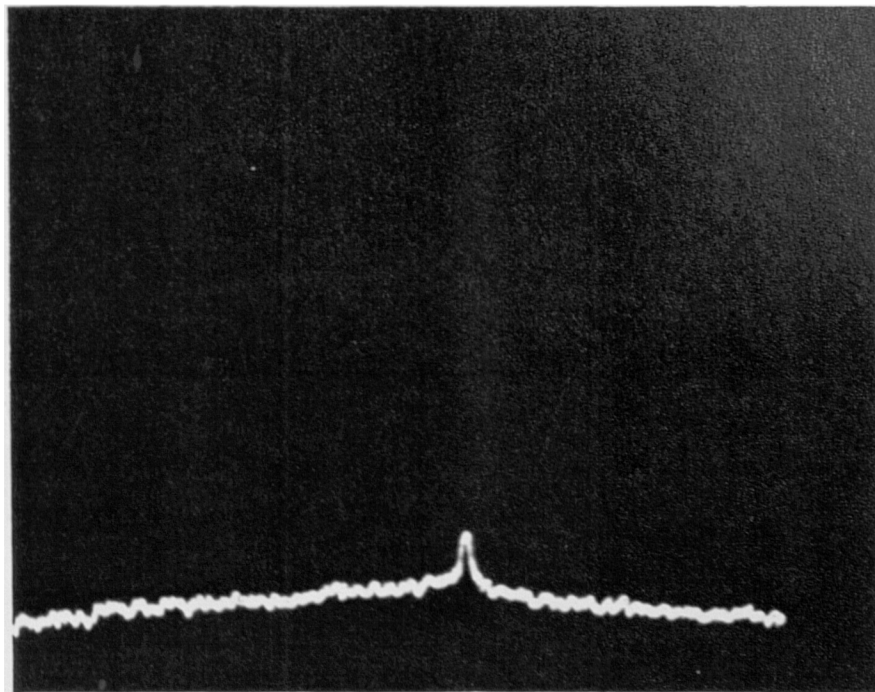


Figure 4.26(b) Cross-correlation response of bandlimited joint transform correlator between Escort view E079 and Mazda view M079

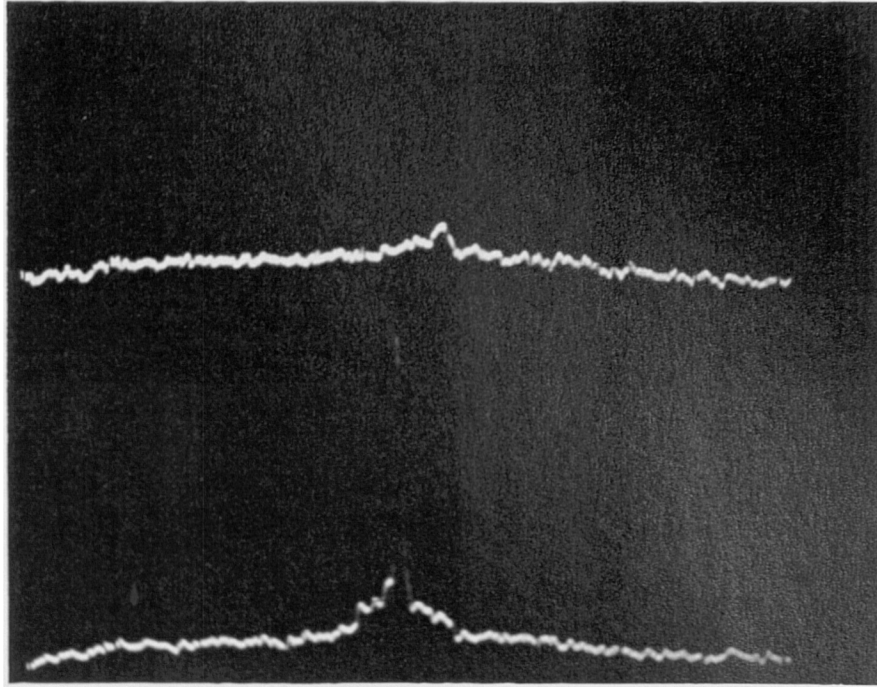


Figure 4.26(c) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E079 and Mazda view M079

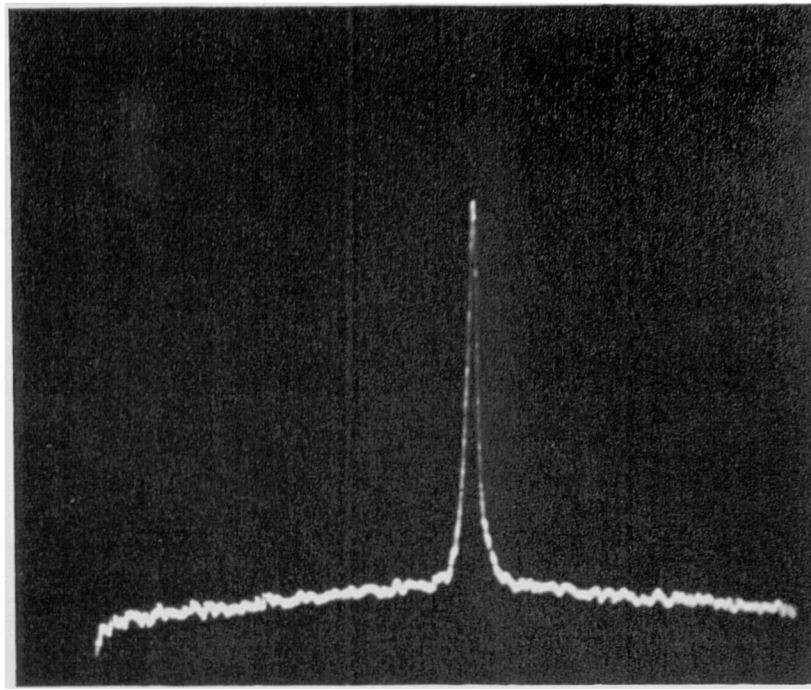


Figure 4.27(a) Auto-correlation response of bandlimited joint transform correlator to Escort view E124

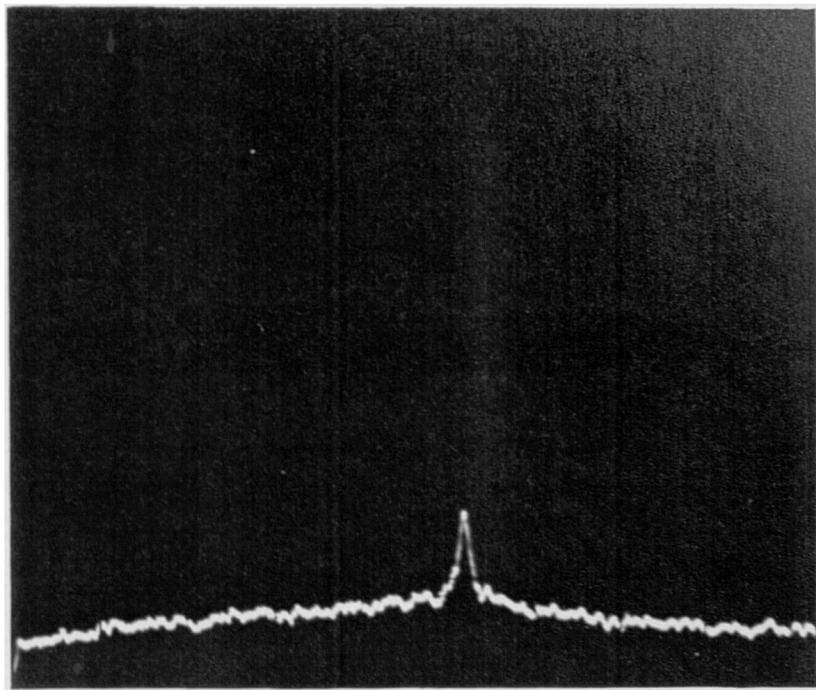


Figure 4.27(b) Cross-correlation response of bandlimited joint transform correlator between Escort view E124 and Mazda view M124

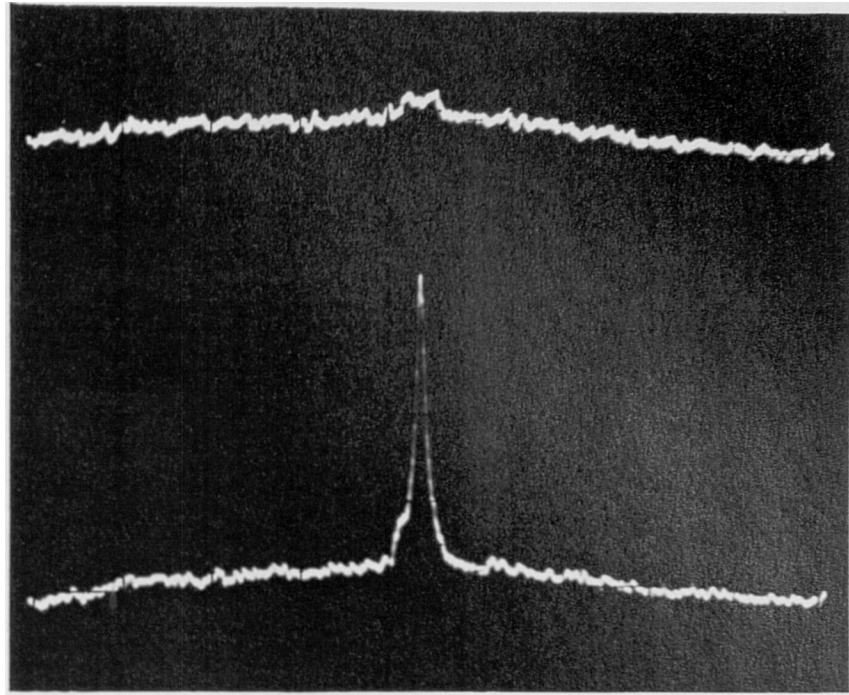


Figure 4.27(c) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E124 and Mazda view M124

CAR ORIENTATION	WITHOUT LOW FREQUENCY REMOVAL		WITH LOW FREQUENCY REMOVAL		INCREASE IN DISCRIMINATION	
	Ratio	dB	Ratio	dB	Ratio	dB
E090 CC M090	2.86	4.56	4.16	6.19	1.45	1.63
E180 CC M180	1.38	1.40	8.00	9.00	5.80	7.60
E000 CC M000	1.42	1.52	3.5	5.44	2.46	3.92
E045 CC M045	1.39	1.43	8.75	9.42	6.29	7.99
E135 CC M135	1.43	1.55	3.55	5.50	2.48	3.95
E011 CC M011	1.46	1.64	3.57	5.53	2.45	3.89
E079 CC M079	7.50	8.75	8.18	9.13	1.10	0.38
E124 CC M124	3.77	5.76	9.47	9.76	2.51	4.00

Table 4 3 Quantitative summary of bandlimited and bandpass joint transform correlator inter-class cross-correlation responses

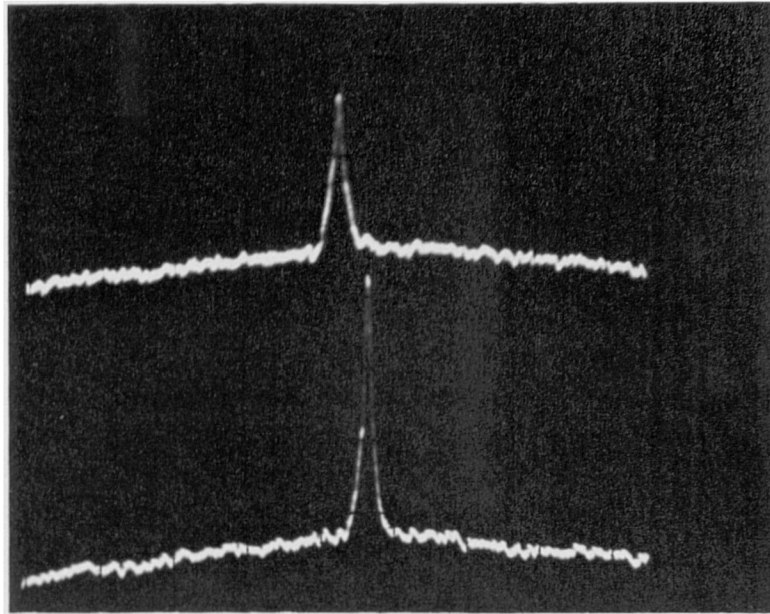


Figure 4.28(a) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E090 and E079

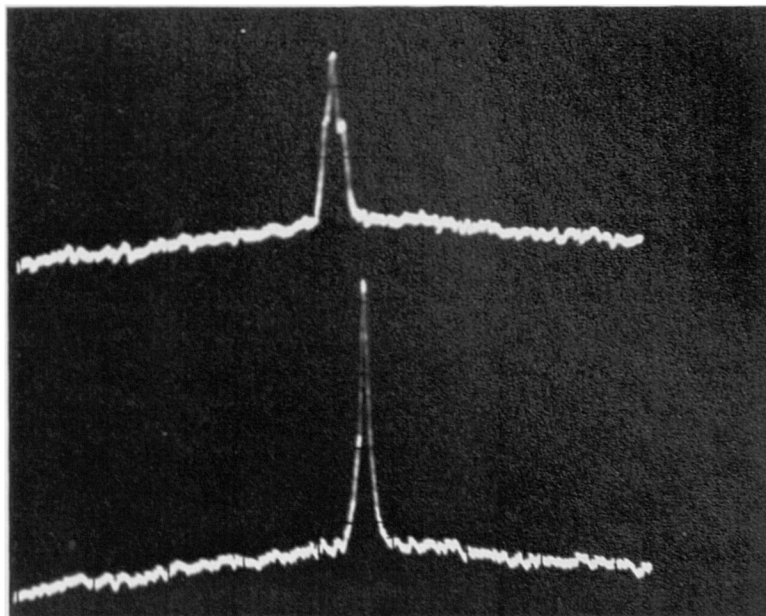


Figure 4.28(b) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E090 and E101

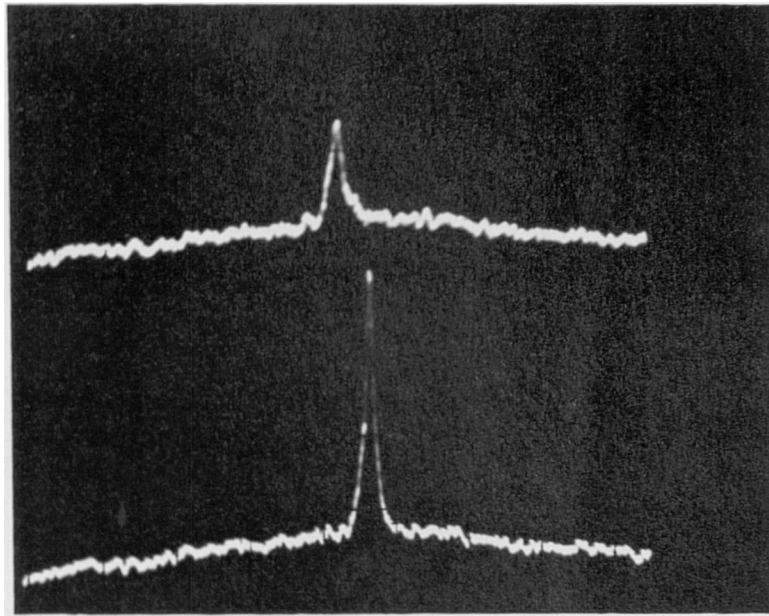


Figure 4.28(c) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E090 and E067

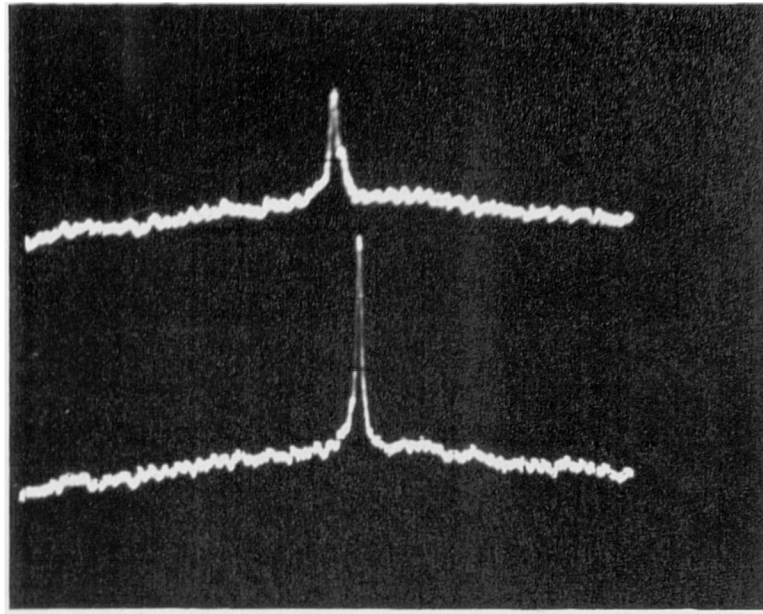


Figure 4.29(a) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E180 and E191

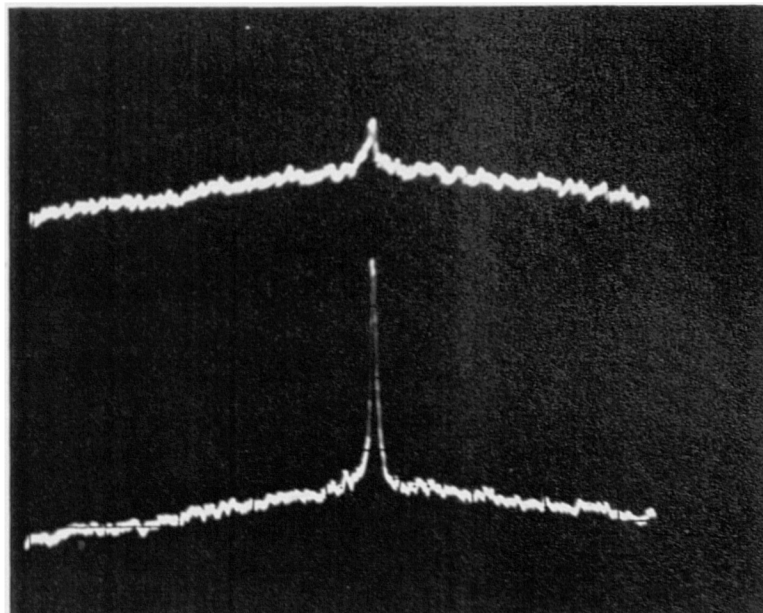


Figure 4.29(b) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E180 and E203

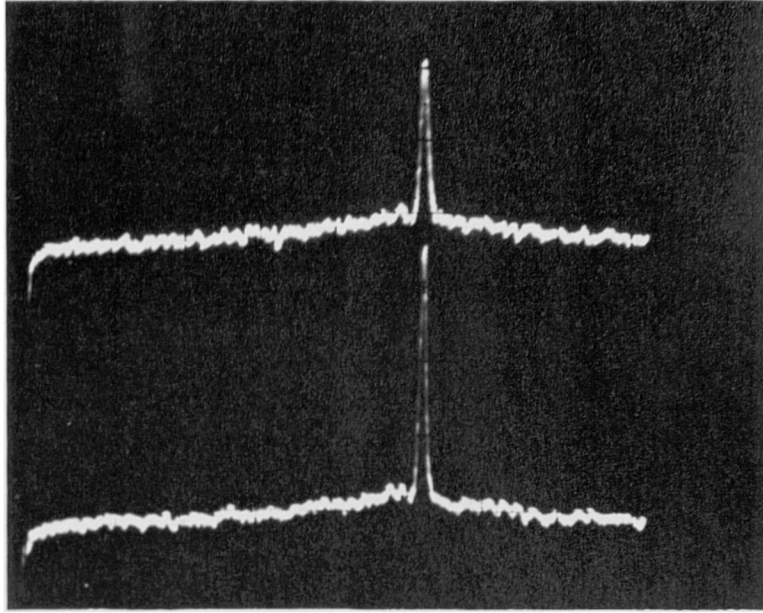


Figure 4.30(a) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E000 and E011

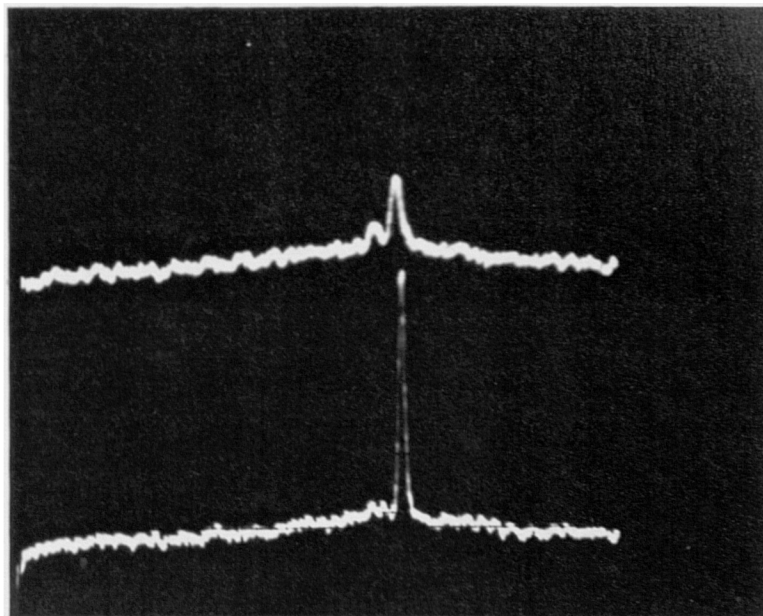


Figure 4.30(b) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E000 and E023

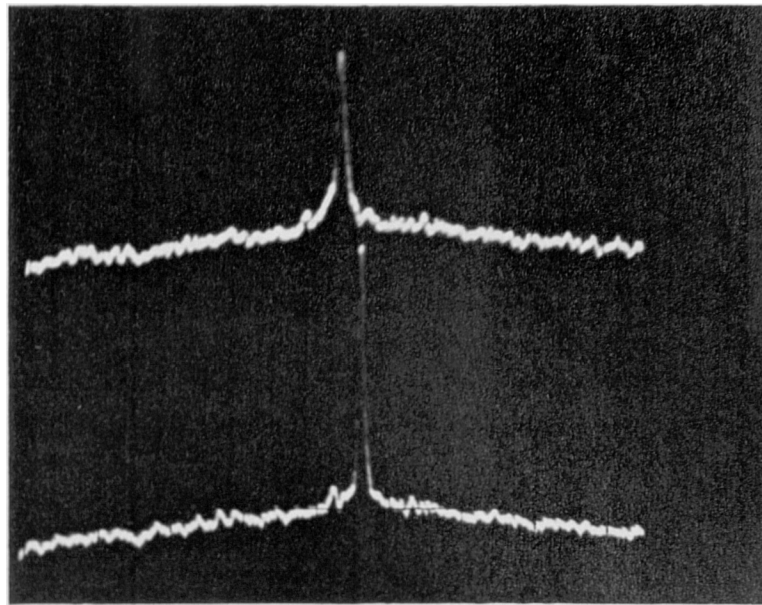


Figure 4.31(a) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E045 and E056

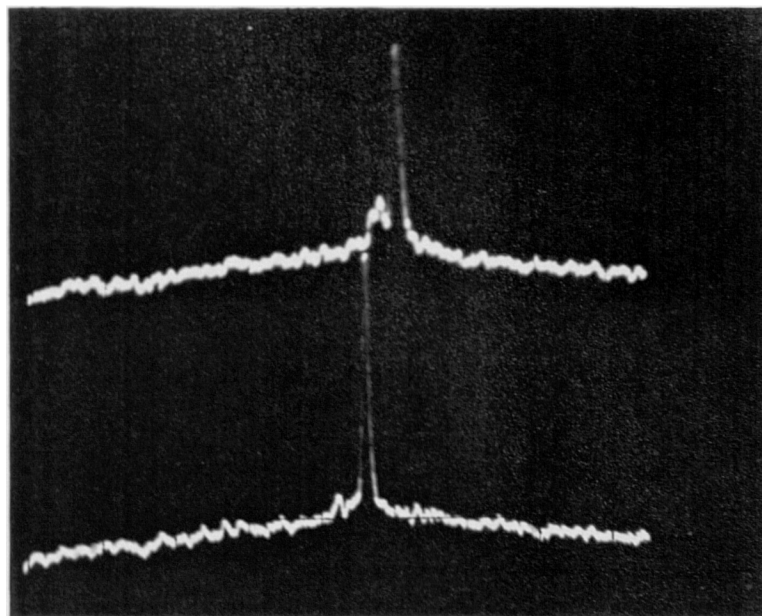


Figure 4.31(b) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E045 and E034

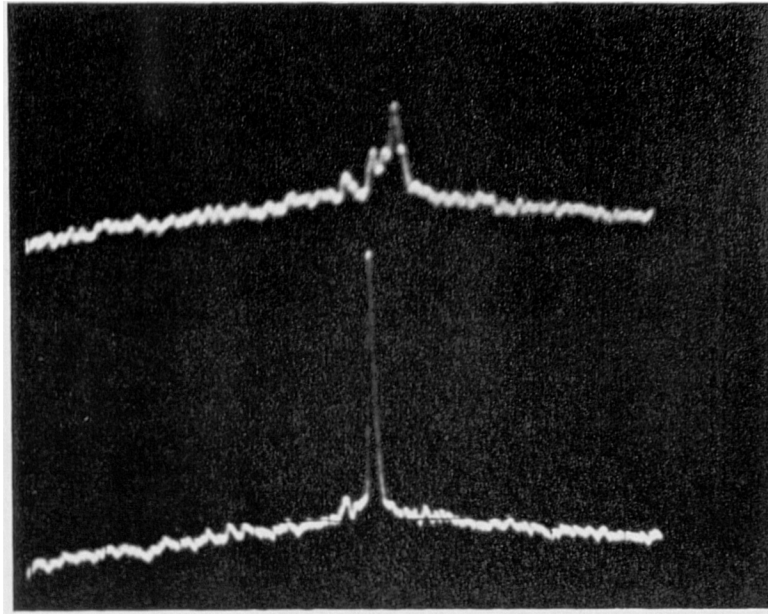


Figure 4.31(c) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E045 and E067

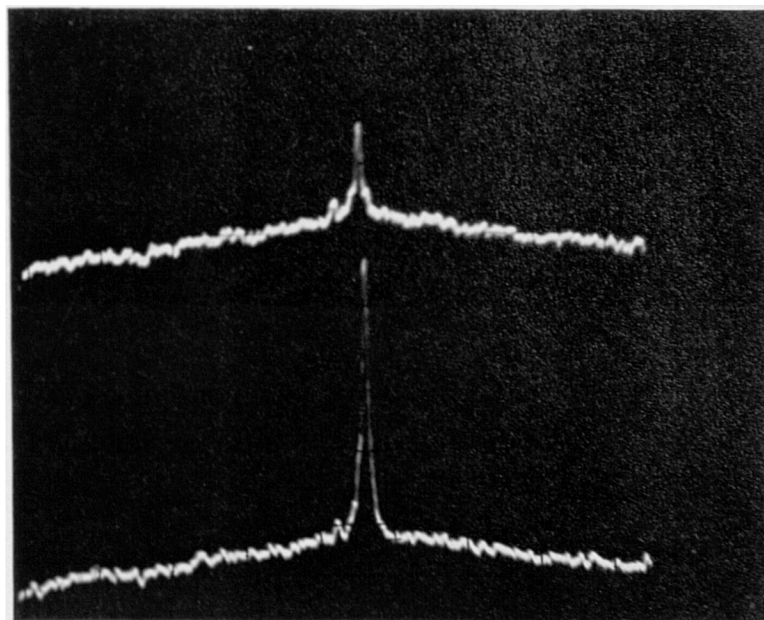


Figure 4.31(d) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E045 and E022

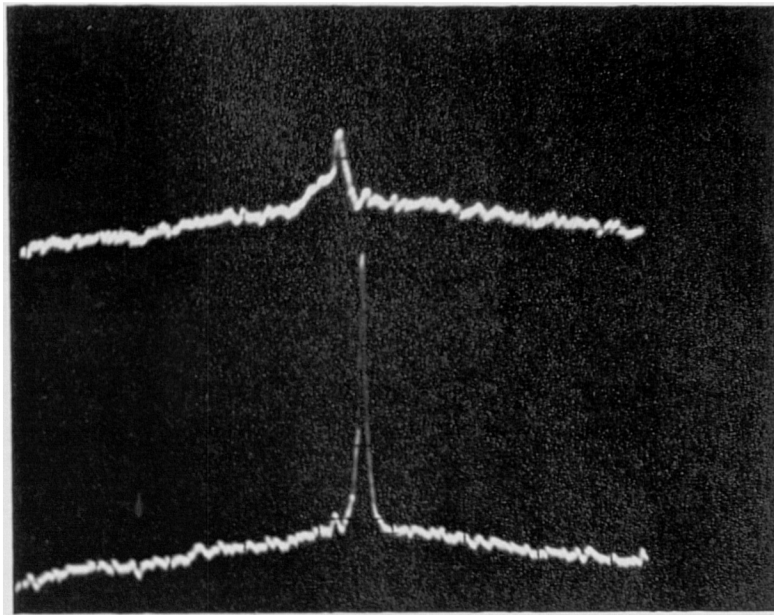


Figure 4.31(e) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E045 and E079

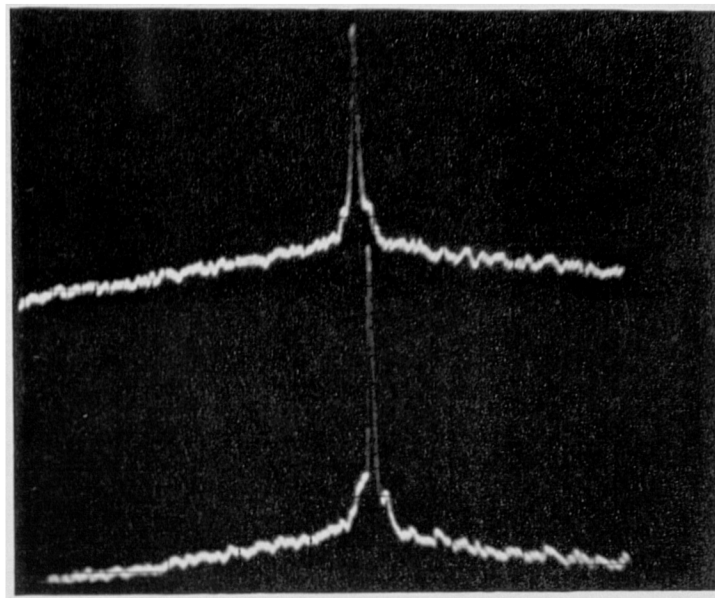


Figure 4.32(a) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E135 and E124

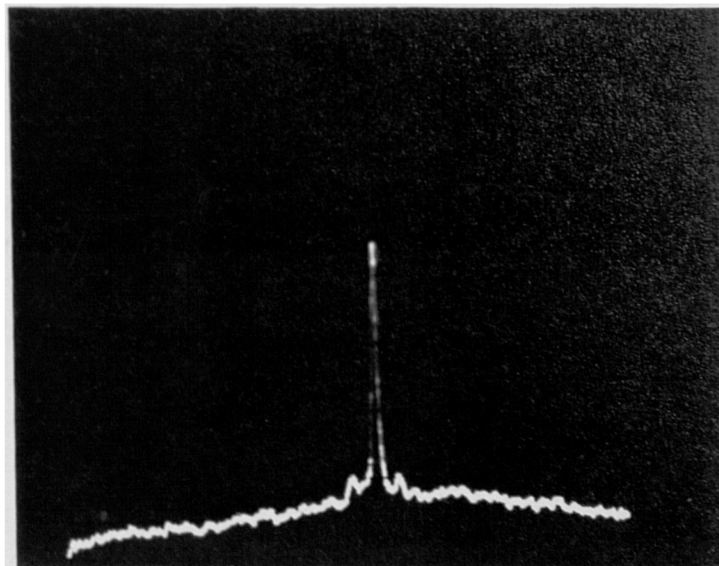


Figure 4.32(b) Cross-correlation response of bandpass joint transform correlator - Escort view E135 and E146

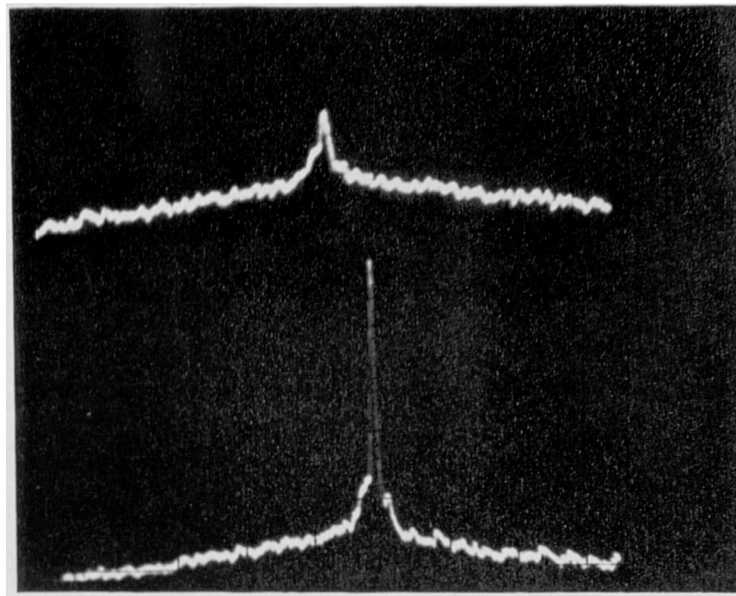


Figure 4.32(c) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E135 and E112

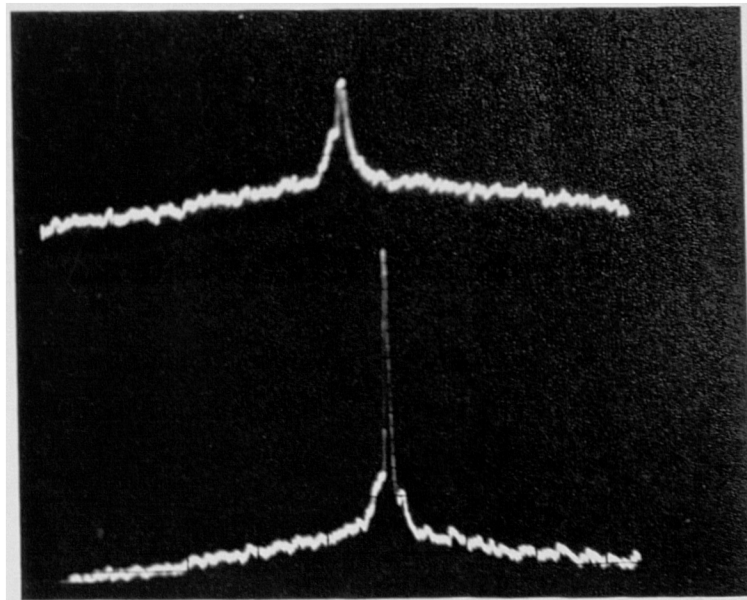


Figure 4.32(d) Auto-correlation and cross-correlation responses of bandpass joint transform correlator - Escort view E135 and E158

CAR ORIENTATION	- 22.5°	- 11.25°	CAR ORIENTATION 0°	11.25°	22.5°
E090 RATIO	0.413	0.625	1.00	0.650	0.438
dB	- 3.85	- 2.04	0.00	- 1.87	- 3.59
E180 RATIO	-	-	1.00	0.603	0.317
dB	-	-	0.00	- 2.20	- 4.99
E000 RATIO	-	-	1.00	0.653	0.280
dB	-	-	0.00	- 1.85	- 5.53
E045 RATIO	0.400	0.688	1.00	0.638	0.375
dB	- 3.98	- 1.62	0.00	- 1.96	- 4.26
E135 RATIO	0.278	0.655	1.00	0.756	0.389
dB	- 5.56	- 1.84	0.00	- 1.21	- 4.10

Table 4 4 Quantitative summary of bandpass joint transform correlator intra-class cross-correlation responses

If $11\frac{1}{4}^\circ$ viewpoint alterations are assumed to be tolerable, sixteen filters would be required to provide 360° of orientation coverage at zero elevation angle. Increasing this to 24 filters would provide for some degree of alteration in elevation angle in addition to accommodating minor variations in in-plane rotation and scale from that of the reference function. At higher elevation angles fewer filters will be necessary. As an estimate of the number of filters required for full 2π steradian coverage, elevations at 10° , 25° , 45° , 75° and 90° will require approximately 20, 16, 12, 8 and 1 filter, respectively. Thus about 80 filters would be required in total. Although extensive simulation and experimentation would be required to confirm this figure, with the car views cross-correlated at all orientations in azimuth and elevation, it is unlikely that this estimate is grossly inaccurate, and is probably within 50% of the correct number. Even if tests were performed over the full range of angles, they would, of course, only be strictly valid for those two particular objects ie. the Escort and Mazda cars. However, it does not seem unreasonable to extrapolate such results to other objects within the same general class as, for instance, other motor vehicles or even military armoured vehicles (although a tank, the example popularly used in optical correlation experiments, is more demanding since it is re-configurable, the turret and main body must be treated as two separate objects each with its own filter bank).

For objects of different types, the number of filters required might be considerably different. For example in Chapter 2, a test industrial component, TC3, was assumed to be capable of resting only in a limited number of stable states, strongly restricting the possible orientations it could present to an overhead camera position. Thus the number of separate filters required was less than ten. However, discrimination between the separate orientations proved to be difficult due to the similar outline shape in each orientation which resulted in a considerable edge overlap between views. The filter had to be biased to high spatial frequencies to provide sufficient discrimination between orientations, thus minimizing sidelobe disruption in the response of the orientation multiplexed synthetic discriminate function filter. Further, discrimination could be assured against another similarly shaped component, TC6, without its specific inclusion in the filter training set. If full orientation independent recognition had been required to accommodate, for example, uncertainties in the component orientation due to machining faults introduced during manufacture, the number of filters needed would have been much larger due to the requirement to bias them towards high spatial frequencies. This would have necessarily resulted in a high sensitivity to alterations in orientation of the component.

As detailed in Chapter 2, the multiplexing of several different orientations into a single filter (or space domain reference image in the case of correlators based on a frequency plane non-linear interaction) results in inevitable deterioration in the peak-to-sidelobe ratio of the correlation response, not only due to the fact that the multiple views produce cross-correlations which cannot be aligned on-axis, but also to the requirement to share the dynamic range within the filter (which with the available SLMs or photographic plate is quite limited). Thus the correlation signal produced by the filter will be lower, with a resultant loss of margin between an intra-class response and that generated by an out-of-class object. This deterioration is likely to be less severe for an object such as a vehicle that changes shape considerably as the orientation alters; different regions of the image or frequency planes will be occupied by the differing high intensity regions of the separate views, in effect resulting in a space, rather than grey level, multiplexing. For this reason, the correlations obtained from multiplexed filters for this type of object may be expected to deteriorate less markedly in strength as compared those achieved with the industrial component example discussed Chapter 2.

However, in-plane rotation must be accommodated in addition to orientation changes. For the cars, as has been previously noted earlier in the Chapter, the spectral energy distribution is highly directional, which trends also to be the case for most man-made objects. Thus for a given orientation, the filter could be effectively angle multiplexed, so utilising fully the available area in the filter plane and avoiding heavy dynamic range multiplexing in the bandpass of importance (the multiple exposure and hence saturation of the unused low frequency regions of the filter near the optical axis being unimportant). The correlation peak generated by such a filter might, therefore, be expected not to be strongly degraded from that of a single image filter, although a compromise between the degree of multiplexing and the acceptable fall-off of the correlation peak at input rotations intermediate to the stored views must be determined for each class of object.

In some applications, particularly those in a controlled environment such as an industrial inspection task, the object to camera distance is known *a priori*. Thus the scale of the object is constant and variations of this parameter are not a source of degradation. For homing tasks, the scale of the object will vary widely and must be controlled at input to the correlator in order to maintain a match to the filter size. To avoid the requirement for filter duplication at the full range of scales likely to be encountered, an auxiliary technique such as laser ranging (to control the setting of a zoom lens system) for compensation of scale changes is preferable. This would only be effective over a certain latitude of scale change, however, and if homing is to be accomplished at scales ranging from an object resolution of a few pixels to that which fills the field of view, a

supplementary initial acquisition technique would be required [11]. Subsequent to this, several sets of reference templates, each covering scale ranges that can be compensated by the zoom lens system, could be employed so that the correlation signal can be maintained over the full distance of the approach. A standard zoom lens providing for a six times zoom factor (for instance, a commonly available CCTV lens with a focal length range of 12.5 mm to 75 mm) would allow a 36 times scale change to be covered with just two sets of separately scaled filters; this should be more than adequate to accommodate the anticipated scale changes.

Thus for the orientation independent recognition of a vehicular type object, the number of individual filters required will be of the order of one to two hundred. If the object is capable of changing orientation with respect to the correlator rapidly, it will be necessary to search this entire filter bank at the input frame up-date rate to ensure a correlation response is maintained. Since the input frame rate must be at least at standard video rates (33 Hz or 40 Hz) and initially the object has to be identified amongst several others, a large reference template data base is required which is capable of rapid random access. In the final Chapter a correlator design is presented in which it is proposed that the input scene is digitally Fourier transformed at video rates by employing recently available FFT chip sets. The spectral phase data so computed is applied to a phase modulating SLM which, since it is in the input rather than reference arm of the correlator, need only be capable of responding at video rates. The template data is stored optically in a high capacity volume holographic memory as angle multiplexed holograms. This large amount of data may then be read-out and optically Fourier transformed at the high rates demanded by the requirement to search several hundred filters per input frame. Calculations indicate that it should be possible to search a 1000 template memory at 10 kHz employing system components currently available although, due to current bandwidth restrictions of the CCD output detector, the recognition cycle will be limited to 3 kHz. Further, the bandpass filtering techniques developed in this and previous Chapters may be directly implemented in this configuration which would allow the reference templates to be designed such that they have good inter-class discrimination while maintaining intra-class tolerance. This hardware configuration, in conjunction with the implementation of the filtering techniques that have been proposed, may provide solutions to demanding pattern recognition problems that are currently beyond the capabilities of algorithms implemented by solely digital means.

4.6 Conclusions

In conclusion to the discussion of filtering techniques that are implementable directly in the photorefractive based joint transform correlator, brief mention will be made of two methods which would require experimental implementation for further assessment. Firstly, the discrete frequency removal technique, which has been discussed and simulated in Chapter 3 with specific reference to the Gaussian apodized joint transform correlator, may be implemented experimentally by attenuation of pre-determined frequency components within the Gaussian envelope of the HeNe read-out beam. Thus localised regions of the beam could be attenuated which correspond to the frequency locations specified for removal by the algorithm described in Chapter 3. This could be accomplished dynamically with a binary amplitude modulating SLM (for example a ferroelectric liquid crystal based device capable of a rapid frame up-date) which is imaged onto the face of the BSO crystal in the frequency plane of the correlator in a similar fashion to the DOG profile. This would lead to increased intra-class tolerance and inter-class discrimination for specific car views with a possible extension to multiplexed filters as discussed in Chapter 3.

A second filtering technique that it is possible to implement directly in the joint transform correlator is the Wiener transfer function discussed in Chapter 1. Consider the application in which an isolated object, such as a vehicle for example, is to be detected and located within a highly cluttered background scene. If the reference template, $h(\mathbf{r})$, modulates beam 3 and the input function containing $h(\mathbf{r})$ together with the background scene, $n(\mathbf{r})$, modulates beam 1, the BSO grating modulation may be written as (see Equation {4.1}) .

$$m(\mathbf{k}) = \frac{2H^*(\mathbf{k})(H(\mathbf{k}) + N(\mathbf{k}))}{|H(\mathbf{k})|^2 + H(\mathbf{k}) + N(\mathbf{k})^2} \quad \{4.14\}$$

As emphasised in Chapter 1 the background scene cannot be assumed to have zero correlation with $h(\mathbf{r})$ since, even if the signals were independent random processes, they are introduced into the correlator with a non-zero mean via space domain SLMs. However, for the present discussion assume that the non-zero mean is removed with a "dc" stop in the frequency plane and that the structured object, $h(\mathbf{r})$, has a relatively low cross-correlation with $n(\mathbf{r})$ at all locations in (x,y) as compared to the individual auto-correlations of $h(\mathbf{r})$ and $n(\mathbf{r})$. If this is the case the cross-spectral terms $H(\mathbf{k})N^*(\mathbf{k})$ and $H^*(\mathbf{k})N(\mathbf{k})$ may be ignored and Equation {4.14} reduces to :

$$m(\mathbf{k}) = \frac{2H(\mathbf{k})}{2|H(\mathbf{k})|^2 + |N(\mathbf{k})|^2} \quad \{4.15\}$$

Comparing this to the Wiener filter transfer function :

$$W(\mathbf{k}) = \frac{G(\mathbf{k})}{|G(\mathbf{k})|^2 + \frac{|N(\mathbf{k})|^2}{|G(\mathbf{k})|^2}} \quad \{4.16\}$$

indicates that to implement the Wiener filter directly in the photorefractive based joint transform correlator beam 2 should be modulated with the following amplitude distribution :

$$U_2(\mathbf{k}) = \frac{|H(\mathbf{k})|^2 + \frac{|N_e(\mathbf{k})|^2}{2}}{|H(\mathbf{k})| + \frac{|N_e(\mathbf{k})|^2}{|H(\mathbf{k})|^2}} \quad \{4.17\}$$

where $|N_e(\mathbf{k})|^2$ is an estimate of the power spectral density of the background at the input previously derived from a statistical model, or formed as an average of sampled data of the background likely to be encountered in a given application. Thus $U_2(\mathbf{k})$ should be capable of modification with changes in the background terrain. This could be accomplished with an amplitude modulating SLM in the HeNe read-out beam, the display of which is imaged onto the frequency plane at the location of the photorefractive crystal

As discussed in Chapter 1, if the assumption of low correlation between signal and noise cannot be made, the Wiener filter transfer function takes the more lengthy form :

$$W(\mathbf{k}) = \frac{H^*(\mathbf{k})|H(\mathbf{k})|^2 + N^*(\mathbf{k})H(\mathbf{k})}{|H(\mathbf{k})|^4 + |N(\mathbf{k})|^2 + (H(\mathbf{k}))^2 N^*(\mathbf{k}) + (H^*(\mathbf{k}))^2 N(\mathbf{k})} \quad \{4.18\}$$

Comparison with Equation {4.14} indicates that to implement this transfer function in the joint transform correlator, $U_2(\mathbf{k})$ should be modulated according to :

$$U_2(\mathbf{k}) = \frac{\left(|H(\mathbf{k})|^2 + \left(\frac{N(\mathbf{k})}{H(\mathbf{k})} \right)^* H(\mathbf{k}) \right) (2|H(\mathbf{k})|^2 + |N(\mathbf{k})|^2)}{|H(\mathbf{k})|^4 + |N(\mathbf{k})|^2 + (H(\mathbf{k}))^2 N^*(\mathbf{k}) + (H^*(\mathbf{k}))^2 N(\mathbf{k}) + H^*(\mathbf{k})N(\mathbf{k}) + H(\mathbf{k})N^*(\mathbf{k})} \quad \{4.19\}$$

Experimentally this would be difficult to achieve since independent amplitude and phase control of $U_2(\mathbf{k})$ would be required since Equation {4.19} is complex; this could only be achieved dynamically by two SLMs in tandem. Simulation work is required with typical terrain data to determine if a significant increase in signal detectability can be achieved by modulating $U_2(\mathbf{k})$ with the functions given by Equations {4.17} and {4.19} and what performance differential can be obtained with the more involved transfer function. If this is significant, implementation details could then be considered further.

References to Chapter 4

- [1] A Yariv, "Optical Electronics," Holt, Rinehart and Winston, (1985)
- [2] P. Guenther, "Holography, coherent light amplification and optical phase conjugation with photorefractive materials," *Physics Reports*, Vol. 93, No. 4, p.199, (1982)
- [3] H. Kogelnik, "Coupled wave theory for thick hologram gratings," *Bell Syst. Tech. J.*, Vol. 48, No. 9, p.2909, (1969)
- [4] T. Slack, "Design of an updateable optical correlator," *Optical computing and processing*, Vol. 2, No. 1, p.19, (1992)
- [5] M. Nicholson, I. Cooper, M. McColl, C. Petts, "Simple computational model of image correlation by four wave mixing in photorefractive media," *Appl. Opt.*, Vol. 26, No 2, p.278, (1987)
- [6] D Gabor, "Theory of communication," *J. IEE*, Vol. 93, Part 3, p.429, (1946)
- [7] I Daubechies, "The wavelet transform time-frequency localisation and signal analysis," *IEEE Trans. Inf. Theory*, IT-36, No. 5, p.961, (1990)
- [8] D Marr, E Hildreth, "Theory of edge detection," *Proc. R. Soc. Lond. B*, Vol. 207, p 187, (1980)
- [9] A Rosenfeld, A Kak, "Digital picture processing," Academic Press, (1976)
- [10] B Guenther, C. Christensen, J. Upatnieks, "Coherent optical processing : another approach," *IEEE J. Quantum Elect.*, Vol. QE 15, No. 12, p.1348, (1979)
- [11] D Casasent, "An optical correlator feature extractor neural net system," *Opt. Eng.*, Vol. 31, No. 5, p.971, (1992)

Chapter 5

A proposed high speed correlator system and its performance assessment

5.1 Introduction

For many years now it has been appreciated that the inherent Fourier transforming property of a lens, together with the capability of performing a complex multiplication between a holographically recorded function and a coherent wavefront, permits the rapid two-dimensional correlation between a reference template and an input scene containing an unknown image [1]. However, as the work described in the previous Chapters has indicated, for the recognition of a solid object unconstrained in scale, rotation, orientation and contained in a generally cluttered background, many reference template searches are required to achieve reliable recognition. The considerable recent effort on filter multiplexing techniques, most notably the synthetic discriminant function (SDF) technique and its variants [2], is aimed at reducing the required number of sequential filter searches by storing many different object views on a single filter. Modifications to this procedure were described in Chapter 2 where it was shown that, in order to maintain correlation peak localisation, these filters are biased towards high spatial frequencies which inevitably means increments between training image viewpoints must be small in order to maintain the correlation response for intermediate object views. This in turn implies that, given the limited dynamic range of recording media or spatial light modulators (SLMs), only a limited degree of alteration in object aspect can be recorded within a single multiplexed filter. Many such filters must be rapidly searched to acquire and maintain a correlation signal at the required rate. Thus recent approaches to real time correlators have largely been hybrid systems, employing either two space domain SLMs together with a non-linear interaction in the frequency domain [3], or a space domain amplitude modulating SLM addressing a frequency domain phase modulating SLM [4,5]. In this way many reference templates may be rapidly passed to the SLM as space domain images or as Fourier transform phase data from the host's digital mass storage device.

From the foregoing discussion it is clear that there is an essential asymmetry in the rate at which the input SLM must respond, and the associated Fourier transform be

computed, as compared to that of the reference SLM and final Fourier transform. The reference template search must be performed between one and three orders of magnitude faster, depending on the exact application, in particular the degree of object variability. Consider the requirement for a CCIR video rate (25 frames/sec, 512 x 512 resolution image) correlator capable of re-acquiring the correlation at video rates. To achieve this the first Fourier transform need be computed no faster than video rates provided the frequency domain phase modulating SLM can be addressed and respond rapidly. Many directed reference template searches may then be performed during the next 40 msec interval in which the following input frame is acquired and Fourier transformed.

Since it is now feasible to perform a video rate 2-D Fourier transform digitally, the following sections conclude the thesis by describing the outline design of an optical/digital hybrid correlator system configuration that exploits this development. A clear speed advantage over a purely digital implementation is demonstrated by the hybridisation of digital and coherent optical processing techniques. In the proposed system the input scene is digitally Fourier transformed and the phase portion of the spectrum applied to a phase modulating SLM. The reference templates are stored as angle multiplexed volume holograms enabling rapid access so that multiple optical correlations may be performed with a single digitally Fourier transformed scene.

5.2 Digital Fourier transform

A system layout diagram is shown in Figure 5.1. The 2-D fast Fourier transform (FFT) board may be hosted from a PC provided a high bandwidth bus is employed to allow video rate communication between processor board and framestore. The data is acquired via a CCD camera and since the initial Fourier transform is performed digitally, an input amplitude modulating SLM is not required. This camera could be CCIR (or RS170) standard provided the framestore clock is set such that a 1:1 pixel aspect ratio is maintained in the framestore and a uniform vertical and horizontal sampling rate achieved. This is particularly important if the image is to be subsequently Fourier transformed. A somewhat preferable alternative is the use of a square pixel, 512 x 512 resolution, digital output camera since this results in an exact mapping of camera pixels to framestore memory locations. This requires a non-standard video interface board but this is also necessary for interfacing a high-speed correlation plane CCD camera.

A schematic of the 2-D FFT board is shown in Figure 5.2. This relies on the recent

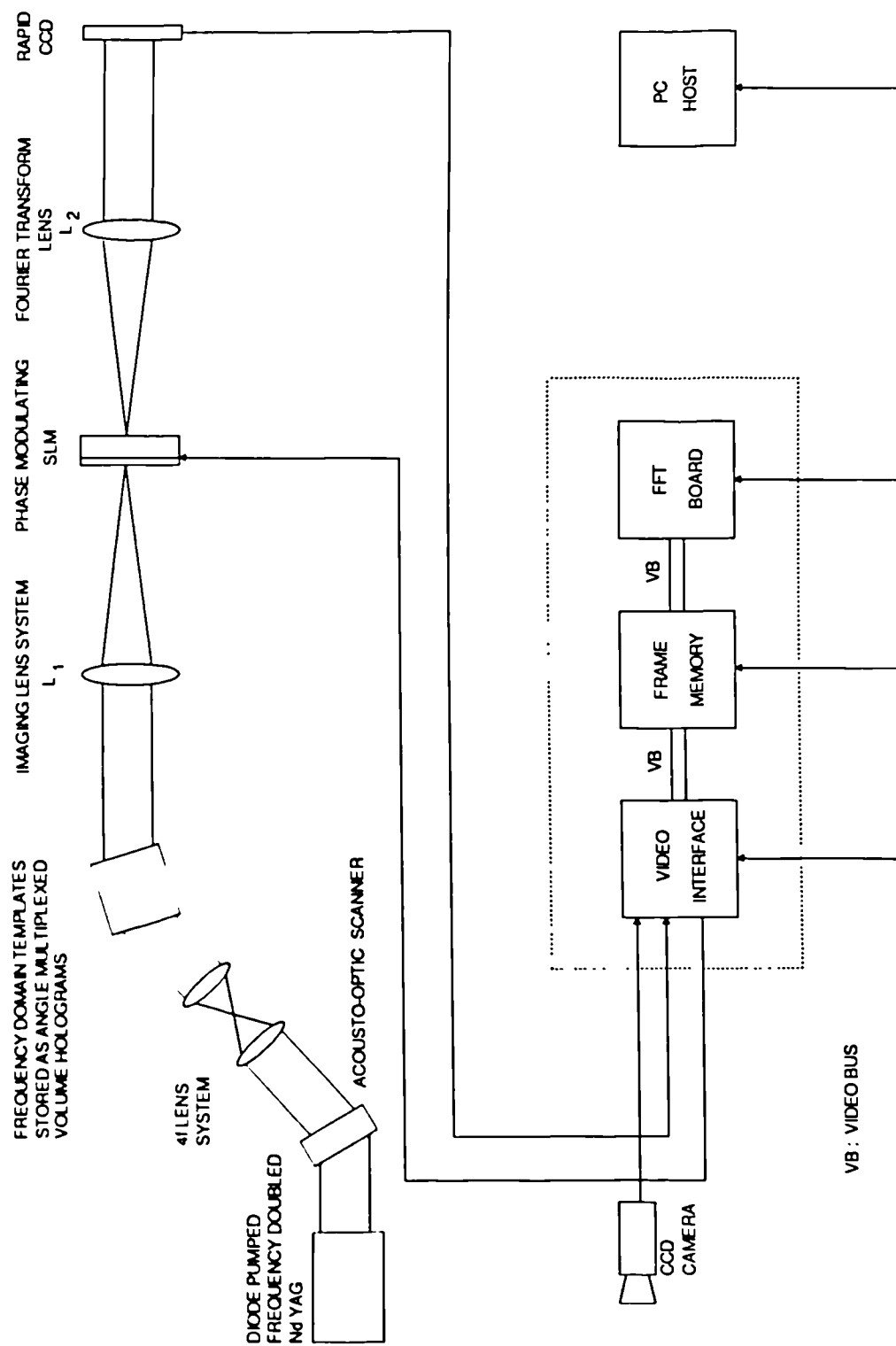


Figure 5.1 Schematic of the hybrid optical/digital correlator

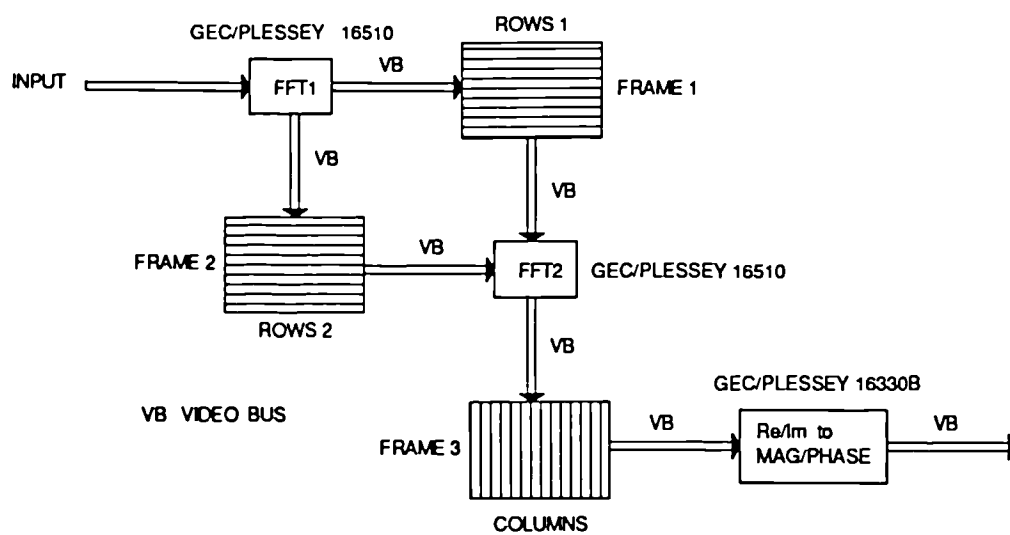


Figure 5 2 Layout of the two-dimensional fast Fourier transform processor

availability of 1-D FFT digital signal processing chip sets capable of video rate data throughput. The GEC/Plessey PDSP 16510 can perform a 1024 complex to complex, 16 bit 1-D FFT calculation in 96 μ secs. [6]. (The device may also be configured to internally perform two simultaneous real to complex transforms at almost double the overall rate.) Thus the transformation of the 14.75 MHz video signal (the uniform aspect ratio sampling rate) is within its capabilities. To accomplish a 2-D FFT, two processors must be employed with an intermediate frame memory, as shown in Figure 5.2. The digitised video signal (8 bits per pixel) is horizontally transformed by the first processor and the intermediate result stored in Frame 1. The next video frame is then horizontally Fourier transformed by this processor, the intermediate results being stored in Frame 2. During the second frame, the data in Frame 1 is transformed vertically by the second processor to complete the 2-D FFT on the first video frame, the result being placed in Frame 3. For the second frame the data input to the second processor is then switched to Frame 2 and the results stored in Frame 3. The GEC/Plessey 16330B DSP converts the real and imaginary components of the computed 2-D FFT to magnitude (16 bit) and phase (12 bit) with a total delay of only two clock cycles. Thus the 2-D FFT of the input can be computed at video rates one frame behind real-time. A Hamming or Blackman-Harris window operator may be internally applied in the PDSP 16510 to reduce the leakage artifact in the computed spectral components. Although 12-bits of phase data is provided by the FFT processor, only the 4 most significant bits are used, allowing 16 levels of phase quantisation. Provided the phase modulating SLM can accurately convert this to a modulation of the coherent wavefront, this quantisation level allows almost complete diffraction into the first order. In this regard, Amako *et al.* [7] have recently demonstrated that the Seiko-Epson liquid crystal display (LCD), modified from a twisted to linear nematic configuration, is capable of conversion of a 4 bit addressing signal to a phase modulation of a coherent wavefront with sufficient accuracy to allow good quality kinoforms to be reconstructed. The device could be similarly employed to display the phase component of the input scene spectrum, a more detailed description of its use for this purpose being given in the next section. In this way a close approximation to a phase only filter [8] may be implemented. A bandpass modification of the filter to give improved intra-class tolerance while maintaining inter-class discrimination is described in Section 5.5.2.

5.3 Phase modulating spatial light modulator

The interface between the electronic and optical processing sections of the correlator is provided by the phase modulating SLM. This must be capable of being addressed and responding in a time interval shorter than the 40 msec during which the following frame is being acquired and Fourier transformed. The remainder of this interval is then available for the template search. A promising technology for this application is a silicon active matrix backplane addressing a liquid crystal layer [9,10]. The silicon backplane allows a high resolution device ($30\text{ }\mu\text{m} \times 30\text{ }\mu\text{m}$ pixels) suitable for insertion into the frequency plane of an optical correlator without the need for a special telephoto lens design to magnify the Fourier transform size. Multiple addressing ports allow the data to be loaded onto the backplane in 1-2 msec. If a linear nematic liquid crystal layer is employed as the phase modulating material a further delay period is required for the tilting of the liquid crystal polar molecules under the locally applied electric field. The linear nematic phase has a response time somewhat slower than the twisted nematic configuration but it is nevertheless of the order of 20 msec. Thus under 20 msec remains to complete the template search before the next video frame must be loaded. Faster responding liquid crystal phases may be interfaced to a silicon backplane. Currently, binary phase modulation has been achieved with a smectic C* based device [11]. However, the smectic A* phase allows a multi-level capability and its interface to a silicon backplane is actively being researched [12], although at present only a limited degree of phase modulation is possible [13]. The smectic liquid crystal phase is capable of responding on sub-millisecond time scales, implying that the overall device speed will be limited only by the obtainable data transfer rates onto the silicon backplane.

If the overall response time of the phase modulating SLM is of the order of a frame time (ie. 40 msec for the CCIR standard) a half video rate correlator may be constructed. The digital FFT may then be accomplished with a single PDSP 16510 operating on the video rows directly from the CCD camera input. These intermediate results are stored in a frame memory and during the next 40 msec vertically transformed to form the 2-D FFT in a total of 80 msec. (In fact the computation could be completed in 60 msec since only half of the vertical FFTs need be computed, the final spectrum of the real input signal having diametric Hermitian symmetry.) During the first 40 msec interval the SLM is loaded and the liquid crystal molecules re-orientate. The second 40 msec is then available for the template search. One of the commercially available devices developed for miniature TV displays could be used for this purpose. The Seiko-Epson LCD has been found to give a good performance as an SLM [14], although the commercially available device employs a twisted

nematic LC layer since it is designed as an amplitude modulator. However, as mentioned above, its use with a linear nematic LC has been reported with 16 levels of phase modulation over a full 2π radians having been achieved [7]. This device could, therefore, be used directly in the proposed system provided a telephoto lens were employed to magnify the reference image spectra to the 25.6 mm x 19.8 mm size of the active area. The LCD consists of 320 x 220 pixels with each pixel measuring 55 μm x 60 μm on a 80 μm x 90 μm pitch. However, it is expected that a higher resolution VGA compatible system will be available shortly. One of its main advantages over earlier LCTVs is its improved optical quality, 0.5λ across the clear aperture having been reported by Kirsch *et al.* [14]. These workers have also determined that the addressing scheme is arranged to take alternate video fields and apply these to the same LCD pixels so creating a refresh rate of twice the video frame rate with a vertical resolution half that of the input video signal.

A single PDSP 16510 could, therefore, be used to perform a video rate, half resolution 2-D FFT. This could be accomplished by horizontally transforming the first field, ie. every second line of data, from the CCD camera and, during the second field period, vertically transforming the first field to complete the 2-D FFT on this data which could then be passed to the LCD twice in succession to minimize the pixel sag on the LC layer. This would allow, therefore, the field time interval of around 20 msec for the template search during which the second identical frame is rewritten to the device. The following frame would then change the display to that corresponding to the Fourier transform of the first field of the next input video frame.

Another SLM type that seems particularly suitable for this application is the Hughes CCD addressed LCD [15,16]. The CCD is operated in reverse to its normal mode in that a charge pattern is loaded onto the array which then modulates the LC layer with this 2-D spatial pattern. The CCD provides the essential serial to parallel interface required to provide a stable charge pattern whilst the device is read-out optically. This addressing technology has the advantage over that of thin film transistors in that the CCD array can be made sufficiently small (< 20 μm x 20 μm pixel size) to be placed directly in the frequency plane of a conventional Fourier transform lens. Again the device developed is configured as an amplitude modulator with a twisted nematic LC layer but could readily be modified for use as a pure phase modulator by using a linear nematic. However, as mentioned above, this liquid crystal configuration is slightly slower than that of the twisted nematic. A further alternative is the electron beam addressed SLM developed by OPTRON Systems Inc. In one variant of the device phase modulation is accomplished with a linear nematic LC layer; in another it is achieved utilising an array of deformable mirrors.

5.4 Volume hologram template memory

As shown in Figure 5.1, it is proposed that the reference templates are stored as angle multiplexed volume holograms in a photorefractive crystal. Thus a large template library may be stored in a compact volume which it is possible to access at the rapid read-out speed required of a real-time correlator. Direct inclusion of the volume holographic medium in the frequency plane of an all-optical correlator is precluded since the very basis of the angle multiplexed storage of the multiple images is the extreme Bragg selectivity of the volume hologram. Thus such a correlator would lose shift-invariance, one of the most attractive features of the correlation algorithm when searching for a small target in a cluttered high spatial resolution input scene. However, if the input Fourier transform is digitally computed, the linear phase factors associated with the target translations in the input scene will be contained in the Fourier transform phase function applied to the phase modulating SLM. This will lead to appropriate tilting of the coherent processing wavefront as it propagates through the SLM so as to generate the correlation signal with correct location upon subsequent Fourier transformation to the output plane. An individual template may be reconstructed from the volume hologram by an appropriately angled plane wave reference beam. A rapid template search may then be performed, during the frame interval in which the display on the SLM is held stable, by rapidly scanning the read-out beam through the swathe of angles originally used in recording the template holograms into the volume medium. This may be accomplished with an acousto-optic scanner and 4f optical system as shown in Figure 5.1, commercially available A/O scanners being capable of more than adequate resolution and scanning speed for this task. For example, the ISOMET LS-110 tellurium dioxide (TeO_2) Bragg cell has a time bandwidth product (equivalent to resolution) of 1100 with an acoustic fill time of 22.7 μsec . Ultimately a laser diode array (SELDA) could be used to address the individual holograms directly [17].

Since the volume hologram must be read-out non-destructively a fixing process is required after recording of the templates. Electrical fixing has been reported in the ferroelectric photorefractives barium titanate [18] and strontium barium niobate [19] and thermal fixing in iron doped lithium niobate [20]. Although it is possible to induce a larger refractive index change in the ferroelectric materials by virtue of their large Pockels' coefficients, and hence an increased diffraction efficiency of the reconstructed holograms, lithium niobate is available as relatively large (1 cm^3) high quality, defect free crystals. This is important since any defects cause scattering leading to increased cross-talk between the reconstructed holograms. Mok [21] has stored 500 images as angle multiplexed holograms in Fe doped LiNbO_3 using an

exposure schedule to obtain a uniform diffraction efficiency between the reconstructions and a thermal fixing process to prevent erasure upon readout. The reconstruction diffraction efficiency was 10^{-4} with minimal cross-talk between individual reconstructions. This has recently been extended to the storage of 5000 images with an altered recording geometry [22]. Even though the diffraction efficiencies of such heavily multiplexed holograms are low, a relatively fast read-out with a moderate power reconstruction beam is still possible as the following example calculation demonstrates.

5.5 Beam energy losses through the system

5.5.1 Source power and component losses

Compact laser diode sources are available giving 1 W of output with sufficient coherence for hologram reconstruction. However, these emit in the near infrared at around 800 nm. If LiNbO_3 is used as the storage medium a shorter wavelength must be employed to record the holograms in order to obtain sufficient sensitivity, the 488 nm or 514.5 nm Argon ion lines being commonly used, or the 532 nm line from a frequency doubled Nd:YAG. Reconstruction of the holograms with a near infrared source would then lead to aberrations in the reconstructed images. In addition, a further problem is the Bragg dephasing introduced by the wavelength shift and a resultant loss of field of view and resolution in the carrier frequency direction of the hologram. Although the field of view can be restored by the incorporation of an appropriate phase curvature into the read-out beam, there remains a resolution loss in the reconstructed image [23]. A more direct solution is to closely match the read-out wavelength to that with which the holograms were originally recorded (making a robust fixing process essential to prevent hologram degradation with read-out). Relatively compact diode pumped, frequency doubled Nd:YAG sources are available with excellent temporal coherence but are limited in power to around 100 mW at 532 nm. Thus a 100 mW reconstruction source will be assumed in the calculation of the possible correlation rate achievable. To assess this rate, losses through the system must be accounted for in addition to those associated with the low diffraction efficiency of the volume hologram. The diffraction efficiency of the A/O scanner will be around 0.80. Kirsch *et al.* [14] have measured the transmission of the Seiko-Epson LCD to be 0.43. There is, however, a further very significant possible energy loss through the system due to the requirement to filter the template images prior to correlation.

5.5.2 Energy losses due to frequency plane filtering

In addition to the losses accounted for above a further absorption of the diffracted wave is required which is best appreciated by considering the spectrum of the reconstructed hologram. If stored as space domain images they may be readily optically Fourier transformed such that their spectra address the phase modulating SLM. Since only the phase of the input spectrum is recorded on this device, the resulting correlation will be similar to a phase only correlation [8] with the difference that the template spectrum amplitude distribution is unattenuated rather than that of the input image. However, this leads to a high sensitivity to mismatches between the input image and reference template due to the unattenuated high spatial frequency components in the spectrum. In addition, the low spatial frequency components also pass through the phase modulating SLM unattenuated. These are of very little use in pattern discrimination and contribute to false cross-correlation peaks. Thus the spectrum of the templates should be bandpassed before addressing the phase modulating SLM. This implies a large reduction in the overall transmitted intensity through the SLM since much of the image power is concentrated in the high intensity, low spatial frequency regions of the spectrum which are removed by the bandpass filter. Similarly, if the images are bandpassed before storage in the volume hologram the resulting edge enhanced imagery will greatly attenuate the reconstruction beam intensity away from the image edge locations.

To alleviate this problem to some degree it is proposed that the templates are stored as frequency domain functions which, upon reconstruction with an appropriately angled reference beam, are imaged to the SLM plane. If the system employs the Seiko-Epson LCD this will require a telephoto lens system to increase the spectrum to the required size. The reference beam amplitude should be such that the grating modulation is maximised in the bandpass of interest. In addition to this, however, the holograms could be illuminated by a difference of Gaussian (DOG) modulated beam [24] so concentrating the available beam power in the required frequency band. The DOG function has several advantages as a bandpass filter as described in Chapter 4. Among these is the property that, since it is composed of Gaussian functions, its Fourier transform does not contain the oscillatory structure that would be generated by a hard-clipped bandpass.

The filter could be implemented directly by modulating the volume hologram read-out wave. However, if done in the manner described below the overall beam energy loss associated with requiring the reconstruction to be in a specific bandpass can be reduced. Since the readout laser will be operating in TEM_{00} , the Gaussian amplitude

profile of the beam can be expanded so as to form the excitatory portion of the DOG filter bandpass when addressing the volume hologram. The centre of the beam may be attenuated with a neutral density filter with a Gaussian fall-off which can simply be made by exposing a photo-emulsion with a TEM₀₀ beam of appropriate size. If the ratio of the Gaussian constants, d_e/d_i is approximately 1.6 a suitable bandpass structure will be generated that approximates the DOG function upon illumination with the read-out beam. The overall beam loss may be calculated from evaluation of the integral :

$$\int_0^\infty 2\pi k \left| \exp\left(-\frac{k^2}{d_e^2}\right) * \left(1 - \exp\left(-\frac{k^2}{d_n^2}\right)\right) \right|^2 dk \quad \{5.1\}$$

where d_e is the distance to the 1/e point of the illuminating beam Gaussian amplitude distribution and d_n is that for the beam used to record the photographic negative. (Note that only an approximation of a Gaussian function can be recorded due to the film's limited dynamic range. To achieve a near zero transmittance in the centre of the Gaussian a high exposure would have to be employed which will compromise the linearity of the photographic process and hence the accuracy in recording the Gaussian fall-off.) If the ratio of 1.6 is desired in the resulting transmitted DOG amplitude distribution, addition of the two exponential mantissa in Equation {5.1} indicates $d_n \approx 0.8d_e$. Evaluation of the integral then shows 26% of the illuminating beam energy is transmitted. A further loss will occur due to amplitude modulation of the beam by the spectrum of the template image within the selected bandpass, resulting in an overall transmission of typically about 10% of the initial beam energy. Although quite severe this energy loss should be compared to that which would be incurred by the uniform illumination of an edge enhanced image. First the Gaussian TEM₀₀ profile of the illuminating beam must be expanded so that it has reasonable uniformity over the area of the image. If a 5% fall in intensity at the edge of the image is allowed, only 5% of the beam energy will be available within the image area. The beam is then modulated by the recorded template which, for strongly edge enhanced imagery, would typically incur a further tenfold decrease in beam energy. The increase in beam power available, using the suggested method, directly translates into an increase in possible template search rate. A more effective solution than that proposed above, although harder and more costly to implement, would be to use a computer generated hologram (CGH) to transform the beam from a Gaussian to a DOG distribution. This could be accomplished almost without power loss provided the CGH was copied into a suitable volume medium such as dichromated gelatin.

5.6 Correlation plane detection and readout

Taking all these losses into account and assuming a 100 mW source, 4 mW will be available at the correlation plane CCD detector before diffraction losses from the volume hologram are considered. Typical rapid frame-rate CCD detectors require between 0.05 and 0.5 $\mu\text{J}/\text{cm}^2$ energy densities for saturation. In the estimate of attainable read-out speed an intermediate sensitivity of 0.25 $\mu\text{J}/\text{cm}^2$ and square pixels 15 μm on a side will be assumed for the CCD. If the correlation is approximated by a relatively localised, uniform intensity distribution in a 4 x 4 pixel area, in order to achieve saturation on the CCD an energy of approximately 10^{-11} J will be required prior to the final Fourier transformation by lens L_2 . Since 4 mW of power is available, a volume hologram diffraction efficiency of 10^{-4} (the value Mok measures with 500 images stored in LiNbO_3) will permit a 40 kHz read-out rate. As the diffracted energy from the volume hologram may be expected to fall as the square of the number of holograms stored [25], the read-out rate falls as the square also. Thus, assuming the quoted attainable diffraction efficiency and the inverse square relationship, a 1000 template library could be read-out at 10 kHz.

However, a critical point to consider is how rapidly the correlation plane signal may be located and measured. A single photodiode cannot be used to detect from which template a correlation response is generated since it would have to be of large enough area (ie. $\approx 0.25 \text{ cm}^2$) to allow for a shift in the correlation peak position. The integrated energy over the area of the photodiode will not vary significantly between a localised correlation response and a more diffused light field characteristic of a poor cross-correlation. What is required for this application is, as pointed out by Anderson *et al.* [26], a photodetector with a non-linear thresholding characteristic that will respond only to a localised relatively high intensity correlation peak but not to a uniform light distribution over the face of the photodiode. The template generating the correlation could then be identified by timing the delay of the response from the initiation of the beam scan. Upon repositioning the reconstruction beam to address this template, a rapid CCD could be used to determine the location of the peak for this single frame. This has the advantage that the relatively slow read-out rate of the large amount of correlation plane data need only be performed once per memory search. Recently the design and fabrication of a thresholding detector array has been undertaken for exactly the requirement described [27]. However, the device is still under development and is not commercially available.

It is thus necessary to consider the template search rate possible using a CCD alone. Current rapid CCD cameras are capable of the read-out of a 512 x 512 array at

around 500 frames/sec. However, some of these devices unload from the CCD array directly through multiple read-out registers. Thus during the 2 msec or so this process requires, the CCD elements are unavailable for light detection. The interline transfer CCD shifts the charges from all the photocells within a column into masked vertical shift registers, a process requiring only 1 μ sec. Thus during the read-out time from these registers the photocells may be integrating the next frame. However, the interline structure results in dead spaces between pixels in the horizontal direction. For the accurate detection of very localised correlation peaks this form of sensor geometry is thus not ideal. A frame transfer CCD consists of a sensor area, with near 100% pixel fill factor, and a masked storage area. Once the integration period is complete the charge distribution is transferred to the storage area in the order of 200 μ sec where it may be read out from the parallel shift registers in a further 2 msec. Thus less than 10% of the frame cycle time is unavailable for charge integration, which together with the 1:1 sampling ratio makes this CCD sensor configuration the most suitable for the present application [28].

If the correlator configuration that employs the Seiko-Epson LCD is considered, 20 msec is available for the template search. Assuming the CCD cycle time is 2.2 msec, 9 read-out cycles may be accomplished during this interval. The mismatch between the 10 kHz rate at which the correlation signals may be generated (in order to give adequate power for detection by the CCD) and the 500 Hz read-out rate of this device can be accommodated by superimposing several correlation responses during a single 2 msec integrate period. The template library can then be initially searched in blocks to determine which set correlates with the object orientation in the current input scene. Once the block in which strong correlation responses have occurred is identified, this may then be searched either with an intermediate sub-block or directly as individual correlations. The integration of multiple template correlations onto a single output frame will lead to a deterioration in the signal to noise ratio (SNR) of the correlation plane. If for an exact template match the ratio of peak correlation signal to mean square noise is taken as 10^3 [26], integration of n frames will reduce this by approximately n times as a worst case. In fact, several templates that closely match the input object may respond in addition to the exact template. Thus if the template memory is arranged with incremental views of object orientation, several correlation responses may sum incoherently on the detector plane to yield a higher, although broader, initial correlation response. In the 2 msec integrate period of the CCD, 20 correlation responses may be integrated to yield an expected correlation peak to mean square noise ratio of 17 dB. This, however, is only to identify the block in which further correlation searches should be made, a final individual template match being used to determine object identity and location. A figure for the overall correlation

rate of the device is complicated by this search scheme. The initial block search may be accomplished at a rate of 4500 template matches per second, a 1000 template memory being searched in a total of 6, 40 msec frame times. A further 2 input frame intervals are required to accomplish the individual correlations within a block. Thus the entire memory may be searched in under $1/3^{\text{rd}}$ of a second ie. an overall rate of 3000 template matches per second. This is, however, a worst case, in which the template match occurs in the last filter of the last block to be searched.

5.7 Filter multiplexing

Synthetic discriminant function techniques [2] may be employed to multiplex several close object orientations into a single template within the memory. This will only lead to a slight increase in the capacity/speed product of the memory since this is governed by the attainable diffraction efficiency from each template. The correlation peak height obtained from a multiplexed filter will be reduced since the available dynamic range of the recording medium is shared between the object views comprising the filter. The situation is slightly improved in that the correlation peak will be the coherent sum of the individual correlations contributed by each view. The advantage of using multiplexed filters is that, for a given number of individual object templates, the number of discrete angles over which the reconstruction beam must be located is reduced, so simplifying the design of the scanning optics. If a 1000 template memory is arranged as 100, 10 fold multiplexed SDF filters, each of these could be allocated a 1 msec integrate period on the CCD output camera. The entire memory could then be searched in 6 input frame times or 0.24 sec. The orientation of the object could probably only be determined to within the ambiguity determined by the ten items stored within a single SDF filter. However, at least in principle, the SDF filters could be designed to give a pre-determined correlation peak height response when one of the items stored in the filter is present at the input. The discriminatory ability of the filter is maintained by designing the SDF in a bandpass as described previously for the individual filters. An adaptation of the SDF filter design procedure for doing this was described in Chapter 2.

5.8 Conclusion

An optical/digital hybrid correlator system has been outlined which allows the potential for a multi-kilohertz reference template search on data acquired at video rates. This is accomplished by storing the reference templates as angle multiplexed volume holograms in a fixable photorefractive material and exploiting the large storage capacity and capability for rapid optical read-out that this affords to perform many correlation searches per input frame. The initial Fourier transform is calculated digitally, the phase portion being applied to a phase modulating SLM. The scheme has the advantage that a relatively rapid response can be achieved with an SLM only capable of video rate operation since this is incorporated into the input arm, rather than the reference arm, of the correlator. However, the anticipated improvements in SLM technology, making available higher resolution and faster response times, will allow an enhancement in performance, as will the rate at which the digital Fourier transforms can be performed which can only be expected to improve. The introduction of higher power short wavelength lasers and the wider availability of large, high quality crystals of photorefractive materials with large Pockels' coefficient together with improved techniques to fix them, will further enhance the read-out rate of volume holographic memories. These advances will permit even faster correlator systems to be developed based on the proposed hybrid architecture.

References to Chapter 5

- [1] A. Vander Lugt, "Signal detection by complex spatial filtering", *IEEE Trans.*, IT-10, 139-145 (1964).
- [2] D. Casasent, "Unified synthetic discriminant function formulation", *Appl. Opt.*, 23 (10), 1620-1627 (1984).
- [3] L. Pichon, J.P. Huignard, "Dynamic joint-Fourier-transform correlator by Bragg diffraction in photorefractive $\text{Bi}_{12}\text{SiO}_{20}$ crystals", *Opt. Comm.*, 36(4), 277-280 (1981).
- [4] D. Psaltis, E.G. Paek, S.S. Venkatesh, "Optical image correlation with a binary spatial light modulator", *Opt. Eng.* 23, 698-704 (1984).
- [5] T.H. Barnes, K. Matsumoto, T. Eiju, K. Matsuda, N. Ooyama, "Phase only liquid crystal light modulator and its application in the Fourier plane of optical correlation systems", in *Optical Pattern Recognition II*, H. J. Caulfield, Ed., *Proc. SPIE*, 1134, 204-208 (1989).
- [6] Full technical details of this and the other related PDSP chip sets are given in the GEC/Plessey publication: "Digital Signal Processing IC Handbook".
- [7] J. Amako, T. Sonehara, "Kinoform using an electrically controlled birefringent liquid-crystal spatial light modulator", *Appl. Opt.* 30(32), 4622-4628 (1991).
- [8] J.L. Horner, P.D. Gianino, "Phase-only matched filtering", *Appl. Opt.* 23(6), 812-816 (1984).
- [9] N. Collings, W.A. Crossland, P.J. Ayliffe, D.G. Vass, I. Underwood, "Evolutionary development of advanced liquid crystal spatial light modulators", *Appl. Opt.* 28(52), 4740-4747 (1989).
- [10] D.A. Jared, R. Turner, K.M. Johnson, "Electrically addressed spatial light modulator that uses a dynamic memory", *Opt. Lett.*, 16(22), 1785-1787 (1991).

- [11] I. Underwood, D.G. Vass, R.M. Sillito, G. Bradford, N.E. Fancey, A. Al-Chalabi, M.J. Birch, W.A. Crossland, A.P. Sparks, S.C. Latham, "A high performance spatial light modulator", in *Devices for Optical Processing*, D.M. Gookin, Ed., *Proc. SPIE*, 1562, 107-115 (1991).
- [12] I. Adulhalim, G. Moddel, K.M. Johnson, "High-speed analog spatial light modulator using hydrogenated amorphous silicon photosensor and an electroclinic liquid crystal", *Appl. Phys. Lett.*, 55(16), 1603-1605 (1989).
- [13] S.E. Day, "An analysis of phase modulation by liquid crystal electro-optic effects", in *Two dimensional opto-electronic device arrays*, IEE Colloquium Group E13, (1991).
- [14] J.C. Kirsch, D.A. Gregory, M.W. Thie, B.K. Jones, "Modulation characteristics of the Epson liquid crystal television", *Opt. Eng.* 31(5), 963-970 (1992).
- [15] M.S. Welkowsky, U. Efron, W.R. Byles, N.W. Goodwin, "Status of the Hughes charge-coupled-device-addressed liquid crystal light valve", *Opt. Eng.* 26(5), 414-417 (1987).
- [16] U. Efron, W.R. Byles, N.W. Goodwin, R.A. Forber, K. Sayyah, C.S. Wu, M.S. Welkowsky, "Charge-coupled-device-addressed liquid-crystal light valve: an update", in *Liquid-Crystal Devices and Materials*, P. S. Drzaic and U. Efron, Eds., *Proc. SPIE*, 1455, 237-247 (1991).
- [17] E.G. Paek, J.R. Wullert, M. Jain, A. Von Lehmen, A. Scherer, J. Harbison, L.T. Florez, H.J. Yoo, R. Martin, J.L. Jewell, Y.H. Lee, "Compact and ultrafast holographic memory using a surface-emitting microlaser array", *Opt. Lett.*, 15(6), 341-343 (1990).
- [18] F. Micheron, G. Bismuth, "Electrical control of fixation and erasure of holographic patterns in ferroelectric materials", *Appl. Phys. Lett.*, 20(2), 79-81 (1972).
- [19] F. Micheron, G. Bismuth, "Field and time thresholds for the electrical fixation of holograms recorded in $(\text{Sr}_{0.75}\text{Ba}_{0.25})\text{Nb}_2\text{O}_6$ crystals", *Appl. Phys. Lett.*, 23(2), 71-72, (1973).

- [20] J.J. Amodei, D.L. Staebler, "Holographic pattern fixing in electro-optic crystals", *Appl. Phys. Lett.* 18(12), 540-542 (1971).
- [21] F.H. Mok, M.C. Tackett, H.M. Stoll, "Storage of 500 high resolution holograms in a LiNbO₃ crystal", *Opt. Lett.*, 16(8), 605-607, (1991).
- [22] F.H. Mok, H.M. Stoll, "Holographic inner-product processor for pattern recognition", in *Optical Pattern Recognition III*, D. P. Casasent and T.-H. Chao, Eds., *Proc. SPIE* 1701, 312-322 (1992).
- [23] H.C. K ulich, "Transfer function for image formation of objects reconstructed from volume holograms with different wavelengths", *Appl. Opt.* 31(14), 2461-2477 (1992).
- [24] D. Marr, E. Hildreth, "Theory of edge detection", *Proc. R. Soc. Lond. B*, 207, 187-217 (1980).
- [25] J.H. Hong, P. Yeh, D. Psaltis, D. Brady, "Diffraction efficiency of strong volume holograms", *Opt. Lett.*, 15(6), 344-346 (1990).
- [26] G.W. Anderson, B.D. Guenther, J.A. Hynecek, R.J. Keyes, A. Vander Lugt, "Role of photodetectors in optical signal processing", *Appl. Opt.* 27(14), 2871-2886 (1988).
- [27] H. Langenbacher, T.-H. Chao, T. Shaw, J. Yu, "64 x 64 thresholding photodetector array for optical pattern recognition" in *Optical Pattern Recognition IV*, D. Casasent, T.-H. Chao Eds., *Proc. SPIE*, 1959, (1993)
- [28] For example the recently available David Sarnoff Research Centre VCCD512H 512 x 512, 500 frames/sec. CCD camera.