

# **Target Characteristics and Cross-Correlation Detectors**

by

**Norman Young**

A thesis submitted in conformity with the requirements  
for the degree of Master of Applied Science  
Graduate Department of Electrical and Computer Engineering  
and the Institute of Biomedical Engineering  
University of Toronto

© by Norman Young 1998



**National Library  
of Canada**

**Acquisitions and  
Bibliographic Services**

**395 Wellington Street  
Ottawa ON K1A 0N4  
Canada**

**Bibliothèque nationale  
du Canada**

**Acquisitions et  
services bibliographiques**

**395, rue Wellington  
Ottawa ON K1A 0N4  
Canada**

*Your file Votre référence*

*Our file Notre référence*

**The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.**

**The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.**

**L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.**

**L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.**

0-612-34142-9

# **Target Characteristics and Cross-Correlation Detectors**

Master of Applied Science 1998

Norman Young

Graduate Department of Electrical and Computer Engineering  
and the Institute of Biomedical Engineering  
University of Toronto

## **Abstract**

Tracking a set of stationary targets in the field of view of a head-mounted camera can provide an estimate of head motion. In this thesis, a theoretical model is developed relating the probability of false detection (POFD) of a stationary target by a cross-correlation detector, to the size, bandwidth and power of the target and spatially disjoint scene noise. Closed-form expressions are developed to provide insight into the fundamental processes that affect the POFD. It is shown, through modelling and computer simulations that 1) by reducing the difference in area between the image and target and by increasing the input SNR, the POFD is lowered and 2) that increasing either the signal or noise bandwidth results in a POFD curve which rises initially, reaches a maximum and then falls. The POFD is shown to be governed by the signal and noise peak variance and the number of independent noise peaks. The results provide a foundation for the examination of “optimal” target selection criteria when targets suffer from additive noise or geometric distortions.

# Acknowledgments

I have come to realize that the meaningfulness of my life is very much shaped—perhaps even defined—by my relationships. I have been blessed with an abundance of such relationships, both professional and personal. The detail included in these acknowledgements may be considered by some as inappropriate, perhaps even unprofessional. Or just plain tacky. However, I feel that simply thanking a list of people without recognizing specifically how each individual has uniquely contributed to my life, would be generic and platitudinous and would fail to convey their true importance to me.

So then, thanks to ...

My supervisor, **Prof. Moshe Eizenman**, who from the very beginning, set the bar high. (I won't ever forget what he said in our very first meeting: "I'll push you, but I won't break you.") He has taught me the necessity of understanding the fundamentals and the importance of never losing sight of the big picture. His sustained, active involvement in this project is very much appreciated. **Steve**, whose encouragement and expertise are why I'm not still running fruitless overnight computer simulations. I am grateful for his friendship and the many truly interesting chats we've had over the past couple of years, as well as his patience during my monopoly of the lab computer. **Thas**, who sac-

rificed a great deal of his time ensuring the computers were working, answered my litany of Linux questions, and on top of that, listened to and encouraged me during the bumpier parts of the ride. **Peter**, who not only provided technical advice, but also the encouragement that I needed early on in this project. **Nathanael**, who entertained my really stupid probability questions and who was always available to listen to my concerns. **Kath**, who always thought more of my ability than I did. Her natural ability to provide encouragement and empathy is perhaps only surpassed by her insight and compassion. Her friendship has been tremendously edifying. **Brian**, whose commitment to a truthful and honest telling of—and response to—reality has been an example to me. The many chats we've shared have been both encouraging and stimulating and have played an integral role in developing my continually evolving view of life and faith.

**Winston**, whose cool, candid objectivity is rare and insight into the human condition unique. His friendship I treasure deeply. **Grant**, whom I have known since childhood. Our friendship deepens with each passing year and his faithfulness to it has always been more than just words. **Alfred and Anna**, who have seen me through the very toughest of times and throughout it all have been unwavering in their concern, support and encouragement. **Gary**, whose drive and determination to make the most of what he's been given has been most admirable. **Janet**, who has the unique ability to brighten up a room like no other. The frequency and sincerity of her words of encouragement and affirmation, I appreciate very much. **Rebecca**, whose faithful stewardship of her many talents has been an example to me. The genuineness of her care and concern, I value a great deal. **Cindy**, whose deep commitment to integrity I respect very much. Her encouragement during this endeavour has been a

source of much comfort. **Jasmine**, who, perhaps better than anyone else, knows what “Life in Grad School” is like. Her empathy and understanding has been tremendously encouraging. **Tina**, who is the closest thing I have to a little sister. Her cheerfulness and support has meant much to me. **Anita**, who truly is one of a kind. Her passion and compassion has been inspiring.

**May**, who taught me the meaning of selflessness . . . I will not forget you.

Finally, my **mother**, whom I know has always said and done everything out of love. Her influence on me will last a lifetime.

“Whatever you do, work at it with all your heart, as working for the Lord . . . it is the Lord Christ you are serving.”

Colossians 3:23,24

NORMAN YOUNG

*University of Toronto*

*March 1998*

# Contents

<b>Abstract</b>	<b>ii</b>
<b>Acknowledgments</b>	<b>iii</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Problem Definition . . . . .	3
1.3 Background . . . . .	5
1.3.1 Pattern Recognition . . . . .	5
1.3.2 Correlation Filters . . . . .	6
1.3.3 Towards Distortion Invariance . . . . .	11
1.3.4 Improving Peak Localization . . . . .	14
1.3.5 Image Characteristics . . . . .	15
1.4 Overview of Approach . . . . .	17
<b>Chapter 2 Theoretical Development</b>	<b>19</b>
2.1 Image Model . . . . .	19
2.2 Cross-Correlation . . . . .	23
2.3 Probability of False Detection . . . . .	27
2.4 Summary . . . . .	34

<b>Chapter 3 Statistical Simulation</b>	<b>36</b>
3.1 Technical Description . . . . .	36
3.2 Simulation Results . . . . .	39
3.3 Model Limitations . . . . .	48
3.4 Summary . . . . .	50
<b>Chapter 4 Considerations in Real Images</b>	<b>51</b>
4.1 Target Selection . . . . .	51
4.2 Performance Evaluation . . . . .	55
4.3 Summary . . . . .	60
<b>Chapter 5 Conclusions</b>	<b>62</b>
5.1 Summary of Work . . . . .	62
5.2 Future Work . . . . .	64
<b>Appendix A Derivations</b>	<b>66</b>
A.1 Synthesis of $s[\bar{n}]$ and $\mathbf{n}[\bar{n}]$ . . . . .	66
A.2 Mean of $\mathbf{c}[\bar{n}]$ . . . . .	68
A.3 Autocorrelation of $\mathbf{c}[\bar{n}]$ . . . . .	69
A.4 Bandwidth of $\bar{\mathbf{n}}[\bar{n}] \bar{\mathbf{s}}[\bar{n}]$ . . . . .	76
A.5 1-D Equations . . . . .	77
<b>Bibliography</b>	<b>78</b>

# Chapter 1

## Introduction

### 1.1 Motivation

There exist numerous contexts in which knowledge of a subject's gaze in space over time, or *scan path*, is of value (Figure 1.1). In neurology, clinical tests based on scan path analysis have been used for diagnostic purposes, since anomalies in a patient's eye movements have been related to certain neurological disorders [59]. In flight training, recordings of the visual behaviour of expert pilots in the cockpit have served as pedagogical examples in the instruction of novice pilots [58]. In research psychology, the concurrent recording of a subject's scan path and verbal description of so-called "ink blots" have allowed researchers to investigate the correlation between what is said and what is seen.

The determination of gaze in space requires knowledge of both the eye position and head position. A number of methods for measuring eye position exist. Among them is *electro-oculography* (EOG) [59] which measures eye position by recording the potential differences between two skin electrodes

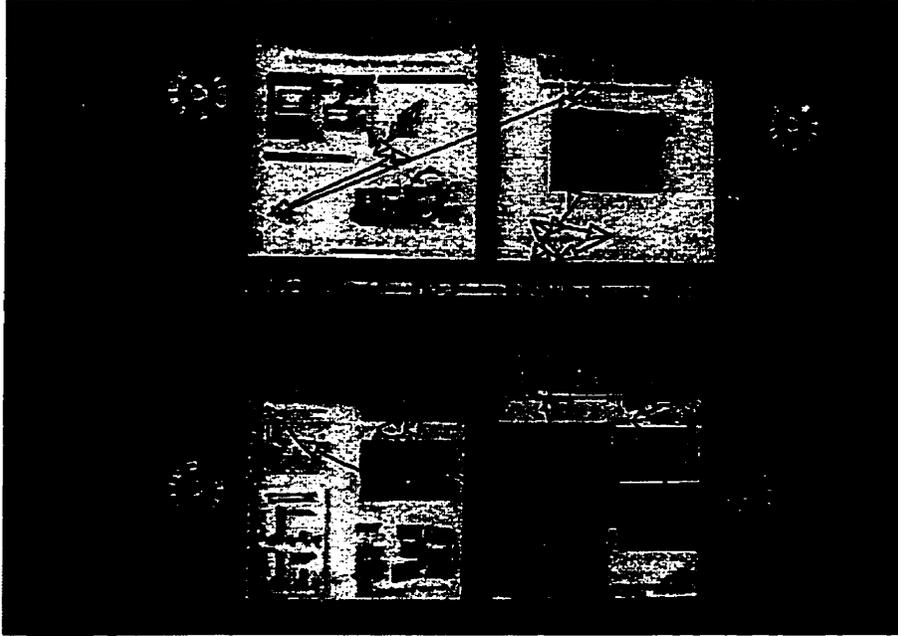


Figure 1.1: A sample scan path. 'S' denotes start position, 'E' denotes end position. Artificial markers currently being used are also shown.

placed near the eye. Although it is the standard clinical technique, EOG suffers from poor sensitivity and a host of noise related problems, such as myographic artifacts. Another technique [2, 59] uses a head-mounted CCD camera directed at the eye to capture an image of the pupil and corneal reflections that are created by the accompanying infrared LEDs. By measuring the position of the pupil relative to the reflections, an estimate of the eye position relative to the head may be obtained.

Head position may also be estimated by one of a number of methods. One approach involves mounting inertial position sensors [59] such as accelerometers or gyroscopes on a helmet to measure the acceleration in each of the three orthogonal directions as a function of time. From those measurements, the head motion may then be reconstructed. Because acceleration data is used, this method can be susceptible to noise. Another method known as

*magnetic head tracking* [2] uses a stationary transmitter that emits a pulsed magnetic field which is sensed by a head-mounted receiver. The signal is then processed to yield an estimate of the head position. Because measurements are sensitive to electromagnetic interference and the presence of large metal objects, this method can be unreliable in certain environments. A third alternative—and the one explored in this thesis—is to estimate head motion from the apparent “movements” of stationary objects in the field of view (FOV) of a head mounted scene camera. One system currently being tested that uses this method involves placing specially designed markers (Figure 1.1) in the FOV of the camera to establish a fixed frame of reference [58]. By tracking their movements within the FOV, an estimate of the head position may be found. By combining this information with the eye-position information, (made relative to the head) an estimate of the subject’s gaze in space may be formed. A shortcoming of this approach however, is that the markers need to be manually placed in the FOV. Clearly, in some situations, this may not be convenient or even possible.

In this thesis, we begin to examine a natural extension of the above approach that does not require the placement of artificial markers in the FOV. Instead, sufficiently “well-defined” objects which occur naturally in the FOV of the scene camera are utilized as frame of reference markers. It is the notion of a “well-defined” object that we endeavour to give meaning to.

## **1.2 Problem Definition**

The problem described above is essentially one of pattern recognition (PR) involving the detection of a number of targets amidst an arbitrary background in each frame of a dynamic image sequence and the estimation of their re-

spective positions. This task is complicated by a number of factors: 1) scene camera motion due to head movements results in targets being subject to geometric distortions which may be approximated as translations, rotations and scalings; 2) targets may temporarily stray beyond the camera's field of view, thus necessitating their recapture; 3) background clutter can be mistaken for the true target; and 4) as the system is moved from one environment to another, changes in ambient illumination cause targets to undergo intensity variations. Any detection scheme employed must therefore be sufficiently robust to account for these factors.

The task maybe subdivided into a static problem and a dynamic problem. The aim of the static problem is to detect the presence of an object(s) *somewhere* in a larger, still image and to estimate its position in that image. In the dynamic problem, the emphasis is on estimating and *predicting* the position of the object(s) over time, given constraints such as continuity. For the purposes of this thesis, we consider only the static problem, leaving the dynamic problem for future work.

The static problem posed by our particular application differs slightly from its more traditional casting. Typically, the goal is to detect the presence of an object which has been determined *a priori* (e.g. aircraft identification). Here, however, because *any* "easily detectable" object in the FOV can serve as a frame of reference marker, the goal is to select the best possible objects for tracking. That is, we are afforded the luxury of selecting the target that is most amenable to detection and tracking. The fundamental issue then—and the focus of this thesis—is *what characterizes a optimal target*. We distill this idea further after a brief survey of the relevant background and literature.

## 1.3 Background

### 1.3.1 Pattern Recognition

A necessary step in any PR system is to measure the similarity between an image of interest (i.e. an object) contained in a larger scene against a reference image or *template* of some kind. Although raw image data may be compared directly, it is often more useful and robust to represent the pattern in a more compact manner with functions of the pattern itself, known as *features*. “Good” features not only eliminate the large amount of redundant information present in most natural patterns [33], but also retain a sufficient amount of *essential* information thereby allowing a pattern to be uniquely identified.

Feature design varies greatly depending on what maybe reasonably assumed about the class of patterns to be recognized. *Structural* features attempt to describe a pattern in a heuristic manner, drawing on concepts such as shape [46], curvature [47, 48] and topology [46]. Much of structural PR has focussed on developing formal codes and “grammars” [40] that are able to describe patterns which share a similar topological or curve structure (e.g. handwritten characters, closed contours) in terms of a set of primitives. Although elegant, the ability to describe only a narrow class of patterns compounded by the frequent requirement that the input images be binary and free of any background, suggest that structural features are less effective in describing naturally occurring patterns about which little may be reasonably assumed. For this reason, our consideration of them ends here.

The second class of features, *statistical* features, are based strictly on mathematical functions of a pattern, and can be somewhat more effective in

representing more arbitrary images. Many statistical techniques have been developed, including the numerous transform domain methods (e.g. Fourier, Radon, Hadamard, etc.) and Hu’s moment invariants [25] [18, 56], although it has been shown that higher order moments suffer from noise sensitivity and information loss [1]. Despite such variety, however, *correlation-based detection filters*, and its numerous variants and descendants, remain predominant. They have been—and still are—the subject of ongoing research even after their initial introduction to PR over 30 years ago. They typically make few assumptions about the input image, generally do not require substantial image preprocessing (e.g. segmentation) and they often have efficient optical and/or digital implementations. As we shall see later, some variants also claim rotation and scale invariance. For these reasons, they are often good candidates for use in real-time PR problems, such as ours, in which little is known about the image *a priori* and geometric distortions are a factor. We pursue them further in the following section, where we develop some of their underlying theory.

### 1.3.2 Correlation Filters

#### Theoretical Basis

The simple binary detection problem involves determining whether or not embedded within a received signal,  $r[\vec{n}]$ , there exists a signal of interest or template,  $s[\vec{n}]$ , at a known position,  $\vec{n} = \vec{n}_0$ , where  $\vec{n} = [n_x, n_y]$  is a two-dimensional, discrete spatial variable. The basic estimation problem, on the other hand, involves determining the position of  $s[\vec{n}]$  within  $r[\vec{n}]$ , *assuming* that  $s[\vec{n}]$  is indeed present. However, in many applications such as ours, neither the signal’s presence or location is known with certainty. Thus, we are concerned with both detection *and* estimation.

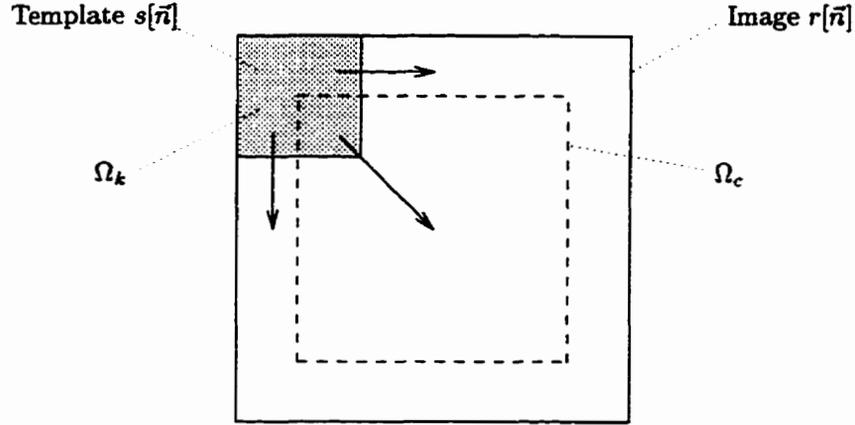


Figure 1.2: Template slides over image. Dotted region denotes domain of template position.

The problem may be approached by first measuring the similarity between the template and every template-sized subregion in the received image. This may be achieved by “sliding” the template over the image (Figure 1.2) and computing the  $L^2$  norm—or mean-square error (MSE)—between  $s[\vec{n}]$  and  $r[\vec{n}]$ , for all shifts,  $\vec{n}$ , of the template [14, 27]. That is,

$$\epsilon^2[\vec{n}] = \sum_{\vec{k} \in \Omega_k} |r[\vec{k}] - s[\vec{k} - \vec{n}]|^2, \quad \vec{n} \in \Omega_c \quad (1.1)$$

where the summation denotes a double summation over the vector  $\vec{k}$ , and  $\vec{k}$  spans the domain of the template,  $\Omega_k$ . The domain of the template is in turn dependent on the position of the template,  $\vec{n}$ . Note that  $\vec{n}$  is such that the template is always completely within the boundaries of the image<sup>1</sup> (Figure 1.2), i.e.,  $\vec{n} \in \Omega_c$ , so that it and the subregion it is being compared to is the same size; the MSE metric is meaningless if they are not. In principle then, finding the shift  $\vec{n} = \vec{n}_0$  at which  $\epsilon[\vec{n}]$  is a minimum provides an estimate of the position of the signal, which is deemed present if the error at that point

<sup>1</sup>For clarity, we intentionally defer the explicit definitions of  $\Omega_k$  and  $\Omega_c$  to Chapter 2

is below a preset threshold, or deemed absent, if it is not. Computationally, equation (1.1) can be made more efficient. Expanding (1.1) yields,

$$\epsilon^2[\vec{n}] = \sum_{\vec{k} \in \Omega_k} |r[\vec{k}]|^2 + \sum_{\vec{k} \in \Omega_k} |s[\vec{k} - \vec{n}]|^2 - 2 \sum_{\vec{k} \in \Omega_k} r[\vec{k}]s[\vec{k} - \vec{n}]. \quad (1.2)$$

We immediately recognize that the last term is simply the spatial cross-correlation,

$$c[\vec{n}] = \sum_{\vec{k} \in \Omega_k} r[\vec{k}]s[\vec{k} - \vec{n}]. \quad (1.3)$$

This is significant for two reasons: Firstly, it can be seen that by assuming that the first two terms in (1.2)—the local image and template energies respectively—are always constant, maximizing  $c[\vec{n}]$  is tantamount to minimizing  $\epsilon^2[\vec{n}]$ . Thus, a cross-correlation peak, if above a preset threshold, indicates the presence of the signal of interest and its location provides an estimate of the signal’s position. Secondly, a cross-correlation may be efficiently computed by optical correlation [41], fast fourier transform (via the correlation theorem) [4], or even statistical methods [3], making it suitable for real-time applications. A couple of heuristic comments are in order:

- Because correlation and convolution are related simply by a time reversal of the kernel, Equation 1.3 may be seen as a *filtering* operation in which the time-reversed signal is the impulse response of the filter; hence the term “correlation filter”. Furthermore, because the filter kernel is “matched” to the signal of interest, cross-correlation is also frequently referred to as *classical matched filtering* [57].
- Cross-correlation may also be thought of as a *projection* operation [52]. Equation 1.3 is essentially an inner product, and thus, the template

vector may be thought of as being projected on to the received image vector. If they are perfectly correlated (i.e. identical) at a template shift  $\vec{n}$ , the “angle” between them is zero and the correlation coefficient  $c[\vec{n}]$  is unity. If, at the other extreme, they are perfectly uncorrelated, they are then orthogonal and the correlation coefficient is zero.

### Drawbacks

Despite its merits, however, cross-correlation is not without its share of caveats. Firstly, maximizing the cross-correlation does not *always* minimize the MSE, and thus its maximum value, even if above threshold, is not *necessarily* indicative of minimum MSE. We recall that that claim was made under the provision that the first two terms in (1.2) are constant. It is easily seen that the template energy is constant; however, because the range of  $\vec{k}$  is dependent on the position of the template, the local image energy in general, can vary. The consequence of this may be illustrated by considering a situation in which the correlation is computed between a template and two separate image regions, one which matches the template identically but is dim (low energy), and one that does not but is bright (high energy). It is not difficult to see that the cross-correlation of the template with the high energy region will be higher than that with the low energy region even if the match is poor, thus suggesting that a correlation coefficient can be high, without the associated MSE being necessarily small. A common solution to this problem is to normalize the cross-correlation by dividing it by the local image energy at every point [12, 14]. That is,

$$c[\vec{n}] = \frac{\sum_{\vec{k} \in \Omega_k} r[\vec{k}]s[\vec{k} - \vec{n}]}{\sum_{\vec{k} \in \Omega_k} |r[\vec{k}]|^2}. \quad (1.4)$$

This operation, however, is very computationally expensive, for it requires that the energy of every template-sized sub-region in the image be calculated, making it impractical for real-time applications. Furthermore, if the image is modelled as a random field, the estimation error (i.e. the variance) associated with the numerator is compounded by that of the denominator, resulting in a greater overall estimation error in the cross-correlation than if it had not been normalized.

Even when the local image and template energies can be considered constant, a problem still remains: target detection is only unique when the cross-correlation maximum is exactly unity. That is, when the peak value is even slightly less than one, there exists in principle, more than one pattern that will map to that value. Consider, for example, a  $10 \times 10$  template that consists entirely of 1's. Cross-correlating the template with a  $10 \times 10$  received signal also consisting entirely of 1's yields a correlation coefficient of 1 (after normalization). Detection is unique; only one signal—an exact replica of the template—can produce a perfect match. However, if the received signal consists of 99 1's and a single 0, yielding a correlation coefficient of 0.99, the location of that single zero is irrelevant—all 100 versions of the signal will map to the same result. This is because cross-correlation is essentially an *energy* measure—it is insensitive to the locations of the various pixel values in the domain of the summation. It is not difficult to see that as the correlation peak decreases, the number of possible signals that will lead to the same peak value grows very quickly.

A third and significant limitation of cross-correlation is its sensitivity to geometric distortions. From (1.3), we can see that cross-correlation is equally sensitive to every pixel in both the test image and the template. This property is somewhat of a double-edged sword. It is desirable, because it is sensitive

to the fine detail in an image; it is unwanted, because the height of the peak can drop off dramatically even with the smallest perturbations in image scale and orientation. Using an aerial image as input, Casasent and Psaltis [8] showed that the output signal-to-noise ratio (SNR) of a classical matched filter decreased from 30 dB to 3 dB when the image was scaled by 2% and from 30 dB to 3 dB when it was rotated by 3.5°. A similar experiment by Gianino and Horner [19] using a simple binary image showed the correlation peak decreased by 30% when it was rotated by 10° and by 50% when it was scaled down by 10%. As we see in the chapters to follow, such substantial losses ultimately manifest themselves as detection errors.

Finally, due to the dominance of low frequency spectral components in natural images, cross-correlation peaks have often been criticized for being broad and poorly localized [23, 52], thus hindering accurate estimation of the peak location.

### 1.3.3 Towards Distortion Invariance

Much effort has been expended in addressing the limitations associated with cross-correlation. In particular, a great deal of attention has been paid to improving cross-correlation's resistance to geometric distortions such as in-plane rotations and scalings. One approach involves exhaustively searching all possible template translations, orientations and scalings and determining which combination provides the best "fit" [10, 27]. In effect, this amounts to performing a 4-dimensional cross-correlation—a very computationally laborious task.

## Fourier Mellin Transform

A somewhat more elegant technique, known as the *Mellin* transform [3, 8, 9, 53, 54, 55], involves mapping the image to log-polar (i.e.  $(\log r, \theta)$ ) coordinates. The use of polar coordinates naturally leads to rotation invariance; rotations are registered as simple translations along the  $\theta$  axis. Scale invariance is achieved by using a logarithmic scale along the radial coordinate. Because the logarithm of a product is the sum of the logarithms, scalings are also reduced to mere translations. Cross-correlation, be it classical or phase-only [11] (see following section), is then performed in the log-polar domain. Further translation invariance can be achieved by applying the Mellin transform to the magnitude of the Fourier transform of the image [3], which is independent of position.

In principle, the continuous Fourier-Mellin transform is an elegant solution to improving cross-correlation's resistance to geometric distortions. However, implementation of the discrete Fourier-Mellin transform can be rather tedious, due to problems relating to finite image size and the non-uniform sampling inherent to the log-polar grid [3, 11].

## Synthetic Discriminant Functions

A third method that has gained much attention since its introduction by Casasent, involves the synthesis of a general-distortion invariant, composite detection filter known as a *Synthetic Discriminant Function* (SDF) [6, 7, 32, 34, 38, 42]. SDFs are composed of a linear combination of training templates<sup>2</sup>, each representing the object at a different distortion (e.g. a different orientation and/or scaling.) The weights corresponding to each training template are

---

<sup>2</sup>For this reason, SDFs are also known as *composite* filters.

chosen such that the projection (i.e. correlation) of the SDF on an unknown image is unity if the image corresponds to one of the templates, and zero if it does not. The principle behind SDFs is similar to that of an exhaustive search over all cross-correlation templates (described above), in that they both attempt to match the image of interest against a number of distorted versions of a template. The difference however, is that SDFs are formed *offline* and thus the very computationally expensive task of searching the parameter space online is avoided, making real-time pattern recognition more feasible. Furthermore, because an SDF is formed from a set of training images, the distortion may be of any form—non-linear, 3-D, etc. [37]. It need not be a well-defined coordinate mapping.

While Casasent's early work provided a rigorous framework for the basic SDF concept, it did not consider the effect of noise. Eventually however, an analytic treatment of the impact of noise on SDF performance was carried out by Kumar [39]. An SDF composed of a number of training images at different rotations was constructed. A noisy version of the image was then correlated with the SDF at several different rotations. As expected, the output SNR of the filter was oscillatory with the rotation of the noisy image; maxima at the rotations corresponding to those of the training images, minima in between. A key conclusion of that study was that the worst-case output SNR (i.e. at the minima) can be increased by increasing the number of training set images, but only at the expense of decreasing the best-case output SNR (i.e. at the maxima).

### 1.3.4 Improving Peak Localization

#### Phase-Only Filters

The task of improving peak localization has also been dealt with at length by a number of authors. Phase-only filters (POF) [10, 11, 23] normalize the magnitude spectrum of the input image to unity across all frequencies by dividing its fourier transform by its magnitude, leaving only the phase information. The result is a whitened signal, which when cross-correlated with the template, yields a very sharp peak, leading to a better estimate of the peak location. Because the remaining phase information has been shown to be very important in encoding highly informative shape information [49], phase correlation is also effective in improving discrimination between objects [23]. A significant drawback of POFs however, is that they are very sensitive to geometric distortions. Because they have a whitening effect, they increase the amplitude of high frequency components that are otherwise negligible. This spectral broadening results in an increase in the filter's space-bandwidth product (SBWP), which has been shown to increase the rate at which the correlation falls with distortion [8]. The consequence then, is that phase-only filters are in general, less resistant to geometric distortions than classical matched filters [19].

A close cousin of the phase-only filter is the *binary* phase-only filter (BPOF) [13, 16, 17, 24]. The principle behind the BPOF is the same as that of the POF, but instead of a continuum of phase values, BPOF phase information is quantized to two bits (i.e. binarized). The motivation for the BPOF is largely practical, since much of the optical and electronic hardware used to generate phase-only filters work particularly well in binary phase-only mode [24].

## Other Techniques

Synthetic discriminant function based techniques, such as the minimum variance [38] and minimum average correlation energy filters [42], have also been proposed to enhance peak localization by suppressing the intensity of correlation sidelobes. More recently, we have seen the introduction of filters which are perhaps best classified as descendants of correlation-based filters, but are not correlation filters *per se*. These filters also attempt to increase the signal output, while suppressing unwanted noise. The method of *Nonorthogonal Image Expansion* [5, 51, 52], proposed by Ben-Arie and Rao has been shown to yield very sharp peaks with relatively low sidelobe response, as have the optimal filters introduced by Javidi *et al* [29, 30, 31]. These filters often improve output response by utilizing some statistic of the test image itself, such as the power spectrum or noise mean.

### 1.3.5 Image Characteristics

In the above discussion, we alluded to how certain image characteristics, such as spatial energy distribution and spectral content, can influence the detection performance of a correlation filter. Indeed, the relationship between a signal and its processor is intimate; one cannot examine the efficacy of the processor without regard for the characteristics of the signal, nor can one determine relevant signal characteristics without regard for the processor. Having briefly examined some of the properties of the cross-correlator, we now turn our attention to some relevant signal characteristics.

A received image is frequently modelled as the sum of a random<sup>3</sup> signal

---

<sup>3</sup>We adopt Papoulis' [50] typographical convention of boldfacing all random variables. (e.g.  $\mathbf{s}[\vec{n}]$ )

centered at  $\vec{n} = \vec{n}_0$  and some form of random noise,  $n[\vec{n}]$ :

$$\bar{r}[\vec{n}] = \bar{s}[\vec{n} - \vec{n}_0] + \bar{n}[\vec{n}], \quad (1.5)$$

For purposes of analytic tractability, it has often been assumed explicitly or otherwise, that the noise in (1.5) satisfies one or more of the following conditions: 1) it overlaps the target [36, 39, 45], 2) it is uncorrelated with the target [45], 3) it is characterized by stationary statistics [45], and/or 4) it is white [39]. While such assumptions may be useful for modeling sensor noise or simple backgrounds such as terrain [21], they often do not hold in real images. A key observation made by Javidi and Wang was that in many PR problems, the unwanted noise is not usually the random system noise associated with the imaging system that is often modeled, but rather the *scene noise* [28]. Anything in an image that is not part of the target, such as background clutter, other objects, etc., may be collectively referred to as scene noise.

Because it can be so varied in nature, scene noise is difficult to model, especially if the image class is unspecified. However, a few general comments can be made about it. Firstly, it is *spatially disjoint* from the target [28, 29]. That is, it does not overlap the target. Secondly, because it often contains signal-like objects, regions within the scene noise are often correlated with the target. Thirdly, it is never white—real images are generally low-pass and thus neighbouring pixels are correlated. Finally, scene noise is non-stationary—its first and second order statistics are typically space-dependent. We shall attempt to describe some of these characteristics in a more quantitative manner in Chapter 2, where we employ a mathematical image model in our examination of the detection performance of the cross-correlator.

## 1.4 Overview of Approach

We recall from Section 1.2 that because our application is such that the targets to be detected are not prescribed but may be selected, we are interested in what characterizes an *optimal* target. At that point, however, we were rather ambiguous as to what we meant by “optimal” nor had we specified a processor. We now give our problem a more precise definition.

Numerous detector performance metrics have been suggested in the literature [22, 29, 37, 45, 52, 57]. Amongst the more common ones are the classical output SNR [57], the peak-to-sidelobe ratio [45] and the peak-to-output-energy ratio (POE) [29]. We refrain from a lengthy treatment of all these measures here. Suffice to say, however, that almost all of these definitions are ratios of some output signal characteristic to some output noise characteristic. In the end, they are all rather arbitrary and ultimately, they all indirectly attempt to describe a most basic idea: how often the detector will fail, that is, the probability of false detection (POFD). It is in this sense—the POFD sense—which we define “optimality”.

Turning our attention to the processor, we recall that at the close of section 1.3.1, we provided some rationale for focusing on correlation-based methods. Because cross-correlation forms the basis for many such methods, it would seem that it is worthy of some theoretical study. However, as we saw in sections 1.3.3 and 1.3.4, most of the work has been devoted to introducing new techniques that address specific problems associated with cross-correlation. In our survey of the literature, some authors have attempted to provide a theoretical basis for spatial cross-correlation performance [15, 26, 28, 35, 36, 37, 39, 43, 44, 45]. Of those however, only Mostfavi/Smith [44] and Ianiello [26] dealt with performance in terms of false detection probabilities. We base

some of our analysis on their work, as well as Javidi and Wang's [28], who introduced the concept of spatially disjoint scene noise.

Implicit in our notion of a "optimal target" is that the characteristics of the target and those of the scene noise may be selected such that the probability of false detection is made small. Characteristics of particular relevance to our application are the

- size, bandwidth and power of the target and scene noise
- correlation between target and scene noise
- image orientation and scale.

In this thesis, we pursue the first of these characteristics. Through the development of a theoretical model, (Chapter 2) and its verification through computer simulations, (Chapter 3) we attempt to gain some insight into how the various target and scene noise characteristics influence the false detection rate. Results from those studies are then used to select "optimal" targets from real images (Chapter 4).

# Chapter 2

## Theoretical Development

In this chapter, we examine how the spatial characteristics of the target in relation to those of the scene influence the fundamental processes that affect the probability of false detection. Through the use of a stochastic image model, we are able to derive closed-form expressions that lend insight into the nature of those processes. The analysis can also serve as a foundation for studying the effects of other factors—such as geometric distortions—on detection performance.

### 2.1 Image Model

We model a discrete<sup>1</sup>, random received image,  $\bar{\mathbf{r}}[\vec{n}]$ , as the sum of a random target signal,  $\bar{\mathbf{s}}[\vec{n}]$ , and random, spatially disjoint scene noise,  $\bar{\mathbf{n}}[\vec{n}]$ , which surrounds the target:

$$\bar{\mathbf{r}}[\vec{n}] = \bar{\mathbf{s}}[\vec{n}] + \bar{\mathbf{n}}[\vec{n}]. \quad (2.1)$$

---

<sup>1</sup>That is,  $s[\vec{n}] = s_c(\vec{n} \Delta x)$ , where  $s_c$  is a continuous signal and  $\Delta x$  is the sampling period.

Note that  $\bar{\mathbf{s}}[\vec{n}]$  is centered at  $\vec{n} = \vec{0}$  with no loss of generality. To simplify the analysis,  $\bar{\mathbf{s}}[\vec{n}]$  and  $\bar{\mathbf{n}}[\vec{n}]$  are assumed to be uncorrelated. The signal, or template, may be written as

$$\bar{\mathbf{s}}[\vec{n}] = \mathbf{s}[\vec{n}] w_s[\vec{n}], \quad (2.2)$$

where  $\mathbf{s}[\vec{n}]$  is a zero-mean<sup>2</sup>, Gaussian, stationary signal with autocorrelation  $R_{ss}[\vec{n}]$ , and  $w_s[\vec{n}]$  is a window that defines its extent:

$$w_s[\vec{n}] = \begin{cases} 1 & \vec{n} \in \Omega_s \\ 0 & \text{otherwise.} \end{cases} \quad (2.3)$$

The region  $\Omega_s$  may be of arbitrary shape; however, to maintain analytic tractability, we assume that it is an  $N_s \times N_s$  sample square<sup>3</sup> (Figure 2.1(a)):

$$\Omega_s = \left[-\frac{N_s}{2}, \frac{N_s}{2} - 1\right] \times \left[-\frac{N_s}{2}, \frac{N_s}{2} - 1\right]. \quad (2.4)$$

The scene noise may be similarly written as

$$\bar{\mathbf{n}}[\vec{n}] = \mathbf{n}[\vec{n}] w_n[\vec{n}], \quad (2.5)$$

where  $\mathbf{n}[\vec{n}]$  is zero-mean, Gaussian, stationary noise with autocorrelation  $R_{nn}[\vec{n}]$ , and  $w_n[\vec{n}]$  is a window that defines its extent. Because the scene noise does not overlap the signal, but surrounds it, the noise window is written as

$$w_n[\vec{n}] = w_r[\vec{n}] - w_s[\vec{n}], \quad (2.6)$$

---

<sup>2</sup>Since any DC bias may be easily subtracted as a part of preprocessing, we may, without any loss in generality, simplify our analysis by assuming that all images are zero-mean.

<sup>3</sup>Because  $n_x, n_y \in I$ , this domain definition is problematic when  $N_s$  is odd, since  $\frac{N_s}{2}$  would be non-integer. To remedy this—and to avoid introducing any more notation—we simply assume that all fractions of this kind—where the quotient is necessarily integer—are rounded to the next highest integer, i.e.  $\frac{N}{2} \equiv \text{ceil}(\frac{N}{2})$ .

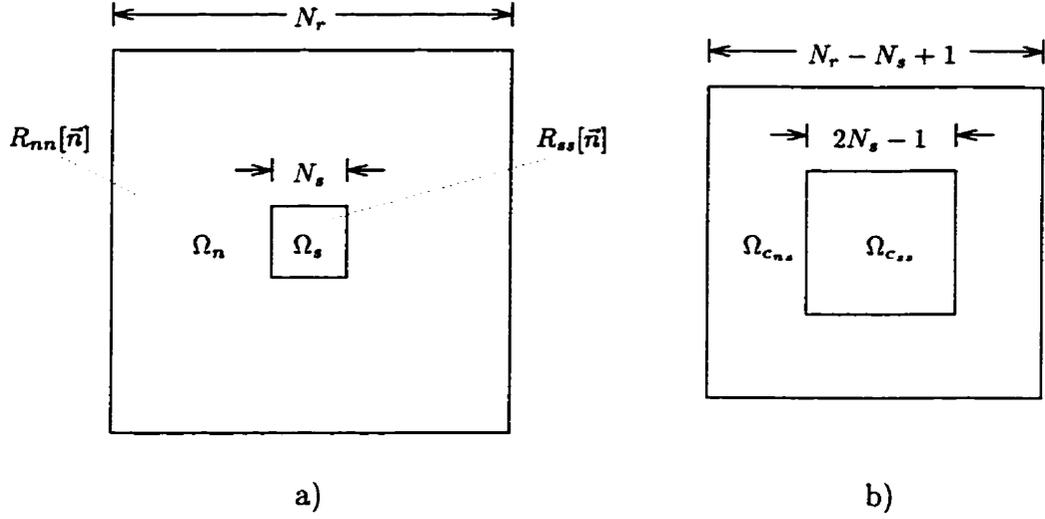


Figure 2.1: (a) Regions in image domain (b) in cross-correlation domain.

where  $w_r[\vec{n}]$  defines the extent of the image:

$$w_r[\vec{n}] = \begin{cases} 1 & \vec{n} \in \Omega_r \\ 0 & \text{otherwise.} \end{cases} \quad (2.7)$$

Again, for simplicity, we assume the image—and thus  $\Omega_r$ —is an  $N_r \times N_r$  square:

$$\Omega_r = \left[-\frac{N_r}{2}, \frac{N_r}{2} - 1\right] \times \left[-\frac{N_r}{2}, \frac{N_r}{2} - 1\right]. \quad (2.8)$$

It is also assumed, of course, that  $\Omega_s \subset \Omega_r$ , i.e.  $N_s < N_r$ . It follows from (2.6) that the domain of  $\bar{\mathbf{n}}[\vec{n}]$  is the “difference” between  $\Omega_r$  and  $\Omega_s$ :

$$\Omega_n = \Omega_s^c \cap \Omega_r, \quad (2.9)$$

where  $^c$  and  $\cap$  denote set complement and set intersection respectively. By combining (2.1), (2.2) and (2.5), the received image may be written as

$$\bar{\mathbf{r}}[\vec{n}] = \bar{\mathbf{s}}[\vec{n}] + \bar{\mathbf{n}}[\vec{n}] \quad (2.10)$$

$$= \mathbf{s}[\vec{n}] w_s[\vec{n}] + \mathbf{n}[\vec{n}] w_n[\vec{n}] \quad (2.11)$$

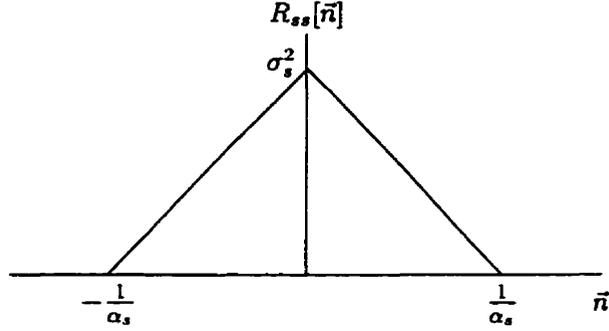


Figure 2.2: Profile of  $R_{ss}[\vec{n}]$ .

The exact form of the signal and noise autocorrelation functions  $R_{ss}[\vec{n}]$  and  $R_{nn}[\vec{n}]$  is somewhat arbitrary, although as noted in chapter 1, image spectra are generally low-pass in nature. To be concrete, however, we choose triangular autocorrelation functions:

$$R_{ss}[\vec{n}] = E \left\{ \mathbf{s}[\vec{k}] \mathbf{s}[\vec{k} + \vec{n}] \right\} = \sigma_s^2 \text{tri}[\alpha_s \vec{n}] \quad (2.12)$$

$$R_{nn}[\vec{n}] = E \left\{ \mathbf{n}[\vec{k}] \mathbf{n}[\vec{k} + \vec{n}] \right\} = \sigma_n^2 \text{tri}[\alpha_n \vec{n}] \quad (2.13)$$

where  $E \{ \cdot \}$  is the expectation operator,  $\sigma_s^2$  and  $\sigma_n^2$  are the signal and noise variances, and  $\alpha_s$  and  $\alpha_n$  are the signal and noise bandwidths<sup>4</sup> in units of inverse samples<sup>5</sup> (Figure 2.2).  $\text{tri}[\alpha \vec{n}]$  is a 2-D triangle function defined as

$$\text{tri}[\alpha \vec{n}] = \begin{cases} (1 - |\alpha n_x|)(1 - |\alpha n_y|) & |n_x| \leq \frac{1}{\alpha}, |n_y| \leq \frac{1}{\alpha} \\ 0 & \text{otherwise.} \end{cases} \quad (2.14)$$

As shown in Section A.1, stochastic signals with triangular autocorrelation functions are easily generated by convolving white noise with a boxcar filter (i.e. a filter with a rectangular impulse response).

A couple of remarks concerning the model:

---

<sup>4</sup>A physical interpretation of bandwidth is given in appendix A.1

<sup>5</sup>Thus,  $\alpha_s, \alpha_n < 1$ .

- In practice, there exist *two* sources of noise: scene noise—which we have modelled—and random *background* noise, associated with the physical environment and imaging system (e.g. effect of lens aberrations during camera movements). We have elected to exclude the effect of background noise, since its inclusion would do little to augment the fundamental results of our study. However, one should be mindful that in the context of real images, its ubiquitous presence can be a significant factor in driving up the probability of false detection in certain situations (Section 3.3).
- Our assumption that the signal and noise are Gaussian may seem somewhat idealized, but not entirely unjustified. According to [20], “Intensity distributions of windows from natural images can usually be approximated by normal distributions”, which when taken in conjunction with the ergodic hypothesis, provides some rationale for our assumption.

The model described here is not intended to be a comprehensive representation of *all* real images. It does however, incorporate certain important features of real images, including their finite extent, low-pass spectral characteristics, and the non-overlapping nature of scene noise. Because it is sufficiently simple, it is to some extent, amenable to closed-form analysis, which is useful in uncovering the fundamental processes that affect the false detection rate.

## 2.2 Cross-Correlation

The cross-correlation between the received image and the template is given by

$$c[\vec{n}] = \frac{1}{a_s} \sum_{\vec{k} \in \Omega_k} \bar{r}[\vec{k}] \bar{s}[\vec{k} - \vec{n}], \quad \vec{n} \in \Omega_c, \quad (2.15)$$

where  $\vec{k}$  spans the domain of the template (region  $\Omega_k$ ),  $\vec{n}$  is such that the template is always within the boundaries of the image (region  $\Omega_c$ ), and the normalizing factor,  $a_s$ , is the average energy of the target given by<sup>6</sup>

$$a_s = N_s^2 \sigma_s^2. \quad (2.16)$$

Explicitly,  $\Omega_k$  and  $\Omega_c$  are defined as

$$\Omega_k = \left[-\frac{N_s}{2} + n_x, \frac{N_s}{2} - 1 + n_x\right] \times \left[-\frac{N_s}{2} + n_y, \frac{N_s}{2} - 1 + n_y\right], \quad (2.17)$$

$$\Omega_c = \left[-\frac{N_r - N_s}{2}, \frac{N_r - N_s}{2}\right] \times \left[-\frac{N_r - N_s}{2}, \frac{N_r - N_s}{2}\right], \quad (2.18)$$

and are  $N_s \times N_s$  and  $(N_r - N_s + 1) \times (N_r - N_s + 1)$  in size, respectively. Substitution of (2.10) into (2.15) yields<sup>7</sup>

$$c[\vec{n}] = a_s^{-1} \sum_{\vec{k}} \bar{s}[\vec{k}] \bar{s}[\vec{k} - \vec{n}] + a_s^{-1} \sum_{\vec{k}} \bar{n}[\vec{k}] \bar{s}[\vec{k} - \vec{n}] \quad (2.19)$$

That is, the total cross-correlation can be expressed as the sum of the cross-correlations between the template and the individual components of the received image. It proves convenient to write (2.19) as

$$c[\vec{n}] = c_{\bar{s}\bar{s}}[\vec{n}] + c_{\bar{n}\bar{s}}[\vec{n}], \quad (2.20)$$

where

$$c_{\bar{s}\bar{s}}[\vec{n}] = a_s^{-1} \sum_{\vec{k}} \bar{s}[\vec{k}] \bar{s}[\vec{k} - \vec{n}] \quad (2.21)$$

$$c_{\bar{n}\bar{s}}[\vec{n}] = a_s^{-1} \sum_{\vec{k}} \bar{n}[\vec{k}] \bar{s}[\vec{k} - \vec{n}]. \quad (2.22)$$

---

<sup>6</sup> $\sigma_s^2$  is the average signal power—the energy per unit area. Thus, multiplying by  $N_s^2$ —the area of the target—yields the target energy.

<sup>7</sup>In the interest of notational brevity, it is hereafter assumed that  $\vec{k} \in \Omega_s$  and  $\vec{n} \in \Omega_c$ , unless otherwise specified.

To develop a better sense of the physical meaning of these equations, it is useful to examine (2.21) and (2.22) in a more heuristic manner.

Concerning (2.21), we observe that  $\mathbf{c}_{\vec{s}\vec{s}}[\vec{n}]$  is nothing more than the autocorrelation of the template itself. It is non-zero in the  $(2N_s - 1) \times (2N_s - 1)$  region defined by

$$\Omega_{\mathbf{c}_{\vec{s}\vec{s}}} = [-N_s + 1, N_s - 1] \times [-N_s + 1, N_s - 1], \quad (2.23)$$

(Figure 2.1(b)) and is a maximum at  $\vec{n} = \vec{0}$ , the shift at which the template lies directly over the target. The manner in which the height of the maximum is distributed may be seen by setting  $\vec{n} = \vec{0}$  in (2.21):

$$\mathbf{c}_{\vec{s}\vec{s}}[\vec{0}] = a_s^{-1} \sum_{\vec{k}} \tilde{s}^2[\vec{k}] \quad (2.24)$$

—it can be approximated by a chi-squared probability density function. It is not *exactly*  $\chi^2$ , since neighbouring samples in  $\tilde{s}[\vec{n}]$  are not independent. Note however, that if the number of terms in (2.21) is increased—thereby increasing the number of independent terms— $\mathbf{c}_{\vec{s}\vec{s}}[\vec{0}]$  may be approximated as a Gaussian RV, via the Central Limit Theorem<sup>8</sup> (CLT). This may be achieved by increasing the size and/or the bandwidth of the target. Thus, for high space-bandwidth product (SBWP) targets, the height of the cross-correlation peak is essentially normally distributed. As shown in Section A.2, the mean of the distribution, i.e. the average peak height, is unity:

$$\mu_{c_0} = 1. \quad (2.25)$$

The variance in peak height when  $N_s > 1/\alpha_s$  is given by:

$$\sigma_{c_0}^2 = 2 \left[ \frac{2}{3N_s\alpha_s} - \frac{1}{6(N_s\alpha_s)^2} \right]^2, \quad (2.26)$$

---

<sup>8</sup>Central Limit Theorem: as the number of arbitrarily distributed *independent* RVs increases, their sum approaches a Gaussian RV.

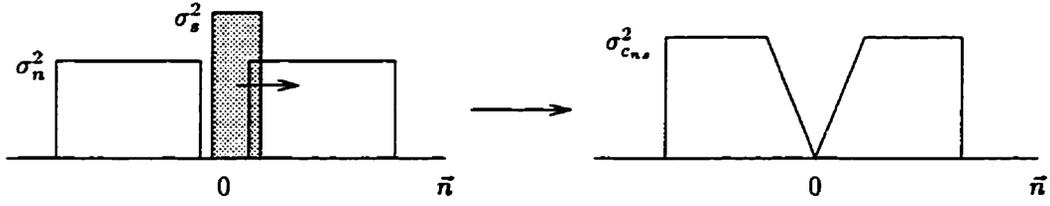


Figure 2.3: Spatial dependence of  $\sigma_{c_{ns}}^2$ .

as shown in A.3. Note that as  $(N_s \alpha_s)^2$ —the signal space-bandwidth product—increases,  $\sigma_{c_0}^2$  decreases monotonically.

Equation 2.22 is the signal-noise cross-correlation, which like  $c[\vec{n}]$ , is defined on region  $\Omega_c$ . Like  $c_{\vec{s}\vec{s}}[\vec{n}]$ , the PDF of  $c_{\vec{n}\vec{s}}[\vec{n}]$  at the limit of large signal or noise SBWP may be reasonably approximated as Gaussian. Because the signal and noise are uncorrelated, the mean of  $c_{\vec{n}\vec{s}}[\vec{n}]$  is zero for all  $\vec{n} \in \Omega_c$ :

$$\mu_{c_{ns}} = 0. \quad (2.27)$$

The variance, however, is space-dependent. A sense of the functional relationship between variance and position may be gained by considering a template that is initially located at  $\vec{n} = \vec{0}$ . Because it exactly overlaps the zero-region of  $\vec{n}[\vec{n}]$  (i.e. region  $\Omega_s$ ), the variance is clearly zero. As the template slides to one side (Figure 2.3), the portion that overlaps the non-zero region of  $\vec{n}[\vec{n}]$  (i.e. region  $\Omega_n$ ) increases linearly, resulting in a corresponding linear increase in variance. Finally, when the template is completely within  $\Omega_n$ , the variance becomes constant. This constant variance region is defined by

$$\Omega_{c_{ns}} = \Omega_{c_{ss}}^c \cap \Omega_c \quad (2.28)$$

(Figure 2.1(b)). As shown in Section A.3, the variance of  $c_{\vec{n}\vec{s}}[\vec{n}]$  in this region

when  $N_s > 1/\alpha_s, 1/\alpha_n$  is given by:

$$\sigma_{c_{n,s}}^2 = \begin{cases} \frac{\sigma_n^2}{\sigma_s^2} \left[ \frac{1}{N_s \alpha_s} \left( 1 - \frac{\alpha_n}{3\alpha_s} \right) - \frac{1}{6(N_s \alpha_s)^2} \left( 2 - \frac{\alpha_n}{\alpha_s} \right) \right]^2, & \alpha_s > \alpha_n \\ \frac{\sigma_n^2}{\sigma_s^2} \left[ \frac{1}{N_s \alpha_n} \left( 1 - \frac{\alpha_s}{3\alpha_n} \right) - \frac{1}{6(N_s \alpha_n)^2} \left( 2 - \frac{\alpha_s}{\alpha_n} \right) \right]^2, & \alpha_n > \alpha_s, \end{cases} \quad (2.29)$$

We see that  $\sigma_{c_{n,s}}^2$  is inversely proportional to  $\sigma_s^2/\sigma_n^2$ , the input SNR ( $\text{SNR}_i$ ); the greater the SNR, the smaller the noise peak variance. Furthermore,  $\sigma_{c_{n,s}}^2$  decreases monotonically with  $(N_s \alpha_s)^2$  (or  $(N_s \alpha_n)^2$ ) as well as with  $(\alpha_n/\alpha_s)^2$  (or  $(\alpha_s/\alpha_n)^2$ ). Note that when  $\sigma_s^2 = \sigma_n^2$  and  $\alpha_s = \alpha_n$ , (2.29) collapses to half of (2.26).

## 2.3 Probability of False Detection

We may measure the detection performance of the cross-correlator by determining the frequency with which a false peak is mistaken for the true peak—that is, the probability of false detection (POFD) or the false alarm rate. More precisely, we wish to find the probability that

$\mathcal{A}$  : one or more points in  $\mathbf{c}[\vec{n}]$ ,  $\vec{n} \in \Omega_c, \vec{n} \neq \vec{0}$  exceeds the true signal peak,  $\mathbf{c}_0$  (i.e.  $\mathbf{c}[\vec{0}]$ ).

We may express this formally as

$$P_{FD} = P\{\mathcal{A}\}, \quad (2.30)$$

where  $P_{FD}$  is the probability of false detection and

$$\mathcal{A} = \bigcup_{\vec{n}} \{\mathbf{c}[\vec{n}] > \mathbf{c}_0\}, \quad (2.31)$$

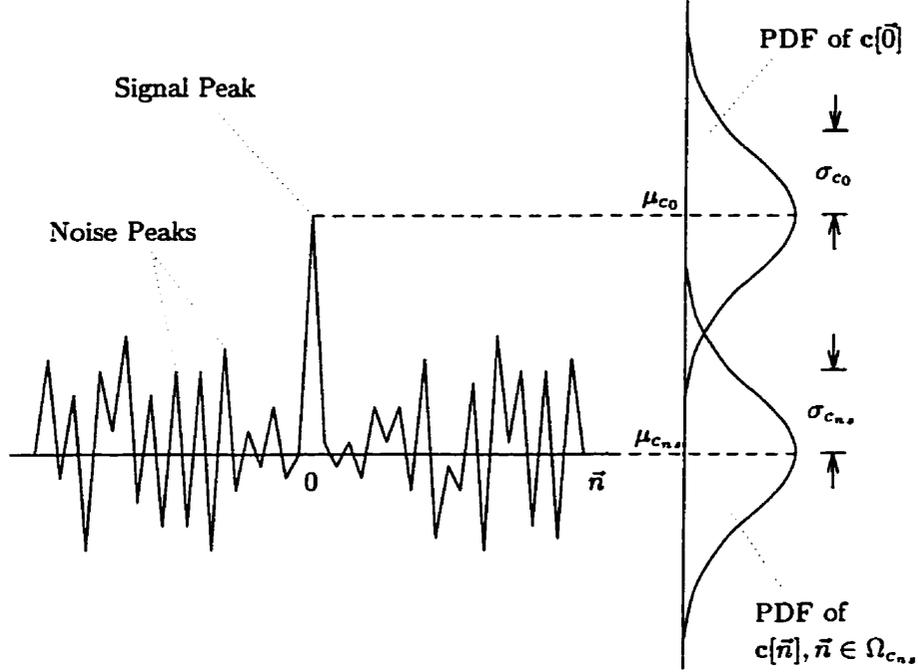


Figure 2.4: Cross-correlation and its associated PDFs.

where  $\cup$  denotes set union and  $\bar{n} \in \Omega_c, \bar{n} \neq \bar{0}$ . It is useful to think of  $P\{\mathcal{A}\}$  in terms of its frequency interpretation [50]: Given a large number of realizations of  $c[\bar{n}]$ , each realization will contain a true, signal peak and a number of noise peaks (Figure 2.4).  $P\{\mathcal{A}\}$  may be thought of as the fraction of realizations that contain one or more noise peaks which exceed the signal peak<sup>9</sup>. Three variables affect  $P\{\mathcal{A}\}$ :

**Signal peak variance.** As  $\sigma_{c_0}^2$  increases, the overlap between the signal and noise peak PDFs also increases (Figure 2.4), thereby increasing the probability of false detection. As shown in (2.26),  $\sigma_{c_0}^2$  is a function of the signal SBWP.

<sup>9</sup>The true peak may be thought of as a variable threshold which changes with each realization.

**Noise peak variance.** Likewise, as  $\sigma_{c_{n,s}}^2$  increases, the probability of false detection also increases. We see in (2.29), that  $\sigma_{c_{n,s}}^2$  is a function of the input SNR, the target size, and the target and noise bandwidth. We note especially that the target bandwidth has a larger effect on the noise peak variance than the noise bandwidth when the target bandwidth is larger, and vice versa.

**Number of independent noise peaks.** As the number of independent noise peaks increases, the probability that *at least one* peak will exceed the signal peak also increases; it approaches unity as the number of peaks approaches infinity. The number of such peaks is partially dependent on the noise SBWP: As  $\Omega_n$  increases in size, so does  $\Omega_{c_{n,s}}$ , thereby increasing in the total number of peaks. Equivalently, as the bandwidth increases, the width of the peaks decreases, resulting in an increase in the number of independent peaks in a given area. The number of noise peaks is also influenced by the signal size and bandwidth. Because the noise does not overlap the target, a large target will reduce the area of the noise, thus reducing the number of peaks. A low bandwidth target will also reduce the number of independent peaks, for it will tend to smooth out higher bandwidth noise, eliminating peaks as it slides over it like a low-pass filter. Conversely, a high bandwidth target will tend to leave noise of lower bandwidth unaffected.

It is important to note that these three variables are not independent of each other. Changing the target size, target bandwidth or noise bandwidth affects *both* the noise peak variance *and* the number of noise peaks simultaneously. A change in the image size however, will affect only the number of noise peaks.

In quantifying  $P\{\mathcal{A}\}$ , we simplify the task by first rewriting (2.30) in

terms of its complement:

$$P\{\mathcal{A}\} = 1 - P\{\bigcap_{\vec{n}} \{c[\vec{n}] < c_0\}\}. \quad (2.32)$$

The computation of the joint probability  $P\{\bigcap_{\vec{n}} \{c[\vec{n}] < c_0\}\}$  is complicated by the fact that the samples in  $c[\vec{n}]$ , are *not* independent. Since cross-correlation is essentially a filtering operation and both the received image and target are coloured, neighbouring samples in  $c[\vec{n}]$  are not independent. The consequence of this is that the joint probability density function (PDF) is not simply the product of marginal densities, and thus

$$P\{\bigcap_{\vec{n}} \{c[\vec{n}] < c_0\}\} \neq \prod_{\vec{n}} P\{c[\vec{n}] < c_0\}. \quad (2.33)$$

The most direct and exact approach for computing the joint probability is to actually form the joint PDF then integrate it over the appropriate range. However, a little foresight reveals that this is a computationally daunting task. Note that the set in (2.32) is over an  $((N_r - N_s + 1)^2 - 1)$  point region. Thus, an  $((N_r - N_s + 1)^2 - 1)$ -variate PDF would be required. Because of (2.33), the resulting integral would also not be separable. Even for images of modest size, the computational time required for such an integral to converge would be enormous. Thus, the means to an exact determination of the joint probability is effectively made unavailable. Another approach is required.

We were not able to use (2.33) because in general, not *all* samples in  $c[\vec{n}]$  are independent. However, if we can estimate the number of samples in  $c[\vec{n}]$  which *are* independent—that is, the number of independent noise peaks—we then *can* use the following independence-based approximation to estimate the joint probability:

$$P\{\bigcap_{\vec{n}} \{c[\vec{n}] < c_0\}\} \approx \prod_{m=1}^M P\{c_{I_m} < c_0\}, \quad (2.34)$$

where  $\mathbf{c}_{I1}, \mathbf{c}_{I2}, \dots, \mathbf{c}_{IM}$  are the  $M$  independent samples. That is, the *effect* of a large number of partially correlated samples may be simulated by replacing them with a smaller number of independent samples.

In general, the number of independent samples in a given region depends on the statistics of that region. In a stationary, high bandwidth region, the number of independent samples is large and the spacing between independent samples is small and uniform. Conversely, in a non-stationary, low bandwidth region, the opposite is true. Strictly speaking, determining  $M$  for  $\mathbf{c}[\vec{n}]$  would involve computing the number of independent samples in region  $\Omega_{c_{ns}}$  (stationary) and region  $\Omega_{c_{ss}}$  (non-stationary) and finding their sum. However, it proves sufficient to evaluate  $M$  over region  $\Omega_{c_{ns}}$  only, provided  $\Omega_{c_{ss}}$  is sufficiently small.

An estimate of  $M$  on  $\Omega_{c_{ns}}$  may be obtained by first determining the *correlation area*,  $N_l^2$ . The correlation area may be thought of as the 2-D analog to the 1-D *correlation length*,  $N_l$ —the number of samples between independent samples<sup>10</sup>. In general, the correlation length cannot be precisely defined, since the correlation between samples typically does not fall to zero exactly, regardless of how large the separation. Fortunately however, our choice of triangular autocorrelation functions for both the signal and noise (Equations (2.12) and (2.13)) affords us an exact definition for the correlation length (see Section A.3):

$$N_l = \frac{1}{\alpha_s} + \frac{1}{\alpha_n}. \quad (2.35)$$

That is, the correlation length of  $\mathbf{c}[\vec{n}]$ ,  $\vec{n} \in \Omega_{c_{ns}}$ , is simply the sum of the correlation lengths of the signal and noise. The number of independent samples in  $\Omega_{c_{ns}}$  is then simply the ratio of the area of  $\Omega_{c_{ns}}$  to the correlation area, i.e.

---

<sup>10</sup>For example, if two adjacent samples are independent of each other,  $N_l = 1$ .

the “area per independent sample”:

$$M = \frac{(N_r - N_s + 1)^2 - N_s^2}{\left(\frac{1}{\alpha_s} + \frac{1}{\alpha_n}\right)^2}. \quad (2.36)$$

An inspection of (2.36) confirms our earlier intuition that the number of independent noise peaks,  $M$ , is driven primarily by the signal if its bandwidth is lower than that of the noise and vice versa. For example, if  $\alpha_s$  is very large,  $1/\alpha_s \approx 0$  and the  $1/\alpha_n$  term dominates.

Having determined  $M$ , we return our attention to equation (2.34). Because we are only concerned about region  $\Omega_{c_{ns}}$ , which is stationary, the independent samples  $\mathbf{c}_{I_m}$  are all identically distributed. It follows then, that (2.34) becomes

$$P\{\cap_{\bar{n}} \{\mathbf{c}[\bar{n}] < \mathbf{c}_0\}\} \approx P\{\mathbf{c}_I < \mathbf{c}_0\}^M \quad (2.37)$$

where we have dropped the subscript  $m$  without ambiguity. If the CLT holds as described in 2.2, then  $\mathbf{c}_I$  is distributed according to  $N(\mu_{c_{ns}}; \sigma_{c_{ns}}^2)$ <sup>11</sup>. Similarly,  $\mathbf{c}_0$  is  $N(\mu_{c_0}; \sigma_{c_0}^2)$ .

In computing  $P\{\mathbf{c}_I < \mathbf{c}_0\}$ , we first note that

$$P\{\mathbf{c}_I < \mathbf{c}_0\} = P\{\mathbf{c}_I - \mathbf{c}_0 < 0\}. \quad (2.38)$$

Because  $\mathbf{c}_I$  and  $\mathbf{c}_0$  are both Gaussian,  $\mathbf{c}_I - \mathbf{c}_0$  is also Gaussian with mean  $\mu_{c_{ns}} - \mu_{c_0}$  and variance  $\sigma_{c_{ns}}^2 + \sigma_{c_0}^2$ . Thus,

$$\begin{aligned} P\{\mathbf{c}_I < \mathbf{c}_0\} &= P\{\mathbf{c}_I - \mathbf{c}_0 < 0\} \\ &= \frac{1}{\sqrt{2\pi(\sigma_{c_0}^2 + \sigma_{c_{ns}}^2)}} \int_{-\infty}^0 \exp\left[-\frac{(y - (\mu_{c_{ns}} - \mu_{c_0}))^2}{2(\sigma_{c_{ns}}^2 + \sigma_{c_0}^2)}\right] dy. \end{aligned} \quad (2.39)$$

---

<sup>11</sup> $N(\eta; \sigma^2)$  denotes a normal RV with mean  $\eta$  and variance  $\sigma^2$ .

Before proceeding to find  $P_{FD}$ , we must first assert an equality in (2.37). We may do this by introducing a constant,  $\gamma$ , into the exponent:

$$P\{\cap_{\vec{n}} \{c[\vec{n}] < c_0\}\} = P\{c_I < c_0\}^{\gamma^2 M}. \quad (2.40)$$

The constant is squared to emphasize that the expression represents a two-dimensional process. Physically,  $\gamma$  adjusts for the presence of the partially correlated samples between the independent  $c_I$ 's, whose probability contribution we have not yet accounted for. One can see that  $P\{c_I < c_0\}^{\gamma^2 M} < P\{c_I < c_0\}^M$ , since, in effect, we are "requiring" that not only the independent samples,  $c_I$ , be less than  $c_0$ , but the partially correlated samples as well—a less probable event. It follows then that  $\gamma > 1$ . The exact value of  $\gamma$  is in general, difficult to determine, although it seems likely that it is dependent on  $R_{c_{ns}}[\vec{n}]$ . There is one exception, however: if there are no correlated samples between independent samples (i.e.  $c[\vec{n}]$  is white), then  $\gamma = 1$ .

Finally, combining (2.30), (2.32), (2.39) and (2.40) yields the probability of false detection:

$$P_{FD} = 1 - P\{c_I < c_0\}^{\gamma^2 M} \quad (2.41)$$

$$= 1 - \left[ \frac{1}{\sqrt{2\pi(\sigma_{c_0}^2 + \sigma_{c_{ns}}^2)}} \int_{-\infty}^0 \exp \left[ -\frac{(y - (\mu_{c_{ns}} - \mu_{c_0}))^2}{2(\sigma_{c_{ns}}^2 + \sigma_{c_0}^2)} \right] dy \right]^{\gamma^2 M}. \quad (2.42)$$

Since  $\mu_{c_{ns}} = 0$  and  $\mu_{c_0} = 1$ , the Gaussian is centered at  $y = -1$ . Translating it to  $y = 0$  and adjusting the upper limit of integration, we obtain

$$P_{FD} = 1 - \left[ \frac{1}{\sqrt{2\pi(\sigma_{c_0}^2 + \sigma_{c_{ns}}^2)}} \int_{-\infty}^1 \exp \left[ -\frac{y^2}{2(\sigma_{c_{ns}}^2 + \sigma_{c_0}^2)} \right] dy \right]^{\gamma^2 M}. \quad (2.43)$$

It is important to recognize the assumptions that underscore equation (2.43):

- The signal and noise are uncorrelated.

- The signal and noise are zero mean.
- The signal and noise variance is constant within their respective domains.
- The height distributions of the signal and noise correlation peaks are Gaussian. This is only true when the signal and noise space-bandwidth products are sufficiently large ( $\gtrsim 250$ ).
- The size of the target is small in comparison to that of the image, and thus, the contribution to the probability of false detection from the non-stationary cross-correlation region (i.e. region  $\Omega_{c_{ss}}$ ) is negligible. In effect then,  $P_{FD}$  is the probability that noise peaks in the *stationary* region only (i.e. region  $\Omega_{c_{ns}}$ ) exceed the signal peak.

## 2.4 Summary

The three variables that influence the probability of false detection,  $P_{FD}$ , are 1) the signal peak variance,  $\sigma_{c_0}^2$ , 2) the noise peak variance,  $\sigma_{c_{ns}}^2$  and 3) the number of independent noise peaks,  $M$ . An increase in any one of these variables results in a corresponding increase in  $P_{FD}$ . The equations that govern these variables are (2.26), (2.29) and (2.36) respectively, which we restate here

	$\sigma_{c_0}^2$	$\sigma_{c_{ns}}^2$	$M$
$N_s \uparrow$	$\downarrow$	$\downarrow$	$\downarrow$
$N_r \uparrow$	-	-	$\uparrow$
$\alpha_s \uparrow$	$\downarrow$	$\downarrow$	$\uparrow$
$\alpha_n \uparrow$	-	$\downarrow$	$\uparrow$
$\text{SNR}_i \uparrow$	-	$\downarrow$	-

Table 2.1: Effects of model parameters.  $\uparrow$  = increase,  $\downarrow$  = decrease, - = none.

for convenience:

$$\sigma_{c_0}^2 = 2 \left[ \frac{2}{3N_s\alpha_s} - \frac{1}{6(N_s\alpha_s)^2} \right]^2 \quad (2.44)$$

$$\sigma_{c_{ns}}^2 = \begin{cases} \frac{\sigma_n^2}{\sigma_s^2} \left[ \frac{1}{N_s\alpha_s} \left( 1 - \frac{\alpha_n}{3\alpha_s} \right) - \frac{1}{6(N_s\alpha_s)^2} \left( 2 - \frac{\alpha_n}{\alpha_s} \right) \right]^2, & \alpha_s > \alpha_n \\ \frac{\sigma_n^2}{\sigma_s^2} \left[ \frac{1}{N_s\alpha_n} \left( 1 - \frac{\alpha_s}{3\alpha_n} \right) - \frac{1}{6(N_s\alpha_n)^2} \left( 2 - \frac{\alpha_s}{\alpha_n} \right) \right]^2, & \alpha_n > \alpha_s, \end{cases} \quad (2.45)$$

$$M = \frac{(N_r - N_s + 1)^2 - N_s^2}{\left( \frac{1}{\alpha_s} + \frac{1}{\alpha_n} \right)^2}. \quad (2.46)$$

The effects of the model parameters on  $\sigma_{c_0}^2$ ,  $\sigma_{c_{ns}}^2$  and  $M$  are summarized in Table 2.1.

# Chapter 3

## Statistical Simulation

In this chapter, we present experimental verification of the theoretical results derived in chapter 2. Because the theoretical model is inherently statistical in nature, so were the computer simulations which were employed to verify its validity. We begin by detailing the technical aspects of the simulation. That is followed by a presentation and discussion of one-dimensional results. The chapter concludes with an examination of the limitations of the theoretical model.

### 3.1 Technical Description

Because our theoretical model is based on 2-D signals, ideally, our simulations should have also been based on 2-D signals. However, due to the large numbers of realizations required for statistically significant results, the computational time required would have been prohibitive. For this reason, we chose to restrict our simulations to 1-D signals. Thus, although we will continue to use the 2-D mathematical notation, the reader should be wary that the results presented

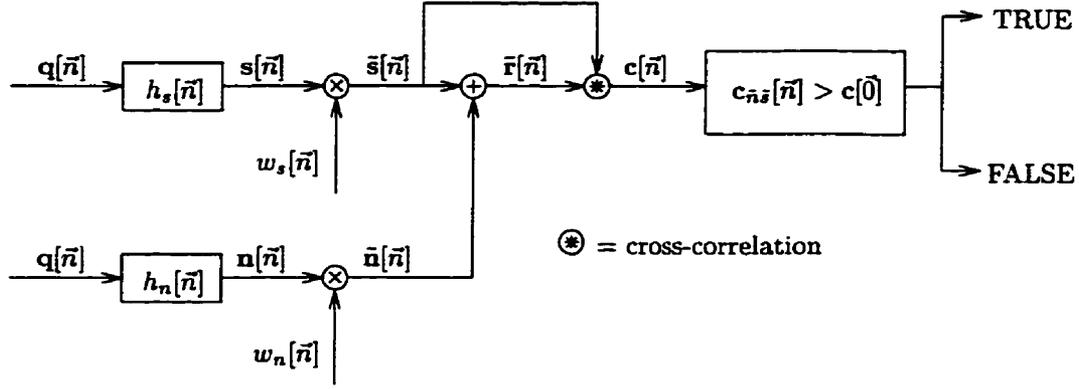


Figure 3.1: System diagram.

in Section 3.2 arise from 1-D signals. For convenience, 1-D versions of key equations have been included in Section A.5.

The various stages of the simulation are summarized in Figure 3.1. Sample functions of the signal and noise were generated by convolving different (i.e. uncorrelated) white noise realizations,  $q[\bar{n}]$ , with boxcar filters,  $h_s[\bar{n}]$  and  $h_n[\bar{n}]$ , respectively (Equations (A.5) and (A.6)), whose bandwidths ranged from 1-25 Hz<sup>1</sup>. To reduce aliasing, a sampling frequency of  $f_s = 100$  Hz was used. Thus, all signals were at least 2x oversampled (Nyquist frequency = 50 Hz). The relatively short sampling period (i.e.  $\Delta x = 1/f_s = 0.01$  s) was also advantageous in enhancing the validity of the integral approximations in Equations A.22 and A.30. The bandwidths in the 1-25 Hz range were chosen such that there were no quantization errors. That is, filters were an integer number of points in length<sup>2</sup>. Filters that were used and that satisfied this requirement had lengths  $N_h$ , of 100, 20, 10, 5 and 4 points, which corresponded to bandwidths of 1, 5, 10, 20 and 25 Hz, respec-

<sup>1</sup>The actual bandwidth in Hertz, BW, may be related to the bandwidth in inverse samples,  $\alpha$ , by:  $BW = \alpha / \Delta x$ , where  $\Delta x$  is the sampling period.

<sup>2</sup>Example: An 20 Hz filter is  $1/(20 \times 0.01) = 5$  samples long. Conversely, a 15 Hz filter is  $1/(15 \times 0.01) = 6.67$  samples long

tively. Effects associated with incomplete convolution at the edges of both the signal and noise were eliminated by requiring that the white noise sequences be  $N_s + N_h - 1$  and  $N_r + N_h - 1$  in length respectively, so that the properly filtered sequences would be exactly  $N_s$  and  $N_r$  points long, respectively.

The signals used were chosen to be at least  $N_s = 1000$  points in length. Because the probability,  $P\{c_I < c_0\}$  (Equation 2.39), was on the order of 0.999 or greater, a very accurate Gaussian approximation was required. This was made possible by using a long signal sequence, so that the signal space-bandwidth product was sufficiently high ( $\gtrsim 250$ ) most of the time for the Central Limit Theorem to hold.

In selecting the appropriate image size, we were reminded that our theoretical model was premised on the image size being somewhat larger than the target size, so as to ensure that most of the interfering noise peaks in the cross correlation originated from region  $\Omega_{c_{ns}}$ . The appropriate ratio of  $N_r$  to  $N_s$  is dependent somewhat on the input SNR. Small input SNRs result in more false detections and thus a smaller  $N_r/N_s$  ratio (i.e. smaller  $\Omega_{c_{ns}}$ ) may be sufficient to “wash out” the effects of the noise peaks from region  $\Omega_{c_{ns}}$ . Because the number of realizations generated in each simulation was  $\sim 10^3$ , the lowest false detection rate possible was  $\sim 10^{-3}$ . To ensure that the false alarm rates were larger than  $\sim 10^{-3}$  (i.e. in the order of  $\sim 10^{-1}$ ), we used an input SNR of -14.77 dB (i.e.  $\sigma_s^2/\sigma_n^2 = 1/30$ ). At this SNR, we found that a ratio of  $N_r/N_s \approx 5$  or greater was sufficient to ignore the false detection contribution from region  $\Omega_{c_{ns}}$ .

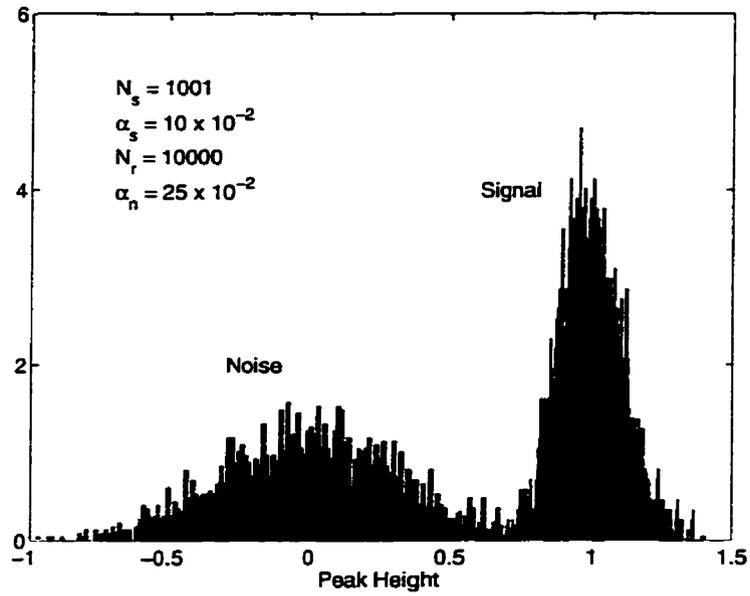
Following the generation of the signal and noise components, the received image was formed and the discrete cross-correlation between the received image and the template was computed. Because only the region  $\vec{n} \in \Omega_c$  is meaningful, only the middle  $N_r - N_s + 1$  samples of the cross-correlation

were retained. The resulting sequence was then scanned for false peaks. Using a different white noise sequence each time, 2000 cross-correlation realizations were generated to obtain an estimate of the probability of false detection.

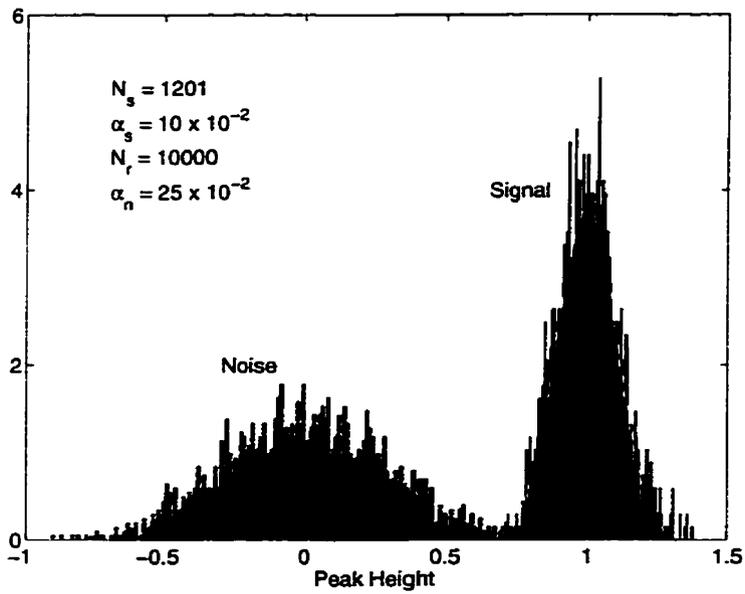
## 3.2 Simulation Results

In this section, we verify the validity of a 1-D version of our model—both in terms of the assumptions made and the theoretical development—by comparing theoretical and experimental results. In doing so, we demonstrate the important functional relationships shown in Table 2.1 which influence the false alarm rate. An understanding of these relationships then provides a qualitative sense of the behaviour of false alarm rates when real images are used.

A key requirement in our model was that the signal and noise peaks be normally distributed. Figure 3.2 shows the signal and noise peak histograms at two different bandwidths. For comparison, the theoretical curves are shown in white. It is not difficult to see that the signal PDF in 3.2(a) ( $N_s\alpha_s = 100$ ) is slightly skewed to the left and thus not quite Gaussian. We recall from section 2.2 that strictly speaking, the signal peak is actually  $\chi^2$  distributed. Because a  $\chi^2$  PDF is one-sided and asymmetric, the CLT summation does not converge as rapidly. As seen in 3.2(b), a higher signal SBWP ( $N_s\alpha_s = 120$ ) results in slightly more symmetric PDF. This however, is still inadequate, since as we see later, at our input SNR of -14.77 dB, a signal SBWP of at least 250 is required for good correspondence between theoretical and experimental false detection curves. At higher input SNRs (i.e.  $\sim 0$  dB) where the false alarm rates are considerably lower (i.e.  $\sim 10^{-12}$ ), SBWPs in excess of 1000 would be required, since we would be operating well into the tail regions of the PDF, where the approximation is poorest.



(a)



(b)

Figure 3.2: Signal and noise PDFs at (a) signal SBWP = 100  
 (b) signal SBWP = 120

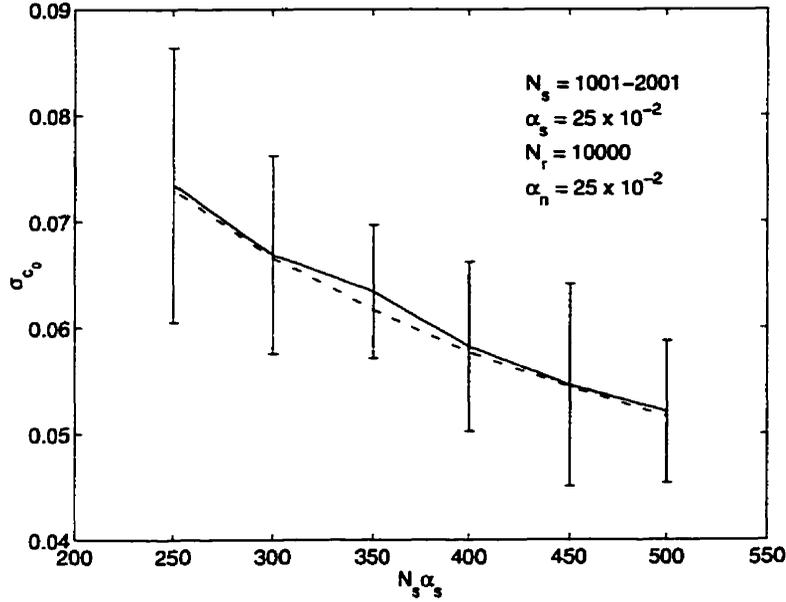


Figure 3.3: Signal peak standard deviation vs. signal SBWP. Solid and broken curves represent the experimental and theoretical data respectively. Error bars represent 95% confidence intervals.

On the surface, it seems curious that the noise PDF in 3.2(a) is not also skewed, since the signal SBWP is the same. The reason for this is twofold. Firstly, because  $\tilde{\mathbf{n}}[\tilde{\mathbf{n}}]$  and  $\tilde{\mathbf{s}}[\tilde{\mathbf{n}}]$  are both Gaussian and *uncorrelated*, the PDF of the product  $\tilde{\mathbf{n}}[\tilde{\mathbf{n}}] \tilde{\mathbf{s}}[\tilde{\mathbf{n}}]$  (equation 2.22) is symmetric about 0, thereby allowing the CLT summation to converge more quickly. The second and more important reason is that the bandwidth of  $\tilde{\mathbf{n}}[\tilde{\mathbf{n}}] \tilde{\mathbf{s}}[\tilde{\mathbf{n}}]$  is determined primarily by the process (signal or noise) with the *higher* bandwidth (see A.4). Since in this case, that process is the noise, with a bandwidth of  $\alpha_n = 25 \times 10^{-2}$ , the SBWP is actually 250. Thus, provided that *either*  $N_s \alpha_s$  or  $N_s \alpha_n$  is large, the PDF of the noise peak may be reasonably approximated as Gaussian. Because of the earlier requirement that  $N_s \alpha_s > 250$ , this condition is always met.

Figure 3.3 shows that the standard deviation of the signal peak is a

decreasing function of signal SBWP, whereas Figure 3.4 shows that the standard deviation of the noise peak decreases with both signal SBWP and noise bandwidth. In all cases, the theoretical results are always within the 95% confidence interval of the simulation results.

The autocorrelation of  $c[\vec{n}]$ ,  $\vec{n} \in \Omega_{c_{ns}}$  is plotted in Figure 3.5. Again, the theoretical and experimental correspondence is good. As expected, the correlation length is dictated primarily by the process (signal or noise) with the lower bandwidth.

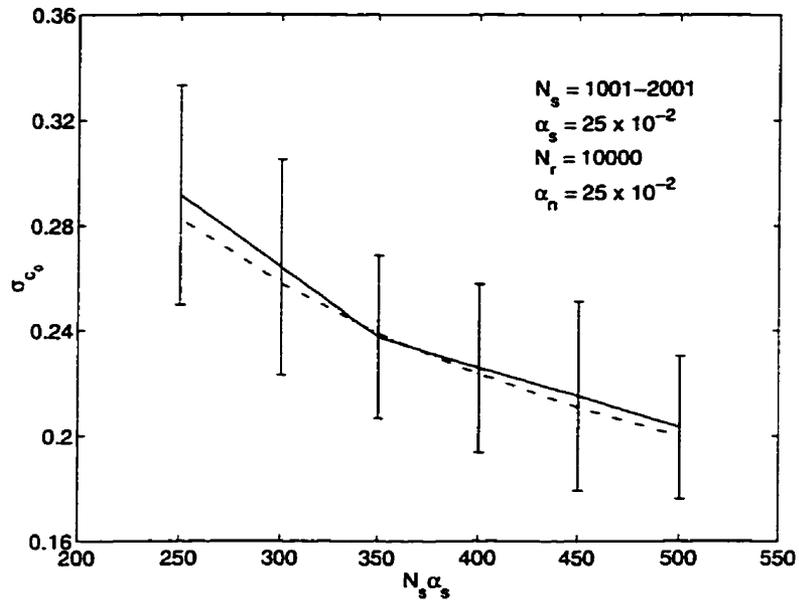
The number of independent noise peaks was not determined experimentally, although Figure 3.6 shows theoretical values of  $M$  as a function of the area of  $\Omega_{c_{ns}}$  and  $\alpha_s$ . Observe in 3.6(b) that when the noise (or signal) bandwidth is constant,  $M$  approaches a limiting value as the signal (or noise) bandwidth grows larger. Note also that the values of  $M$  are quite large.

Finally, the probability of false detection is plotted against each of the model parameters in Figures 3.7, 3.8 and 3.9. The constant  $\gamma$  was determined such that the error between the experimental and theoretical curves in Figure 3.8(a) was minimized<sup>3</sup>; a value of  $\gamma = 3.0$  was chosen.

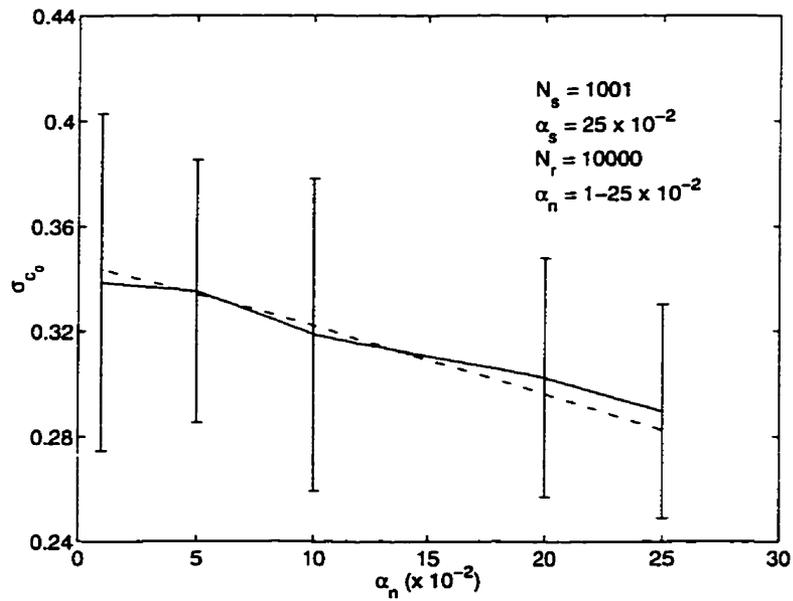
The validity of the theoretical model appears to be supported by Figures 3.7(a), 3.8(a) and 3.8(b). The physical process underlying the behaviour in Figure 3.7(a) is apparent from Table 2.1. We see that an increase in the target size results in a decrease in both the signal and noise peak variances, which in turn leads to a decrease in the false alarm rate. Conversely, in Figure 3.8(a), the variable that affects the behaviour of the graph is the number of independent noise peaks. As the image size grows larger, the noise peaks become more numerous, thereby increasing  $P_{FD}$ . In both of these cases, as

---

<sup>3</sup>The choice of Figure 3.8(a) was entirely arbitrary. Minimizing the error in any of the other plots (with the exception of Figure 3.7(b)) would also have been acceptable.



(a)



(b)

Figure 3.4: (a) Noise peak standard deviation vs. signal SBWP (b) vs. noise bandwidth.

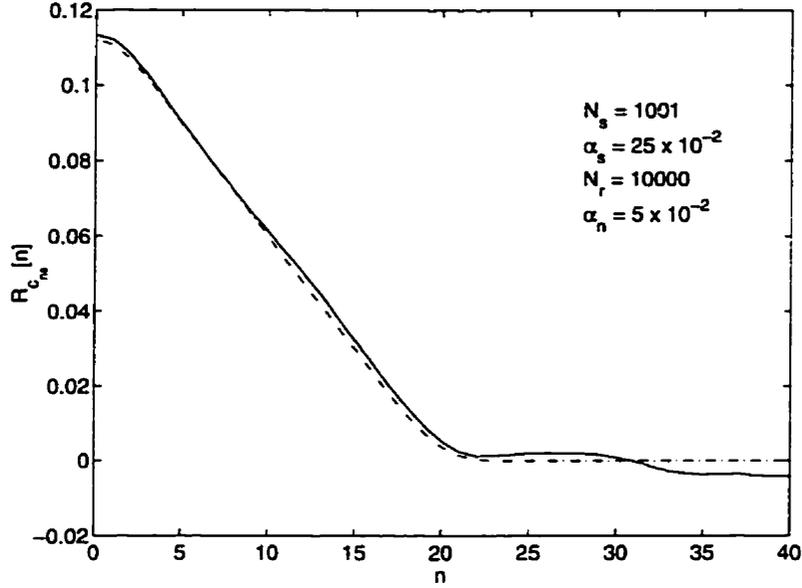
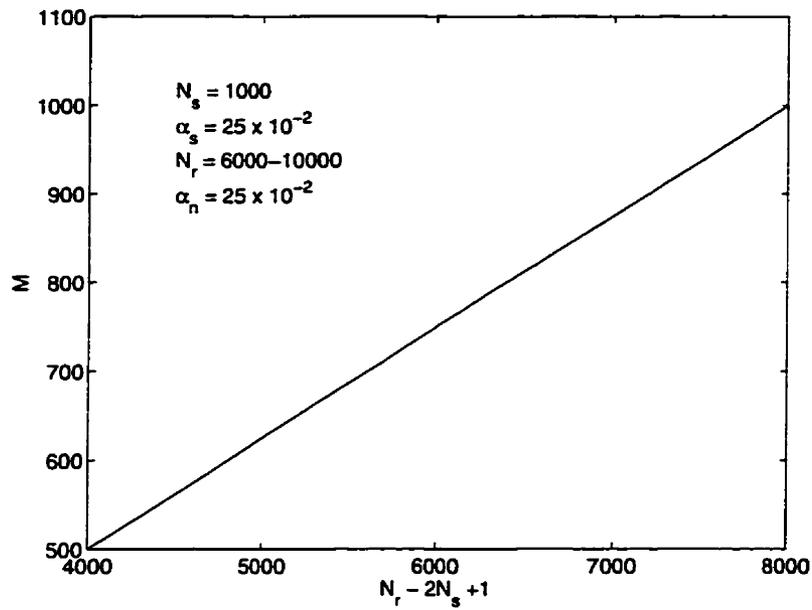


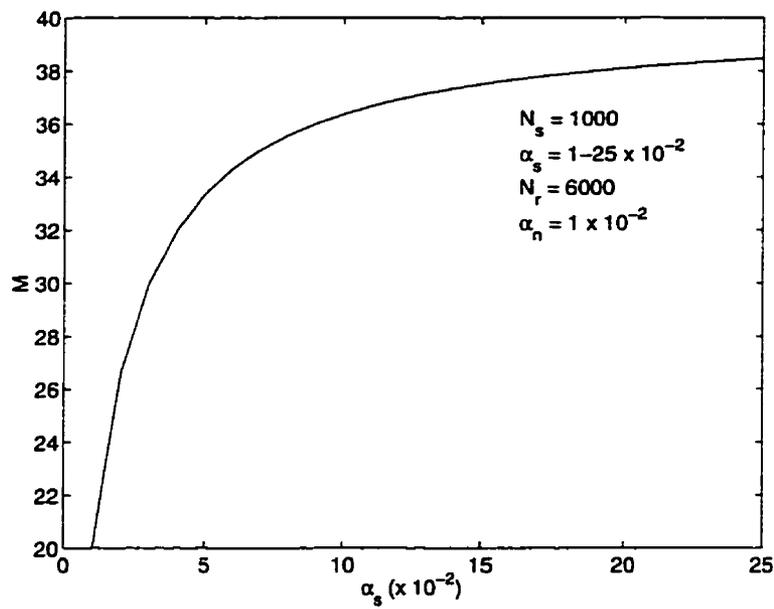
Figure 3.5: Autocorrelation of  $c[\vec{n}]$ ,  $\vec{n} \in \Omega_{c_{n_s}}$  at  $\alpha_n = 5$

the target or image size changes, only one variable changes at a time—*either* the peak variance *or* the number of noise peaks. However, in Figures 3.7(b) (experimental curve) and 3.8(b), *both* change simultaneously and in *opposite directions* with increasing bandwidth; hence the presence of an extremum in both plots. Initially, the false detection rate is driven up by the increasing number of noise peaks. At a certain point—the maximum—the variance becomes sufficiently small that it begins to drive the probability back down. Because of the complex form of Equation 2.42, a closed-form expression for the bandwidth at which the maximum occurs was not found.

The lack of agreement between the theoretical and experimental curves at the first three signal bandwidths in Figure 3.7(b) may be attributed to the invalidity of the Central Limit Theorem in that region. The signal SBWPs at those points are low, ranging from 12 at  $\alpha_n = 1 \times 10^{-2}$  to 120

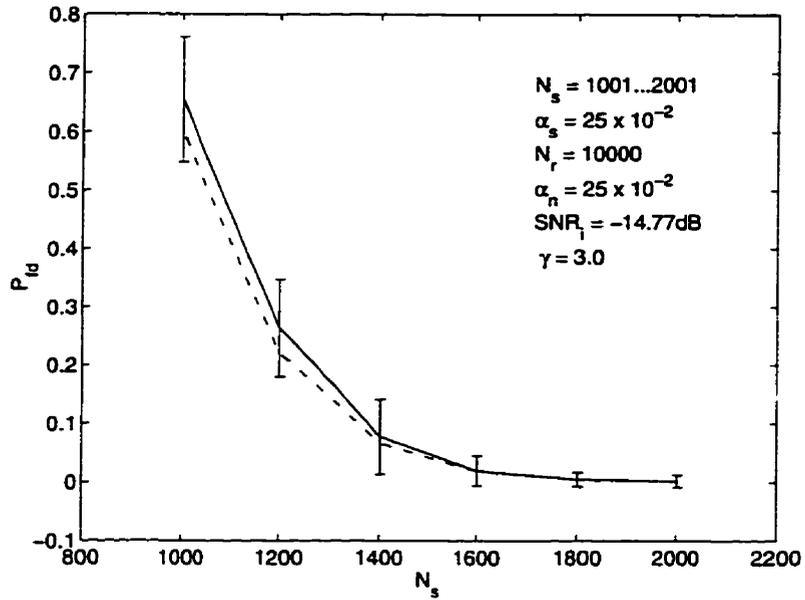


(a)

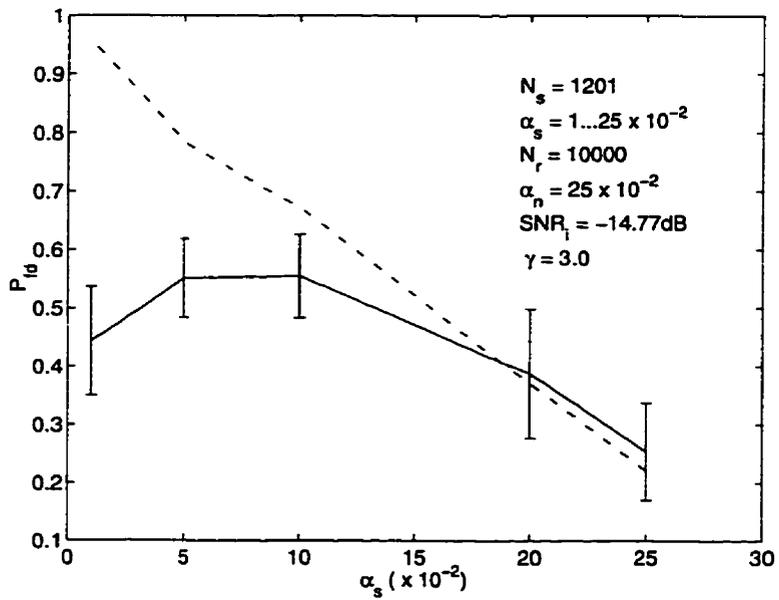


(b)

Figure 3.6: (a) Number of independent noise peaks vs. size of  $\Omega_{c,n,s}$  (b) vs. signal bandwidth.

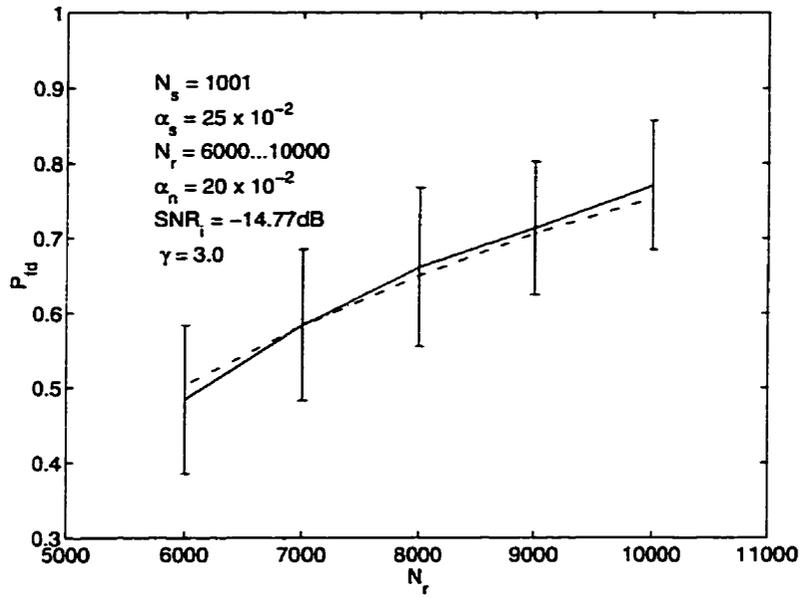


(a)

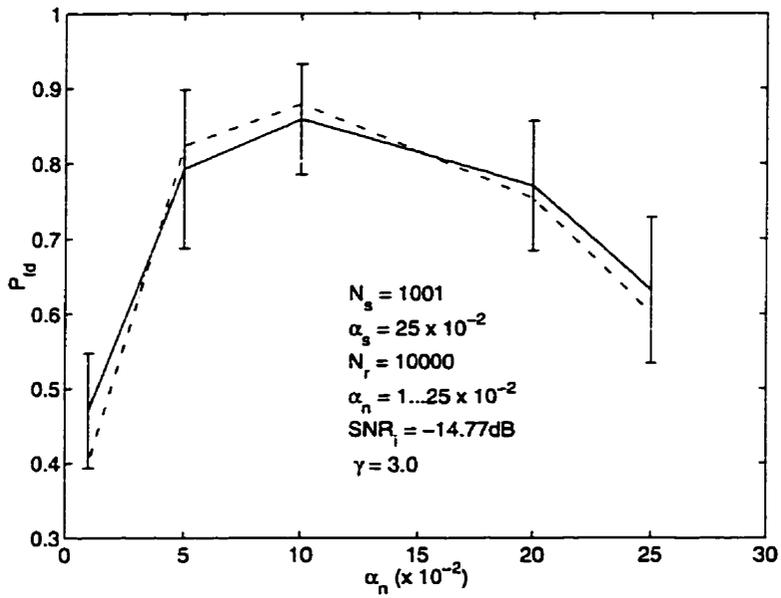


(b)

Figure 3.7: (a) Probability of false detection vs. signal size (b) vs. signal bandwidth.



(a)



(b)

Figure 3.8: (a) Probability of false detection vs. image size (b) vs. noise bandwidth.

at  $\alpha_n = 10 \times 10^{-2}$ . We have already seen in Figure 3.2(a) that at a signal SBWP of 120, the  $\chi^2$  asymmetry in the signal PDF is visible. Because the tail values of a  $\chi^2$  density function for values smaller than the mean is lower than that of a symmetric Gaussian function, there is less overlap between the signal and noise peak PDFs. The result is fewer false detections (experimental curve of Figure 3.7(b)) than if both densities were Gaussian (theoretical curve of Figure 3.7(b).) We note that at a signal SBWP of approximately 250 ( $\alpha_n \approx 20$ )—the requirement that was introduced earlier without justification—the correspondence between the two curves improves substantially.

The rather significant discrepancy in Figure 3.7(b) suggests that the false alarm rate is extremely sensitive to variations in the PDF. This however, is not surprising, since  $P\{c_I < c_0\}$  is raised to the power of  $M$ , (Equation 2.41) which is typically very large. Thus, even relatively small errors in  $P\{c_I < c_0\}$  are greatly amplified.

The theoretical dependence of  $P_{FD}$  on input SNR is shown in Figure 3.9. The relationship was not determined by simulation, since as is apparent from the plot, the false detection rates can become very small and thus the number of realizations required to incur a false detection would be enormous. The trend is expected; as the input SNR increases, the probability of false detection decreases.

### 3.3 Model Limitations

Figure 3.9 suggests that at realistic input SNRs, ( $\sim 0$  dB<sup>4</sup>) false detection rates on the order of  $10^{-12}$ —and even lower in 2-D—are possible. However, one must

---

<sup>4</sup>Because the illumination from region to region in many natural images is on the same order of magnitude, the power of the target and scene noise are comparable, and thus the typical input SNR is  $\sim 0$  dB

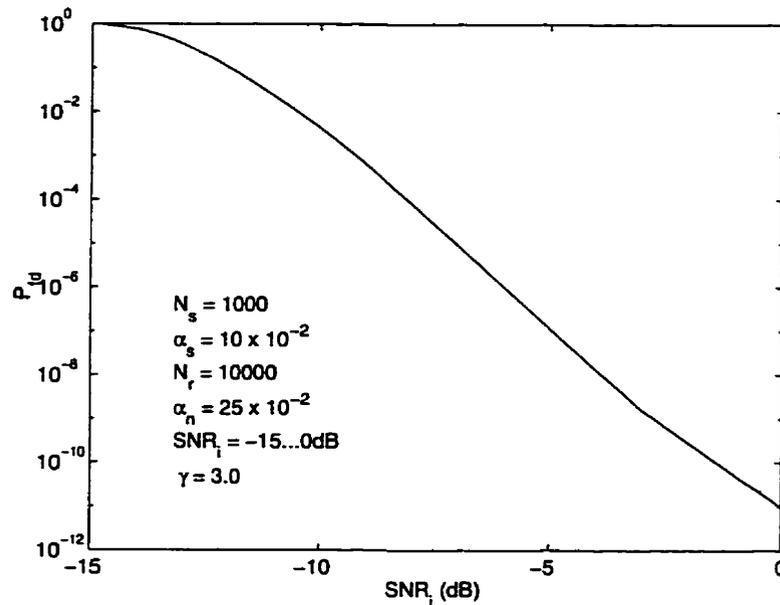


Figure 3.9: Probability of false detection vs. input SNR.

interpret such results in light of the assumptions made by the theoretical model, which generally do not hold in practice. For example, real images are typically not stationary, nor are their autocorrelation functions exactly triangular. Signal and noise peak PDFs are also never perfectly Gaussian. Of greater consequence, however, is the correlation that exists between the target and scene. If there exist regions in the scene which are similar in structure to the target, the probability of making a false detection can increase significantly. Because the model in its current form does not specifically account for the correlation between target and scene, it is not able to accurately *predict* false alarm rates. However, we suggest in Chapter 4 that even when the target and scene are correlated, the *qualitative* effects on the false alarm rate of varying the model parameters are still preserved.

## 3.4 Summary

Through computer simulation, we have experimentally confirmed the theory developed in Chapter 2. Because our model does not attempt to describe the complex statistics of natural images, specific false alarm rates computed using the model should not be considered a true representation of expected performance when real images are used. However, the model *does* provide a qualitative sense of how each model parameter influences the false alarm rate, since the model does describe the fundamental processes that occur when real targets are cross-correlated with real images. Specifically, we note that

- reducing the difference in area between the image and target results in fewer independent noise peaks, thus lowering the false detection rate.
- increasing the input SNR also lowers the false detection rate by reducing the variance of the noise peaks
- increasing either the signal or noise bandwidth increases the number of independent noise peaks and simultaneously reduces the signal and noise peak variance. These two opposing processes result in a false detection curve which rises initially, but then falls after having reached a maximum.

# Chapter 4

## Considerations in Real Images

In this chapter we use the insight gained from our theoretical and experimental results to select “optimal” targets from real images. As a control, “sub-optimal” targets are also selected. The results for both classes of targets are then compared. In our discussion, we make no attempt to be rigorous. Rather, through examples, we make arguments that suggest the *plausibility* of certain claims.

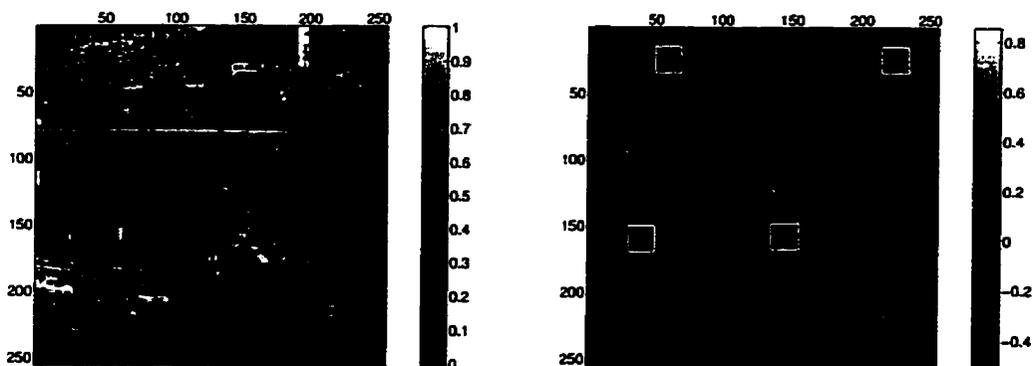
### 4.1 Target Selection

In our simulation, all five model parameters were allowed to vary. However in a real image, the image size and the bandwidth of the scene noise are fixed, leaving only the target size, bandwidth and power (i.e. the input SNR) as variables. The cost function that relates these parameters has already been determined—the probability of false detection. Thus, the task is to search the image for the sub-region that possesses the optimal combination of the above three parameters, i.e., that which yields the lowest cost.

In our experiment, a theoretically optimal target,  $20 \times 20$  pixels in size, was selected from each of the two  $256 \times 256$  test images shown in Figures 4.1(a) and 4.2(a). Each image was brought into better conformity with the model by first removing the local mean (Figures 4.1(b), 4.2(b)). Assuming local stationarity, this was achieved by applying a target-sized moving window to each image and subtracting the estimate of the mean at each pixel. The same moving window was then applied to each of the resulting zero-mean images and an estimate of the autocorrelation at each pixel was obtained. The  $x$  and  $y$  correlation lengths were given by the width at half maximum of the autocorrelation estimate, which was assumed to be triangular. A local bandwidth estimate was obtained by averaging the two correlation lengths and inverting the result (Figures 4.1(c), 4.2(c)). The power within the window at a particular point was given simply by the value of the peak of the autocorrelation (Figures 4.1(d), 4.2(d)). The bandwidth and power of the noise were estimated by averaging the local bandwidth and power estimates over the *entirety* of each image. Because the target size was small in comparison to that of the image, the contributions of the target to those averages in each case were sufficiently small such that they did not need to be specifically accounted for. All the parameters were then inserted into the model (Equation 2.43) and a cost computed for each subregion (Figures 4.1(e), 4.2(e)). The variances,  $\sigma_{c_0}^2$  and  $\sigma_{c_{n,s}}^2$ , were scaled by an arbitrary, positive factor to ensure that none of the costs were smaller than machine precision<sup>1</sup>. This had no effect on the location of the minimum. The subregion with the lowest cost in each image—Target 1 in Figures 4.1(e) and 4.2(b)—was then designated the “optimal” target. For comparison, three “sub-optimal” targets were also selected from various regions in each image. Parameter estimates are shown in Table 4.1,

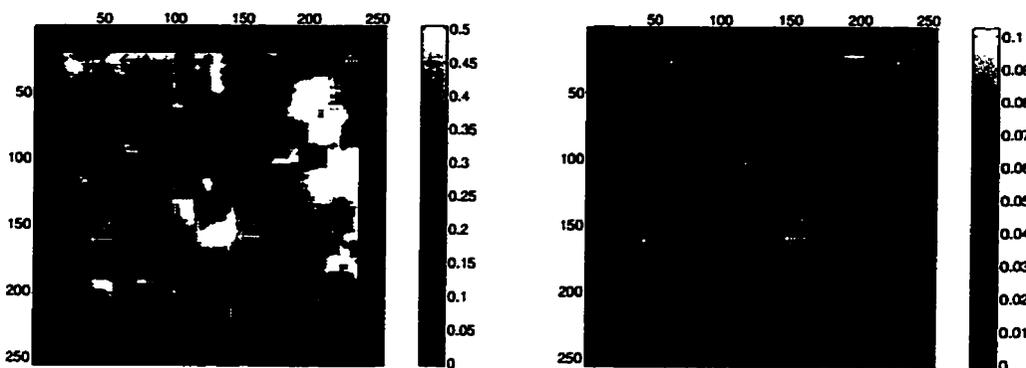
---

<sup>1</sup>When this was not done, many of the sub-regions had a cost of 0.



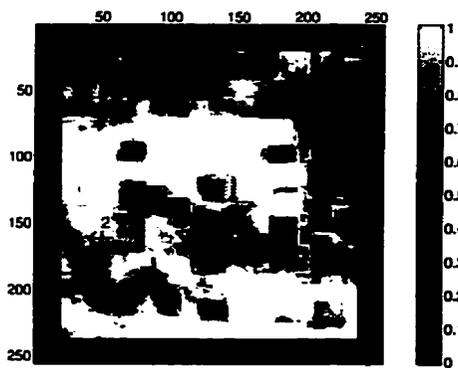
(a) Raw image

(b) Mean removed



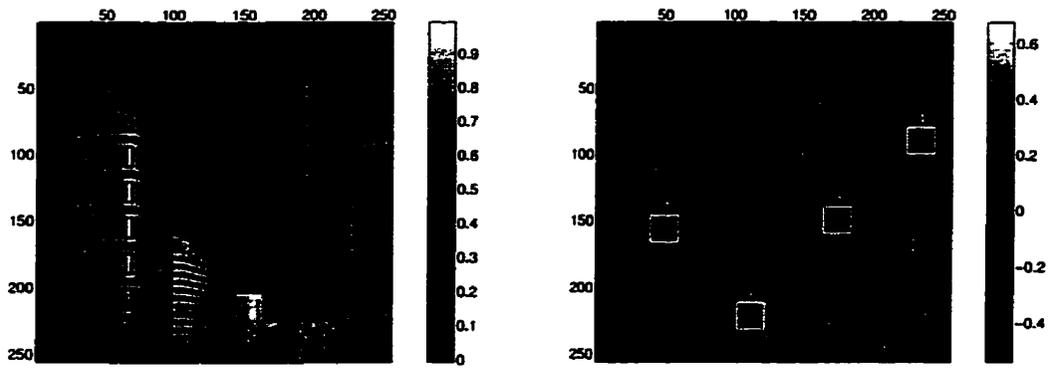
(c) Bandwidth

(d) Power



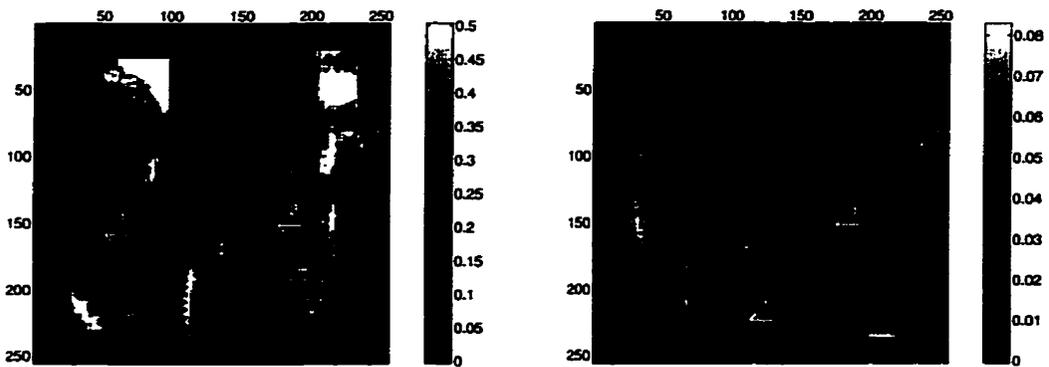
(e) Cost

Figure 4.1: Intensity maps for Image 1.



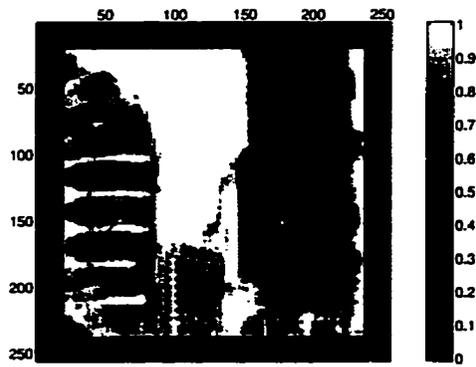
(a) Raw image

(b) Mean removed



(c) Bandwidth

(d) Power



(e) Cost

Figure 4.2: Intensity maps for Image 2.

Image	$\alpha_n$	$\sigma_n^2$	Target	$\alpha_s$	$\sigma_s^2$	Cost
1	0.26	0.035	1	0.50	0.047	$2.2 \times 10^{-8}$
			2	0.20	0.024	$8.6 \times 10^{-1}$
			3	0.50	0.037	$9.2 \times 10^{-8}$
			4	0.25	0.034	$2.1 \times 10^{-1}$
2	0.25	0.021	1	0.50	0.057	$7.2 \times 10^{-8}$
			2	0.17	0.061	$9.7 \times 10^{-1}$
			3	0.33	0.043	$1.2 \times 10^{-2}$
			4	0.25	0.026	$6.1 \times 10^{-1}$

Table 4.1: Parameter estimates.

## 4.2 Performance Evaluation

To gain a sense of the *relative* “optimality” of each of the targets in a realistic situation, a limited simulation was carried out. Changes in the image due to factors such as small angle rotations, background noise, etc., was simulated by adding white, Gaussian noise of variance 0.5 and 0.2 to Figures 4.1(b) and 4.2(b) respectively. Each of the targets in the original zero-mean images was then cross-correlated with each of the images containing the additive noise (total of 100 realizations). The actual targets detected in each case, as well as the associated false detection rate and *primary-to-secondary ratio* (PSR) [37] are shown in Figures 4.3 and 4.4. The PSR is a measure of the correlation between the target and scene and is defined as the ratio of the true correlation peak to the closest interfering peak in the absence of added background noise.

The PSRs in Image 1 (Figure 4.3) suggest that the regions from which the targets were selected are weakly correlated with the surrounding scene. Because this is consistent with the assumption of our model that the target and scene are uncorrelated, we can assume that the relative optimality of the selected targets can be inferred from the cost function. We observe, in fact, that the relative optimality of the selected targets as indicated by the

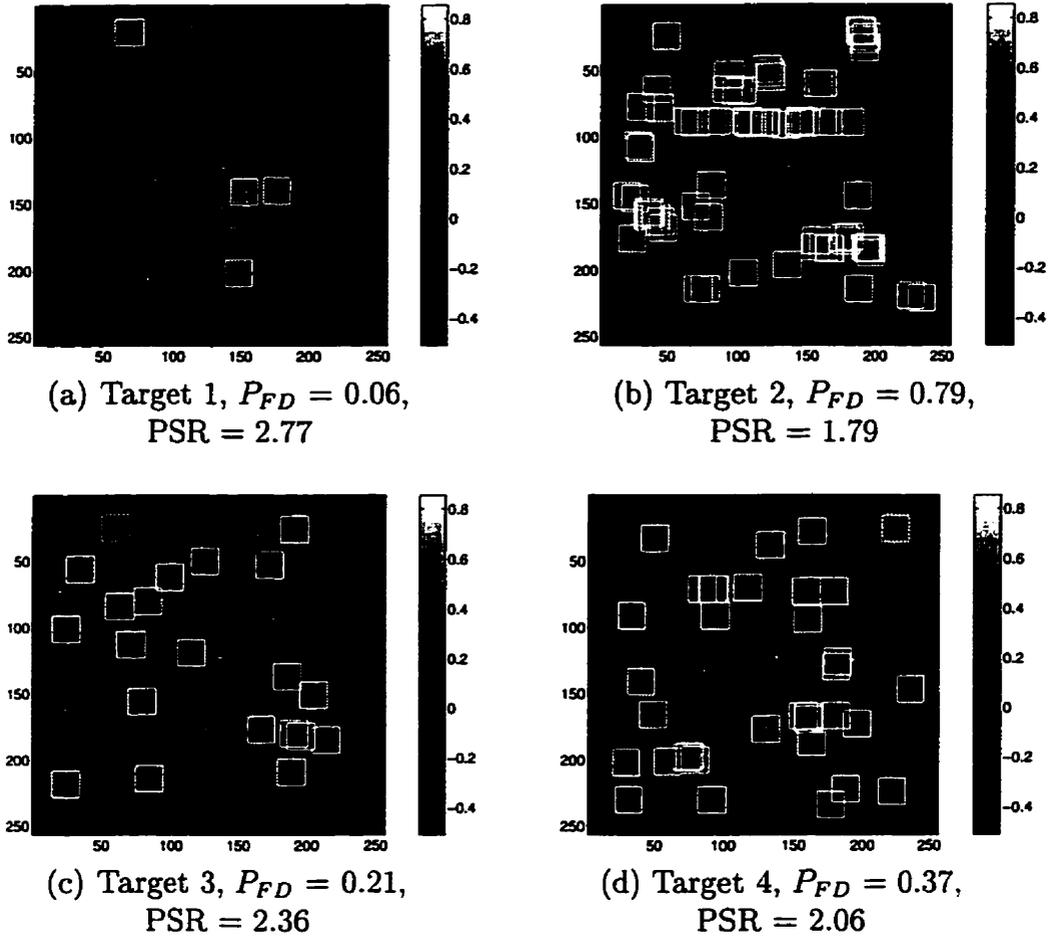


Figure 4.3: Squares indicate the location of the peak of the cross-correlation for all 100 realizations of Image 1. The dotted square denotes the true target. Notice that some detections overlap. False detection rates ( $P_{FD}$ ) and primary-to-secondary ratios (PSR) are also shown.

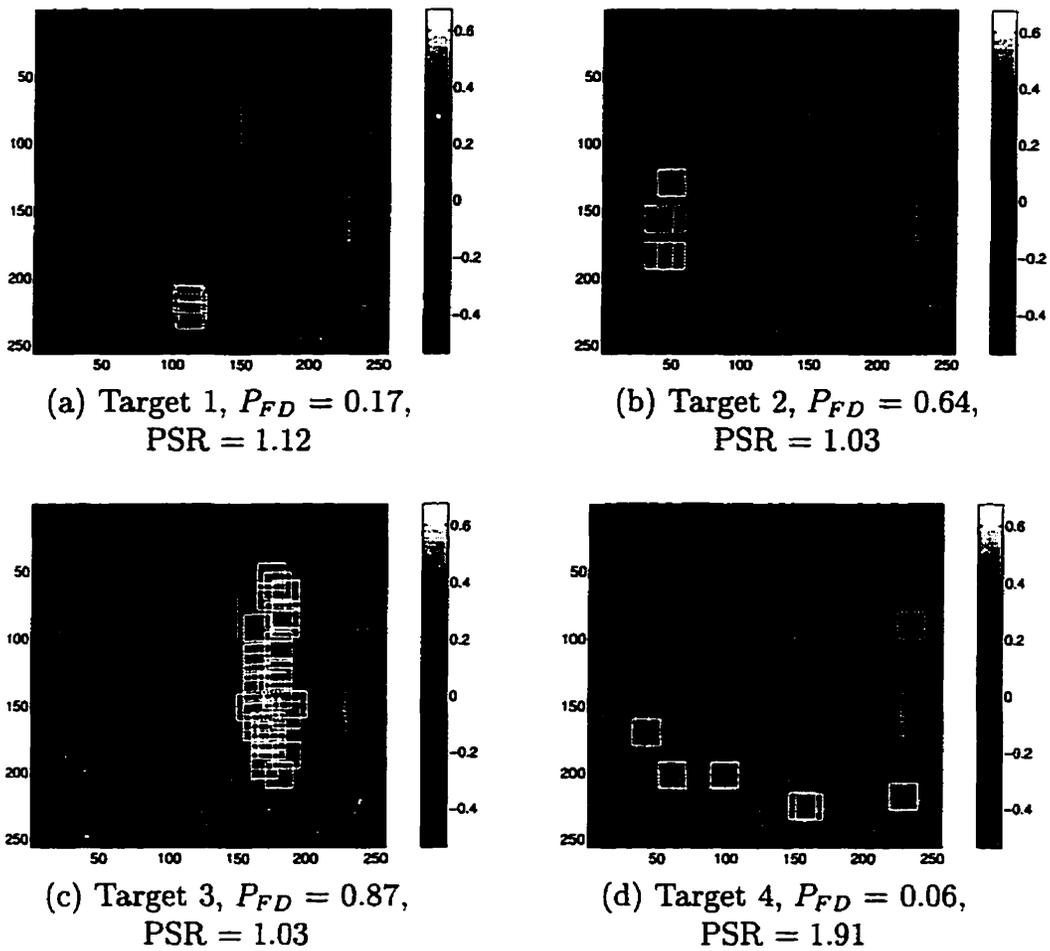


Figure 4.4: Detection locations,  $P_{FD}$ , and PSRs in Image 2.

cost column in Table 4.1 is consistent with the false detection rates that were observed in the simulation (Figure 4.3). For example, of the four targets, Target 1—the “optimal” target—has a combined bandwidth and power which is larger than all other targets, as shown in Table 4.1. We know from our model that this contributes to a reduction in the variance of the scene noise peaks, which results in a false detection rate that is smaller than those of the other targets. In contrast, because Target 2 has the lowest combined bandwidth and power, the variance of the noise peaks is larger than that of the other targets, and thus so is the false detection rate. Although Image 1 is an example of an image that conforms fairly well to the assumptions of the theoretical model, it is important to note that other factors, which we discuss below, also affect the error rate.

### **Effect of Target-Scene Correlation**

It is apparent (Figure 4.4) that the “optimal” target (Target 1) in Image 2 has a larger false detection rate than a “sub-optimal” target (Target 4). This is directly attributable to the correlation between the targets and the scene. It is apparent that the PSR of Target 4 is much larger (1.91) than that of Target 1, and thus its false detection rate is lower. Conversely, Target 1 is strongly correlated with the immediately surrounding scene—as is suggested by its PSR of nearly 1—and its false detection rate is higher.

Targets 2 and 3 are also strongly correlated with the surrounding scene and thus their PSRs are approximately unity. However, because the region that is correlated with Target 3 is larger than the regions that are correlated with Targets 1 or 2, the false detection rate of Target 3 is substantially higher. This is seen in the number of different detection locations localized in each

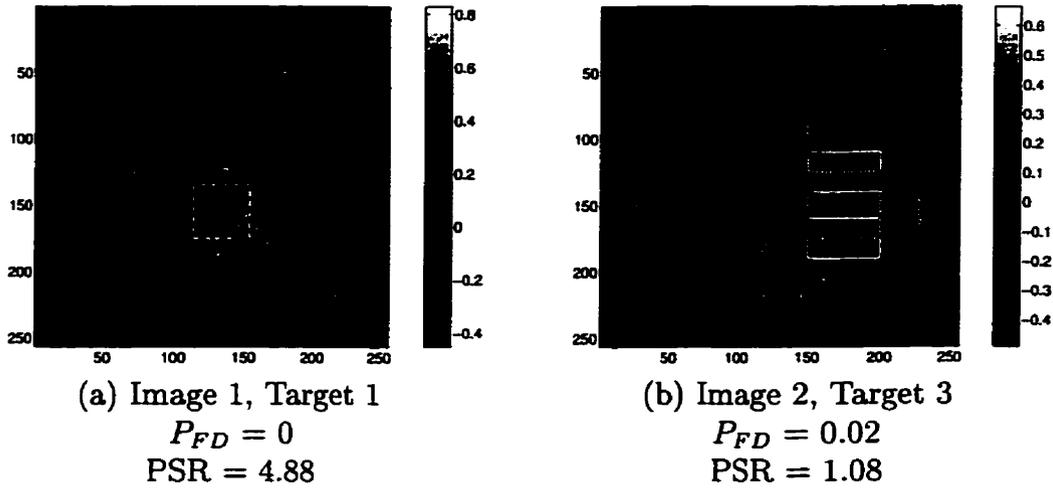


Figure 4.5: Effect of increasing target size on target-scene correlation.

region of Figures 4.4(a), (b) and (c).

Our above discussion suggests that in some practical situations where the image contains strongly correlated structures, selecting “optimal” targets only according to their size, bandwidth and power is insufficient; their correlation with the scene must also be accounted for.

### Effect of Target Size on Target-Scene Correlation

It is important to recognize that the correlation between the target and scene is not independent of the target characteristics. To illustrate this, we increased the size of Target 1 in Image 1 and Target 3 in Image 2 to  $40 \times 40$  and  $50 \times 50$  pixels respectively. The simulation above was then repeated, with the noise variances unchanged from before. Figure 4.5 shows the results. It is evident that increasing the target size has an effect on the correlation between the target and scene and thus on the false detection rate as well.

The effect may be partially explained by examining the influence of the

target size on the cross-correlation between the target and the original image. In Chapter 2, expressions for the variance of the scene noise peaks as well as the number of noise peaks were derived (Equations 2.45 and 2.46). It was shown that increasing the target size decreased the variance as well as the number of noise peaks. Thus, increasing the target size in a single image—as we have done here—results in a lower spatial variance of the cross-correlation in the scene and thus the relative height of competing correlation peaks is reduced. This, in conjunction with fewer peaks, results in fewer false detections in the presence of additive background noise. This is seen in the increased PSR and decreased false detection rate of the target in Figure 4.5(a).

The nominal increase in the PSR of Figure 4.5(b) is attributable to the strong correlation that still exists between the target and scene, even though the target size has increased. For the PSR to increase, the target would have to include regions that are weakly correlated with the scene.

### 4.3 Summary

Using the optimality criteria developed in Chapter 2, (i.e. the theoretical probability of detection) an “optimal” target was selected from each of two real images. After adding noise to each image, the false detection rates of “optimal” and “sub-optimal” targets were then determined. The relative optimality of the targets in the first image as indicated by their assigned costs, was consistent with the actual false detection rates. This was because the image conformed well to our model, in that the correlation between the targets and scene was weak. In contrast, the second image contained regions which were strongly correlated with the “optimal” target. This resulted in the “optimal” target having a false detection rate that was higher than one of the “sub-optimal”

targets.

Increasing the target size decreased the relative height of the scene noise peaks, which led to a corresponding decrease in the false detection rate. This was consistent with the results derived in Chapters 2 and 3, where it was shown that the variance of the noise peaks decreases as the target size is increased.

# Chapter 5

## Conclusions

### 5.1 Summary of Work

Spatial cross-correlation detectors are commonly used to detect the presence of a known target in an image. In such a case, all target characteristics are given and the “quality” of the target is irrelevant. However, when targets are used as frame of reference markers for head tracking, optimal targets may be chosen, where optimality is defined in terms of the probability of false detection.

In this thesis, we examined the influence of the size, bandwidth and power of the target and scene on the false detection rate. In an effort to characterize certain key features of real images, a stochastic image model was developed in which a target of a certain bandwidth and power is surrounded by uncorrelated scene noise of a different bandwidth and power. This differed from the models used by Mostafavi/Smith and Ianiello, in that the signal and noise were spatially disjoint, and thus the image as a whole was non-stationary.

The false detection probability was derived by considering the signal peak variance, the noise peak variance and the number of independent noise

peaks. These three variables represented the fundamental processes that drive the false detection rate. Because specific (i.e. triangular) signal and noise autocorrelation functions were chosen, it was possible to express these variables as closed-form functions of the model parameters. This allowed us to uncover the individual effects of each of the parameters on the above processes. These effects were summarized in Section 3.4.

The validity of our model was confirmed experimentally through computer simulations. Theoretical false detection rates were generally within the 95% confidence intervals of the experimentally determined rates. Deviations were attributed to the invalidity of our assumption regarding the Gaussian nature of the probability density function at low space-bandwidth products.

It was recognized that the model made certain simplifying assumptions which, while of great value in the theoretical analysis, were not necessarily valid in real images. In particular, target and scene are not uncorrelated. Correlation between the target and other regions in the image results in an actual probability of false detection that is potentially much higher than what is predicted theoretically.

In our examination of real images, “optimal” targets were selected from each of two real images according to our optimality criteria. Changes in the image due to factors such as background noise or small-angle rotations were simulated by adding white noise to each image. False detection rates were then determined for both “optimal” and “sub-optimal” targets. In the first image, correlation between the targets and scene was weak and thus the image conformed well to the model. Consequently, the relative optimality of the targets as predicted by the model was consistent with the actual false detection rates. In contrast, many of the targets—including the “optimal” target—in the second image were strongly correlated with the scene. The relative optimality

of the targets as predicted by the model thus was not reflected in the actual error rates.

It is apparent that because the model does not completely describe the statistics of the scene, the expected false detection rate cannot be computed accurately. However, because the underlying processes are the same irrespective of the scene, the model *is* able to predict the qualitative false detection trends when particular target characteristics are varied.

## 5.2 Future Work

Although we have a qualitative sense of the effect of target-scene correlation and geometric distortions on the probability of false detection, a formal modelling of their contribution still remains.

We have already seen that correlation between the target and scene can increase the false alarm rate significantly. One approach to studying this effect is to include finite regions in the scene which are partially correlated with the target. By specifying the mean and variance of each of the correlated peaks (which may be computed from the statistics of each region) and the number of such peaks, their contribution to the probability of false detection may be determined in the same way that it was determined for the uncorrelated noise.

Much of the foundational work relating to the effect of geometric distortions, such as rotation and scale, on cross-correlation has already been done by Mostafavi and Smith [45]. Thus, what remains is the tailoring of their basic mathematical techniques to accommodate the specific features of our image model, such as the spatial disjointness between target and scene. Some of the qualitative effects of geometric distortions on the false detection rate are not difficult to see. For example, we know from our existing model that high

bandwidth targets are desirable. However, as we noted in Section 1.3.2, correlation peaks of high bandwidth targets are particularly sensitive to geometric distortions. This conflict of requirements suggests that at a given distortion (say, a particular angle of rotation) there exists an optimal target bandwidth such that the false detection rate is minimized.

One must be mindful that making these amendments to the model will not necessarily improve the model's ability to predict actual false detection rates; to do that would require detailed knowledge of the statistics of the operating environment. However, because the fundamental processes at work are common to both the theoretical and actual situations, an understanding of how the various factors above affect such processes provides insight into how to select targets such that detection errors are minimized.

# Appendix A

## Derivations

### A.1 Synthesis of $\mathbf{s}[\vec{n}]$ and $\mathbf{n}[\vec{n}]$

For the linear system defined by

$$\mathbf{s}[\vec{n}] = \mathbf{q}[\vec{n}] * h_s[\vec{n}], \quad (\text{A.1})$$

where  $\mathbf{q}[\vec{n}]$  is an arbitrary input signal and  $h[\vec{n}]$  is a linear filter, the output autocorrelation,  $R_{ss}[\vec{n}]$ , is shown [50] to be:

$$R_{ss}[\vec{n}] = R_{qq}[\vec{n}] * h_s[\vec{n}] * h_s[-\vec{n}]. \quad (\text{A.2})$$

This suggests that our signal,  $\mathbf{s}[\vec{n}]$ , which has an autocorrelation given by (2.12), may be synthesized by applying a linear filter to an input signal with autocorrelation  $R_{qq}[\vec{n}]$ . In the simplest case,  $\mathbf{q}[\vec{n}]$  is zero-mean, stationary white noise, with variance  $\sigma_s^2$  i.e.  $R_{qq}[\vec{n}] = \sigma_s^2 \delta[\vec{n}]$ . Thus, (A.2) becomes

$$\begin{aligned} R_{ss}[\vec{n}] &= \sigma_s^2 \delta[\vec{n}] * h_s[\vec{n}] * h_s[-\vec{n}] \\ &= \sigma_s^2 (h_s[\vec{n}] * h_s[-\vec{n}]) \\ &= \sigma_s^2 \sum_{\vec{k}} h_s[\vec{k}] h_s[\vec{k} - \vec{n}], \end{aligned} \quad (\text{A.3})$$

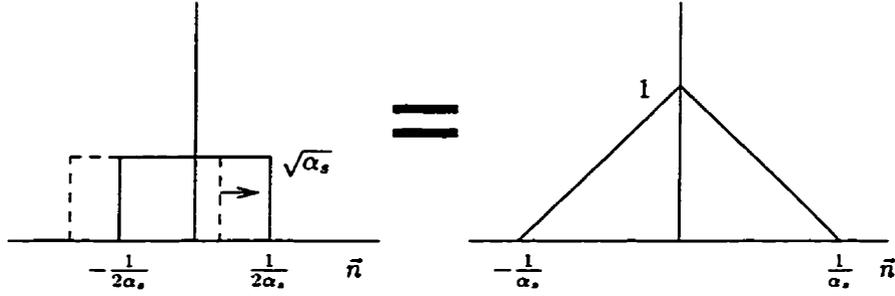


Figure A.1: Graphical representation of Equation A.4.

where  $\vec{k}$  is over the domain of  $h_s[\vec{n}]$ . We may determine  $h_s[\vec{n}]$  by equating (A.3) and (2.12) and dividing both sides by  $\sigma_s^2$ :

$$\sum_{\vec{k}} h_s[\vec{k}] h_s[\vec{k} - \vec{n}] = \text{tri}[\alpha_s \vec{n}] \quad (\text{A.4})$$

That is, the auto-correlation of some filter function  $h_s[\vec{n}]$  generates a triangle function  $\frac{2}{\alpha_s}$  samples in length. We recognize that a rectangular pulse  $\frac{1}{\alpha_s}$  samples in width satisfies (A.4) (Figure A.1). Thus, after normalizing the energy of the filter to unity, we obtain

$$h_s[\vec{n}] = \begin{cases} \sqrt{\alpha_s} & \vec{n} \in \left[ \frac{1}{2\alpha_s}, \frac{1}{2\alpha_s} - 1 \right] \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A.5})$$

Similarly, for the noise,

$$h_n[\vec{n}] = \begin{cases} \sqrt{\alpha_n} & \vec{n} \in \left[ \frac{1}{2\alpha_n}, \frac{1}{2\alpha_n} - 1 \right] \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A.6})$$

Thus, a random 1-D signal with a triangular autocorrelation may be generated by convolving white noise with a rectangular impulse response filter.

The “physical” meaning of bandwidth here may be seen by noting that the frequency response of  $h_s[\vec{n}]$ —which is  $\frac{1}{2\alpha_s}$  samples in length—is a sinc function whose main lobe is  $2\alpha_s$  inverse samples in width. Thus, as the length of

the filter (or equivalently, the width of the autocorrelation function) increases, the bandwidth decreases, as expected.

## A.2 Mean of $\mathbf{c}[\vec{n}]$

In determining the mean and covariance of  $\mathbf{c}[\vec{n}]$ , we follow the approach of Mostafavi and Smith [45].

Taking the expected value of (2.20) gives an expression for the ensemble mean of the cross-correlation:

$$E \{ \mathbf{c}[\vec{n}] \} = E \{ \mathbf{c}_{\bar{s}\bar{s}}[\vec{n}] \} + E \{ \mathbf{c}_{\bar{n}\bar{s}}[\vec{n}] \}. \quad (\text{A.7})$$

From (2.21) and (2.2), the first term in (A.7) becomes

$$\begin{aligned} E \{ \mathbf{c}_{\bar{s}\bar{s}}[\vec{n}] \} &= E \left\{ a_s^{-1} \sum_{\vec{k}} \bar{\mathbf{s}}[\vec{k}] \bar{\mathbf{s}}[\vec{k} - \vec{n}] \right\} \\ &= a_s^{-1} \sum_{\vec{k}} E \left\{ \mathbf{s}[\vec{k}] \mathbf{s}[\vec{k} - \vec{n}] \right\} w_s[\vec{k}] w_s[\vec{k} - \vec{n}] \\ &= a_s^{-1} R_{ss}[-\vec{n}] \sum_{\vec{k}} w_s[\vec{k}] w_s[\vec{k} - \vec{n}]. \end{aligned} \quad (\text{A.8})$$

Recognizing that the summation in (A.8) is a 2-D triangle function  $N_s^2$  in height and  $2N_s \times 2N_s$  in area, (A.8) becomes

$$E \{ \mathbf{c}_{\bar{s}\bar{s}}[\vec{n}] \} = a_s^{-1} R_{ss}[\vec{n}] N_s^2 \text{tri} \left[ \frac{1}{N_s} \vec{n} \right] \quad (\text{A.9})$$

$$= R_{ss}[\vec{n}] \text{tri} \left[ \frac{1}{N_s} \vec{n} \right], \quad (\text{A.10})$$

where we have also inserted (2.16) and used the fact that  $R_{ss}[\vec{n}] = R_{ss}[-\vec{n}]$ .

From (2.5) and (2.22), the second term in (A.7) becomes

$$\begin{aligned}
E \{ \mathbf{c}_{\bar{n}\bar{s}}[\bar{\mathbf{n}}] \} &= E \left\{ a_s^{-1} \sum_{\bar{k}} \bar{\mathbf{n}}[\bar{k}] \bar{\mathbf{s}}[\bar{k} - \bar{\mathbf{n}}] \right\} \\
&= a_s^{-1} \sum_{\bar{k}} E \left\{ \mathbf{n}[\bar{k}] \mathbf{s}[\bar{k} - \bar{\mathbf{n}}] \right\} w_n[\bar{k}] w_s[\bar{k} - \bar{\mathbf{n}}] \\
&= 0.
\end{aligned} \tag{A.11}$$

where we have used the fact that  $\mathbf{n}[\bar{\mathbf{n}}]$  and  $\mathbf{s}[\bar{\mathbf{n}}]$  are uncorrelated, i.e.,  $E \left\{ \mathbf{n}[\bar{k}] \mathbf{s}[\bar{k} - \bar{\mathbf{n}}] \right\} = 0$ .

Inserting (A.9) and (A.11) into (A.7) yields the mean:

$$\begin{aligned}
E \{ \mathbf{c}[\bar{\mathbf{n}}] \} &= E \{ \mathbf{c}_{\bar{s}\bar{s}}[\bar{\mathbf{n}}] \} \\
&= R_{ss}[\bar{\mathbf{n}}] \text{tri} \left[ \frac{1}{N_s} \bar{\mathbf{n}} \right].
\end{aligned} \tag{A.12}$$

Thus, the ensemble averaged cross-correlation has only one contributor: the signal itself. The mean at  $\bar{\mathbf{n}} = \vec{0}$ , i.e., at the true signal peak, is thus

$$\begin{aligned}
\mu_{c_0} &= E \left\{ \mathbf{c}[\vec{0}] \right\} \\
&= R_{ss}[0] \text{tri}[0] \\
&= 1,
\end{aligned} \tag{A.13}$$

as expected.

### A.3 Autocorrelation of $\mathbf{c}[\bar{\mathbf{n}}]$

From (2.20), the autocorrelation of  $\mathbf{c}[\bar{\mathbf{n}}]$  is given by

$$\begin{aligned}
E \{ \mathbf{c}[\bar{\mathbf{n}}] \mathbf{c}[\bar{\mathbf{m}}] \} &= E \{ \mathbf{c}_{\bar{s}\bar{s}}[\bar{\mathbf{n}}] \mathbf{c}_{\bar{s}\bar{s}}[\bar{\mathbf{m}}] \} + E \{ \mathbf{c}_{\bar{s}\bar{s}}[\bar{\mathbf{n}}] \mathbf{c}_{\bar{n}\bar{s}}[\bar{\mathbf{m}}] \} \\
&\quad + E \{ \mathbf{c}_{\bar{n}\bar{s}}[\bar{\mathbf{n}}] \mathbf{c}_{\bar{s}\bar{s}}[\bar{\mathbf{m}}] \} + E \{ \mathbf{c}_{\bar{n}\bar{s}}[\bar{\mathbf{n}}] \mathbf{c}_{\bar{n}\bar{s}}[\bar{\mathbf{m}}] \}.
\end{aligned} \tag{A.14}$$

This expression may be simplified by observing that because  $\bar{\mathbf{s}}[\bar{\mathbf{n}}]$  and  $\bar{\mathbf{n}}[\bar{\mathbf{n}}]$  are uncorrelated, terms 2 and 3 vanish. This leaves

$$E \{ \mathbf{c}[\bar{\mathbf{n}}] \mathbf{c}[\bar{\mathbf{m}}] \} = E \{ \mathbf{c}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}[\bar{\mathbf{n}}] \mathbf{c}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}[\bar{\mathbf{m}}] \} + E \{ \mathbf{c}_{\bar{\mathbf{n}}\bar{\mathbf{s}}}[\bar{\mathbf{n}}] \mathbf{c}_{\bar{\mathbf{n}}\bar{\mathbf{s}}}[\bar{\mathbf{m}}] \}. \quad (\text{A.15})$$

Thus, the total autocorrelation of the cross-correlation is equal to the sum of the autocorrelations contributed by the cross-correlations of the signal and the noise. We consider each term in turn.

Inserting (2.21) into  $E \{ \mathbf{c}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}[\bar{\mathbf{n}}] \mathbf{c}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}[\bar{\mathbf{m}}] \}$  yields

$$\begin{aligned} E \{ \mathbf{c}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}[\bar{\mathbf{n}}] \mathbf{c}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}[\bar{\mathbf{m}}] \} &= E \left\{ a_s^{-1} \sum_{\bar{\mathbf{k}}} \bar{\mathbf{s}}[\bar{\mathbf{k}}] \bar{\mathbf{s}}[\bar{\mathbf{k}} - \bar{\mathbf{n}}] a_s^{-1} \sum_{\bar{\mathbf{l}}} \bar{\mathbf{s}}[\bar{\mathbf{l}}] \bar{\mathbf{s}}[\bar{\mathbf{l}} - \bar{\mathbf{m}}] \right\} \\ &= a_s^{-2} \sum_{\bar{\mathbf{k}}} \sum_{\bar{\mathbf{l}}} E \left\{ \bar{\mathbf{s}}[\bar{\mathbf{k}}] \bar{\mathbf{s}}[\bar{\mathbf{l}}] \bar{\mathbf{s}}[\bar{\mathbf{k}} - \bar{\mathbf{n}}] \bar{\mathbf{s}}[\bar{\mathbf{l}} - \bar{\mathbf{m}}] \right\}. \end{aligned} \quad (\text{A.16})$$

Equation (A.16) may be simplified by noting that for four zero-mean, jointly Gaussian random variables [50],

$$\begin{aligned} E \{ P_1 P_2 P_3 P_4 \} &= E \{ P_1 P_2 \} E \{ P_3 P_4 \} + E \{ P_1 P_3 \} E \{ P_2 P_4 \} \\ &\quad + E \{ P_1 P_4 \} E \{ P_2 P_3 \}. \end{aligned}$$

Thus,

$$\begin{aligned} E \{ \mathbf{c}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}[\bar{\mathbf{n}}] \mathbf{c}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}[\bar{\mathbf{m}}] \} &= a_s^{-2} \sum_{\bar{\mathbf{k}}} \sum_{\bar{\mathbf{l}}} E \left\{ \bar{\mathbf{s}}[\bar{\mathbf{k}}] \bar{\mathbf{s}}[\bar{\mathbf{l}}] \right\} E \left\{ \bar{\mathbf{s}}[\bar{\mathbf{k}} - \bar{\mathbf{n}}] \bar{\mathbf{s}}[\bar{\mathbf{l}} - \bar{\mathbf{m}}] \right\} \\ &\quad + a_s^{-2} \sum_{\bar{\mathbf{k}}} \sum_{\bar{\mathbf{l}}} E \left\{ \bar{\mathbf{s}}[\bar{\mathbf{k}}] \bar{\mathbf{s}}[\bar{\mathbf{k}} - \bar{\mathbf{n}}] \right\} E \left\{ \bar{\mathbf{s}}[\bar{\mathbf{l}}] \bar{\mathbf{s}}[\bar{\mathbf{l}} - \bar{\mathbf{m}}] \right\} \\ &\quad + a_s^{-2} \sum_{\bar{\mathbf{k}}} \sum_{\bar{\mathbf{l}}} E \left\{ \bar{\mathbf{s}}[\bar{\mathbf{k}}] \bar{\mathbf{s}}[\bar{\mathbf{l}} - \bar{\mathbf{m}}] \right\} E \left\{ \bar{\mathbf{s}}[\bar{\mathbf{l}}] \bar{\mathbf{s}}[\bar{\mathbf{k}} - \bar{\mathbf{n}}] \right\}. \end{aligned} \quad (\text{A.17})$$

We note that the second term in (A.17) may be rewritten as:

$$\begin{aligned}
& a_s^{-2} \sum_{\vec{k}} \sum_{\vec{l}} E \left\{ \bar{\mathbf{s}}[\vec{k}] \bar{\mathbf{s}}[\vec{k} - \vec{n}] \right\} E \left\{ \bar{\mathbf{s}}[\vec{l}] \bar{\mathbf{s}}[\vec{l} - \vec{m}] \right\} \\
&= a_s^{-1} \sum_{\vec{k}} E \left\{ \bar{\mathbf{s}}[\vec{k}] \bar{\mathbf{s}}[\vec{k} - \vec{n}] \right\} a_s^{-1} \sum_{\vec{l}} E \left\{ \bar{\mathbf{s}}[\vec{l}] \bar{\mathbf{s}}[\vec{l} - \vec{m}] \right\} \\
&= E \left\{ a_s^{-1} \sum_{\vec{k}} \bar{\mathbf{s}}[\vec{k}] \bar{\mathbf{s}}[\vec{k} - \vec{n}] \right\} E \left\{ a_s^{-1} \sum_{\vec{l}} \bar{\mathbf{s}}[\vec{l}] \bar{\mathbf{s}}[\vec{l} - \vec{m}] \right\} \\
&= E \left\{ \mathbf{c}_{\bar{s}\bar{s}}[\vec{n}] \right\} E \left\{ \mathbf{c}_{\bar{s}\bar{s}}[\vec{m}] \right\}. \tag{A.18}
\end{aligned}$$

Substitution of (2.2) and (A.18) into (A.17) gives

$$\begin{aligned}
& E \left\{ \mathbf{c}_{\bar{s}\bar{s}}[\vec{n}] \mathbf{c}_{\bar{s}\bar{s}}[\vec{m}] \right\} - E \left\{ \mathbf{c}_{\bar{s}\bar{s}}[\vec{n}] \right\} E \left\{ \mathbf{c}_{\bar{s}\bar{s}}[\vec{m}] \right\} \\
&= a_s^{-2} \sum_{\vec{k}} \left[ \sum_{\vec{l}} w_s[\vec{k}] w_s[\vec{l}] w_s[\vec{k} - \vec{n}] w_s[\vec{l} - \vec{m}] \right] \\
&\quad \left[ E \left\{ \mathbf{s}[\vec{k}] \mathbf{s}[\vec{l}] \right\} E \left\{ \mathbf{s}[\vec{k} - \vec{n}] \mathbf{s}[\vec{l} - \vec{m}] \right\} \right. \\
&\quad \left. + E \left\{ \bar{\mathbf{s}}[\vec{k}] \bar{\mathbf{s}}[\vec{l} - \vec{m}] \right\} E \left\{ \mathbf{s}[\vec{l}] \mathbf{s}[\vec{k} - \vec{n}] \right\} \right] \\
&= a_s^{-2} \sum_{\vec{k}} \left[ \sum_{\vec{l}} w_s[\vec{k}] w_s[\vec{l}] w_s[\vec{k} - \vec{n}] w_s[\vec{l} - \vec{m}] \right] \\
&\quad \left[ R_{ss}[\vec{l} - \vec{k}] R_{ss}[\vec{l} - \vec{k} - \vec{m} + \vec{n}] \right. \\
&\quad \left. + R_{ss}[\vec{l} - \vec{k} - \vec{m}] R_{ss}[\vec{l} - \vec{k} - \vec{n}] \right]. \tag{A.19}
\end{aligned}$$

Making the change of variable  $\vec{z} = \vec{l} - \vec{k}$  in (A.19) yields an exact expression for the autocovariance of  $\mathbf{c}_{\bar{s}\bar{s}}[\vec{n}]$ :

$$\begin{aligned}
& E \left\{ \mathbf{c}_{\bar{s}\bar{s}}[\vec{n}] \mathbf{c}_{\bar{s}\bar{s}}[\vec{m}] \right\} - E \left\{ \mathbf{c}_{\bar{s}\bar{s}}[\vec{n}] \right\} E \left\{ \mathbf{c}_{\bar{s}\bar{s}}[\vec{m}] \right\} = \\
& a_s^{-2} \sum_{\vec{z}} \left[ \sum_{\vec{k}} w_s[\vec{k}] w_s[\vec{z} + \vec{k}] w_s[\vec{k} - \vec{n}] w_s[\vec{z} + \vec{k} - \vec{m}] \right] \\
& \quad \left[ R_{ss}[\vec{z}] R_{ss}[\vec{z} - \vec{m} + \vec{n}] + R_{ss}[\vec{z} - \vec{m}] R_{ss}[\vec{z} - \vec{n}] \right]. \tag{A.20}
\end{aligned}$$

The variance of the peak may be found by setting  $\vec{m} = \vec{n} = \vec{0}$  in (A.20):

$$\begin{aligned}
\sigma_{c_0}^2 &= E \left\{ c_{\vec{s}\vec{s}}^2[\vec{0}] \right\} - E \left\{ c_{\vec{s}\vec{s}}[\vec{0}] \right\}^2 \\
&= a_s^{-2} \sum_{\vec{z}} \left[ \sum_{\vec{k}} w_s[\vec{k}] w_s[\vec{z} + \vec{k}] \right] 2R_{\vec{s}\vec{s}}^2[\vec{z}] \\
&= 2a_s^{-2} \sum_{\vec{z}} N_s^2 \text{tri} \left[ \frac{1}{N_s} \vec{z} \right] R_{\vec{s}\vec{s}}^2[\vec{z}]. \tag{A.21}
\end{aligned}$$

Inserting (2.12) into (A.21), and assuming that  $N_s > 1/\alpha_s$ —which is always true in our simulations—we obtain

$$\begin{aligned}
\sigma_{c_0}^2 &= 2a_s^{-1} \sigma_s^4 \sum_{\vec{z}} \text{tri} \left[ \frac{1}{N_s} \vec{z} \right] \text{tri}^2[\alpha_s \vec{z}] \\
&\approx 2a_s^{-1} \sigma_s^4 \int \text{tri} \left[ \frac{1}{N_s} \vec{z} \right] \text{tri}^2[\alpha_s \vec{z}] d\vec{z} \\
&= 2 \left[ \frac{2}{3N_s \alpha_s} - \frac{1}{6(N_s \alpha_s)^2} \right]^2. \tag{A.22}
\end{aligned}$$

In determining the second term of (A.15), we first substitute (2.22) into  $E \{ c_{\vec{n}\vec{s}}[\vec{n}] c_{\vec{n}\vec{s}}[\vec{m}] \}$  and use the fact that  $\vec{n}[\vec{n}]$  and  $\vec{s}[\vec{n}]$  are uncorrelated and independent:

$$\begin{aligned}
E \{ c_{\vec{n}\vec{s}}[\vec{n}] c_{\vec{n}\vec{s}}[\vec{m}] \} &= E \left\{ a_s^{-1} \sum_{\vec{k}} \vec{n}[\vec{k}] \vec{s}[\vec{k} - \vec{n}] a_s^{-1} \sum_{\vec{l}} \vec{n}[\vec{l}] \vec{s}[\vec{l} - \vec{m}] \right\} \\
&= a_s^{-2} \sum_{\vec{k}} \sum_{\vec{l}} E \left\{ \vec{n}[\vec{k}] \vec{n}[\vec{l}] \vec{s}[\vec{k} - \vec{n}] \vec{s}[\vec{l} - \vec{m}] \right\} \\
&= a_s^{-2} \sum_{\vec{k}} \sum_{\vec{l}} E \left\{ \vec{n}[\vec{k}] \vec{n}[\vec{l}] \right\} E \left\{ \vec{s}[\vec{k} - \vec{n}] \vec{s}[\vec{l} - \vec{m}] \right\}. \tag{A.23}
\end{aligned}$$

Inserting (2.2) and (2.5) into (A.23) we obtain

$$\begin{aligned}
E \{ \mathbf{c}_{\vec{n}\vec{s}}[\vec{n}] \mathbf{c}_{\vec{n}\vec{s}}[\vec{m}] \} &= a_s^{-2} \sum_{\vec{k}} \sum_{\vec{l}} w_n[\vec{k}] w_n[\vec{l}] w_s[\vec{k} - \vec{n}] w_s[\vec{l} - \vec{m}] \\
&\quad E \{ \mathbf{n}[\vec{k}] \mathbf{n}[\vec{l}] \} E \{ \mathbf{s}[\vec{k} - \vec{n}] \mathbf{s}[\vec{l} - \vec{m}] \} \\
&= a_s^{-2} \sum_{\vec{k}} \sum_{\vec{l}} w_n[\vec{k}] w_n[\vec{l}] w_s[\vec{k} - \vec{n}] w_s[\vec{l} - \vec{m}] \\
&\quad R_{nn}[\vec{l} - \vec{k}] R_{ss}[\vec{l} - \vec{k} - \vec{m} + \vec{n}]. \tag{A.24}
\end{aligned}$$

Making the change of variable  $\vec{z} = \vec{l} - \vec{k}$  in (A.24) yields an exact expression for the autocorrelation of  $\mathbf{c}_{\vec{n}\vec{s}}[\vec{n}]$ :

$$\begin{aligned}
E \{ \mathbf{c}_{\vec{n}\vec{s}}[\vec{n}] \mathbf{c}_{\vec{n}\vec{s}}[\vec{m}] \} &= a_s^{-2} \sum_{\vec{z}} \left[ \sum_{\vec{k}} w_n[\vec{k}] w_n[\vec{z} + \vec{k}] w_s[\vec{k} - \vec{n}] w_s[\vec{z} + \vec{k} - \vec{m}] \right] \\
&\quad R_{nn}[\vec{z}] R_{ss}[\vec{z} - \vec{m} + \vec{n}]. \tag{A.25}
\end{aligned}$$

Equation (A.25) is rather unwieldy and difficult to evaluate in general. However, since we wish only to find the variance of  $\mathbf{c}_{\vec{n}\vec{s}}[\vec{n}]$  in the stationary region  $\Omega_{c_{ns}}$ , it may be simplified. We observe that for  $\vec{n} \in \Omega_{c_{ns}}$  (i.e. when the template is completely in  $\Omega_n$ ), the inner summation in (A.25) collapses to

$$\sum_{\vec{k}} w_n[\vec{k}] w_n[\vec{z} + \vec{k}] w_s[\vec{k} - \vec{n}] w_s[\vec{z} + \vec{k} - \vec{m}] = \sum_{\vec{k}} w_s[\vec{k} - \vec{n}] w_s[\vec{z} + \vec{k} - \vec{m}]. \tag{A.26}$$

Making the change of variable  $\vec{u} = \vec{k} - \vec{n}$  in the RHS of (A.26), we obtain

$$\begin{aligned}
\sum_{\vec{k}} w_s[\vec{k} - \vec{n}] w_s[\vec{z} + \vec{k} - \vec{m}] &= \sum_{\vec{u}} w_s[\vec{u}] w_s[\vec{z} + \vec{u} + \vec{n} - \vec{m}] \\
&= N_s^2 \text{tri} \left[ \frac{1}{N_s} (\vec{z} + \vec{n} - \vec{m}) \right]. \tag{A.27}
\end{aligned}$$

Back substitution of (A.27) through (A.25), setting  $\vec{v} = \vec{n} - \vec{m}$  and inserting (2.12), (2.13) and (2.16) yields the autocorrelation in region  $\Omega_{c_{ns}}$ :

$$\begin{aligned}
R_{c_{ns}}[\vec{v}] &= E \{ \mathbf{c}_{\vec{n}s}[\vec{n}] \mathbf{c}_{\vec{n}s}[\vec{m}] \} \\
&= a_s^{-2} \sum_{\vec{z}} N_s^2 \text{tri} \left[ \frac{1}{N_s} (\vec{z} + \vec{v}) \right] R_{nn}[\vec{z}] R_{ss}[\vec{z} + \vec{v}] \\
&= a_s^{-1} \sigma_s^2 \sigma_n^2 \sum_{\vec{z}} \text{tri} \left[ \frac{1}{N_s} (\vec{z} + \vec{v}) \right] \text{tri}[\alpha_n \vec{z}] \text{tri}[\alpha_s (\vec{z} + \vec{v})]. \quad (\text{A.28})
\end{aligned}$$

The value of  $\vec{v}$  at which  $R_{c_{ns}}[\vec{v}]$  falls to zero (i.e. the correlation length) may be determined by inspecting fig. A.3, where the three triangle functions in the integrand of (A.28) have been graphed. Assuming that  $N_s > \frac{1}{\alpha_s}, \frac{1}{\alpha_n}$ —which is always true in our simulations—we observe in Figure. A.3a) that the integrand is non-zero in the region delimited by the broken lines when  $|v| < 1/\alpha_s + 1/\alpha_n$ . However, at  $|v| = 1/\alpha_s + 1/\alpha_n$  (Figure. A.3b), the integrand—and consequently the integral—is zero. Since the noise and signal are both isotropic, the correlation length (in samples) in both  $x$  and  $y$  directions is thus  $N_l = 1/\alpha_s + 1/\alpha_n$ . That is, the correlation length of  $\mathbf{c}[\vec{n}]$ ,  $\vec{n} \in \Omega_{c_{ns}}$ , is simply the sum of the correlation lengths of the signal and noise. The correlation area then, is simply given by

$$N_l^2 = \left( \frac{1}{\alpha_s} + \frac{1}{\alpha_n} \right)^2. \quad (\text{A.29})$$

The variance in region  $\Omega_{c_{ns}}$  may be determined by setting  $\vec{v} = \vec{0}$  in

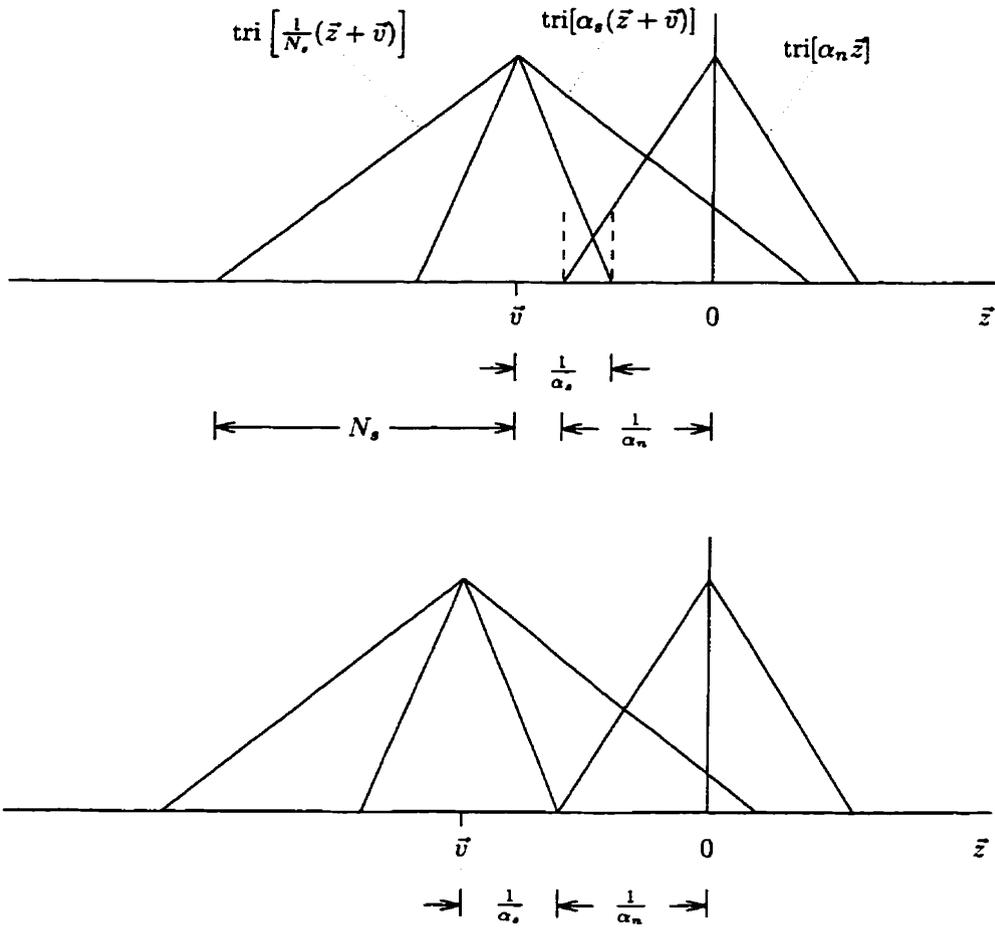


Figure A.2: Correlation falls to zero when  $|v| = 1/\alpha_s + 1/\alpha_n$ .

(A.28):

$$\begin{aligned}
\sigma_{\epsilon_{ns}}^2 &= a_s^{-1} \sigma_s^2 \sigma_n^2 \sum_{\vec{z}} \text{tri} \left[ \frac{1}{N_s} \vec{z} \right] \text{tri} [\alpha_n \vec{z}] \text{tri} [\alpha_s(\vec{z})] \\
&\approx a_s^{-1} \sigma_s^2 \sigma_n^2 \int \text{tri} \left[ \frac{1}{N_s} \vec{z} \right] \text{tri} [\alpha_s \vec{z}] \text{tri} [\alpha_n \vec{z}] d\vec{z} \\
&= \begin{cases} \frac{\sigma_n^2}{\sigma_s^2} \left[ \frac{1}{N_s \alpha_s} \left( 1 - \frac{\alpha_n}{3\alpha_s} \right) - \frac{1}{6(N_s \alpha_s)^2} \left( 2 - \frac{\alpha_n}{\alpha_s} \right) \right]^2, & \alpha_s > \alpha_n \\ \frac{\sigma_n^2}{\sigma_s^2} \left[ \frac{1}{N_s \alpha_n} \left( 1 - \frac{\alpha_s}{3\alpha_n} \right) - \frac{1}{6(N_s \alpha_n)^2} \left( 2 - \frac{\alpha_s}{\alpha_n} \right) \right]^2, & \alpha_n > \alpha_s, \end{cases}
\end{aligned} \tag{A.30}$$

where we have again assumed that  $N_s > 1/\alpha_s, 1/\alpha_n$ .

## A.4 Bandwidth of $\bar{\mathbf{n}}[\vec{n}] \bar{\mathbf{s}}[\vec{n}]$

An estimate of the bandwidth of  $\bar{\mathbf{n}}[\vec{n}] \bar{\mathbf{s}}[\vec{n}]$  may be found by inverting the width of its autocorrelation. If  $\mathbf{y}[\vec{n}] = \bar{\mathbf{n}}[\vec{n}] \bar{\mathbf{s}}[\vec{n}]$ , then the autocorrelation is given by

$$E \{ \mathbf{y}[\vec{n}] \mathbf{y}[\vec{m}] \} = E \{ \bar{\mathbf{n}}[\vec{n}] \bar{\mathbf{s}}[\vec{n}] \bar{\mathbf{n}}[\vec{m}] \bar{\mathbf{s}}[\vec{m}] \} \tag{A.31}$$

$$= E \{ \bar{\mathbf{n}}[\vec{n}] \bar{\mathbf{n}}[\vec{m}] \} E \{ \bar{\mathbf{s}}[\vec{n}] \bar{\mathbf{s}}[\vec{m}] \} \tag{A.32}$$

$$= R_{nn}[\vec{m} - \vec{n}] R_{ss}[\vec{m} - \vec{n}], \tag{A.33}$$

where we have used the fact that  $\bar{\mathbf{n}}[\vec{n}]$  and  $\bar{\mathbf{s}}[\vec{n}]$  are uncorrelated. Since  $R_{nn}[\vec{n}]$  and  $R_{ss}[\vec{n}]$  are triangular functions of widths  $2/\alpha_n$  and  $2/\alpha_s$  respectively, ((2.13) and (2.12)), the region of support of their product is equal to the width of the narrower of the two. The reciprocal of that width is the bandwidth of the higher bandwidth process and is also approximately the bandwidth of  $\bar{\mathbf{n}}[\vec{n}] \bar{\mathbf{s}}[\vec{n}]$ .

## A.5 1-D Equations

The 1-D equivalents to Equations 2.44, 2.45 and 2.46 are given by

$$\sigma_{c_0}^2 = 2 \left[ \frac{2}{3N_s\alpha_s} - \frac{1}{6(N_s\alpha_s)^2} \right] \quad (\text{A.34})$$

$$\sigma_{c_{ns}}^2 = \begin{cases} \frac{\sigma_n^2}{\sigma_s^2} \left[ \frac{1}{N_s\alpha_s} \left( 1 - \frac{\alpha_n}{3\alpha_s} \right) - \frac{1}{6(N_s\alpha_s)^2} \left( 2 - \frac{\alpha_n}{\alpha_s} \right) \right], & \alpha_s > \alpha_n \\ \frac{\sigma_n^2}{\sigma_s^2} \left[ \frac{1}{N_s\alpha_n} \left( 1 - \frac{\alpha_s}{3\alpha_n} \right) - \frac{1}{6(N_s\alpha_n)^2} \left( 2 - \frac{\alpha_s}{\alpha_n} \right) \right], & \alpha_n > \alpha_s, \end{cases} \quad (\text{A.35})$$

$$M = \frac{(N_r - N_s + 1) - N_s}{\frac{1}{\alpha_s} + \frac{1}{\alpha_n}}. \quad (\text{A.36})$$

# Bibliography

- [1] Y. S. Abu-Mostafa and D. Psaltis. Recognitive aspects of moment invariants. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 6(6):698–706, 1984.
- [2] R. S. Allison, M. Eizenman, and B. S. K. Cheung. Combined head and eye tracking system for dynamic testing of the vestibular system. *IEEE Transactions on Biomedical Engineering*, 43:1073–1082, 1996.
- [3] J. Altmann and H. J. P. Reitbock. A fast correlation method for scale- and translation-invariant pattern recognition. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 6(1):46–57, 1984.
- [4] D. I. Barnea and H. F. Silverman. A class of algorithms for fast digital image registration. *IEEE Transactions On Computers*, 21(2):179–186, 1972.
- [5] J. Ben-Arie and K. R. Rao. A novel approach for template matching by nonorthogonal image expansion. *IEEE Transactions On Circuits And Systems For Video Technology*, 3(1):71–84, 1993.
- [6] D. Casasent. Unified synthetic discriminant function computational formulation. *Applied Optics*, 23(10):1620–1627, 1984.

- [7] D. Casasent and W. T. Chang. Correlation synthetic discriminant functions. *Applied Optics*, 25(14):2343–2350, 1986.
- [8] D. Casasent and D. Psaltis. Position, rotation, and scale invariant optical correlation. *Applied Optics*, 15(7):1795–1799, 1976.
- [9] D. Casasent and D. Psaltis. Space-bandwidth product and accuracy of the optical mellin transform. *Applied Optics*, 16(6):1472, 1977.
- [10] E. De Castro and C. Morandi. Registration of translated and rotated images using finite fourier transforms. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 9(5):700–703, 1987.
- [11] Q. S. Chen, M. Defrise, and F. Deconinck. Symmetric phase-only matched filtering of fourier-mellin transforms for image registration and recognition. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 16(12):1156–1168, 1994.
- [12] F. M. Dickey and L. A. Romero. Normalized correlation for pattern recognition. *Optics Letters*, 16(15):1186–1188, 1991.
- [13] F. M. Dickey, K. T. Stalker, and J. J. Mason. Bandwidth considerations for binary phase-only filters. *Applied Optics*, 27(18):3811–3818, 1988.
- [14] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, 1973.
- [15] V. N. Dvornychenko. Bounds on (deterministic) correlation functions with application to registration. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 5(2):206–213, 1983.

- [16] M. W. Farn and J. W. Goodman. Optimal binary phase-only matched filters. *Applied Optics*, 27(21):4431–4437, 1988.
- [17] D. L. Flannery, J. S. Loomis, and M. E. Milkovich. Design elements of binary phase-only correlation filters. *Applied Optics*, 27(20):4231–4235, 1988.
- [18] M. O. Freeman and B. E. A. Saleh. Moment invariants in the space and frequency domains. *Journal of the Optical Society of America A*, 5(7):1073–1084, 1988.
- [19] P. D. Gianino and Joseph L. Horner. Additional properties of the phase-only correlation filter. *Optical Engineering*, 23(6):695–697, 1994.
- [20] A. Goshtasby, H. Gage, and J. F. Bartholic. A two-stage cross correlation approach to template matching. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 6(3):374–378, 1984.
- [21] P. R. Hinrichs. Advanced terrain correlation techniques. In *Proceedings of the IEEE Position Location, and Navigation Symposium*, pages 89–96, 1976.
- [22] J. L. Horner. Metrics for assessing pattern-recognition performance. *Applied Optics*, 31(2):165–166, 1992.
- [23] J. L. Horner and P. D. Gianino. Phase-only matched filtering. *Applied Optics*, 23(6):812–816, 1984.
- [24] J. L. Horner and J. R. Leger. Pattern recognition with binary phase-only filters. *Applied Optics*, 24(5):609–611, 1985.

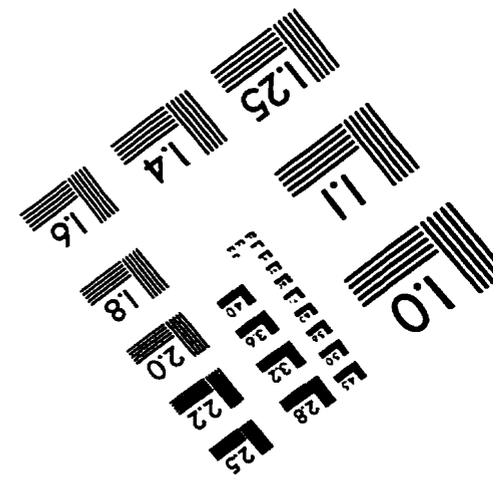
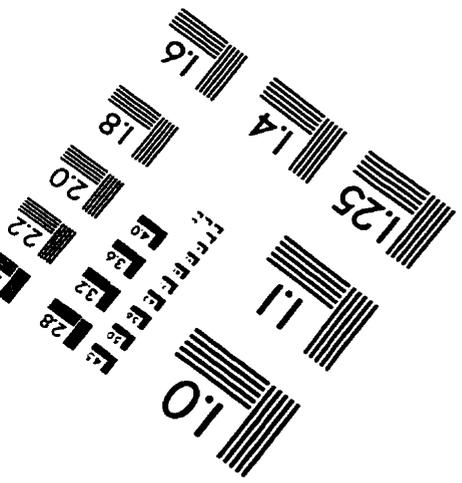
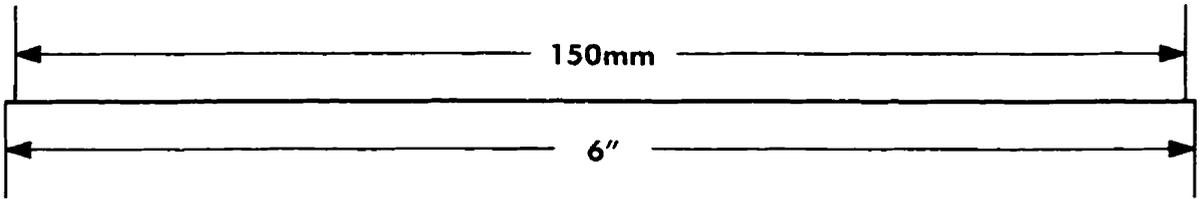
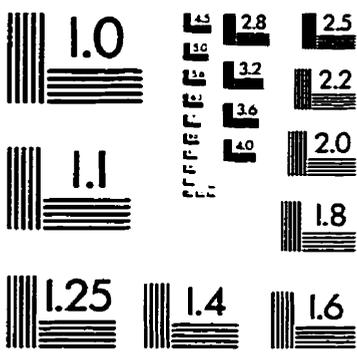
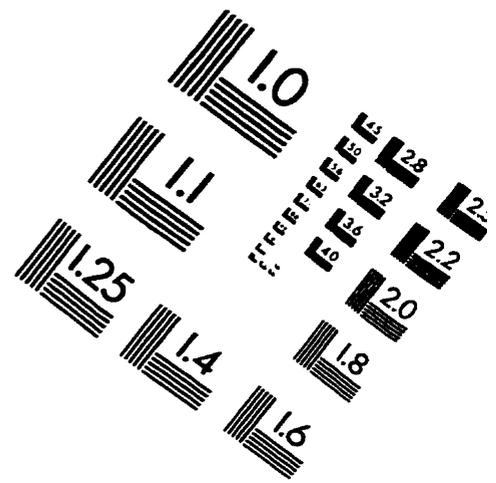
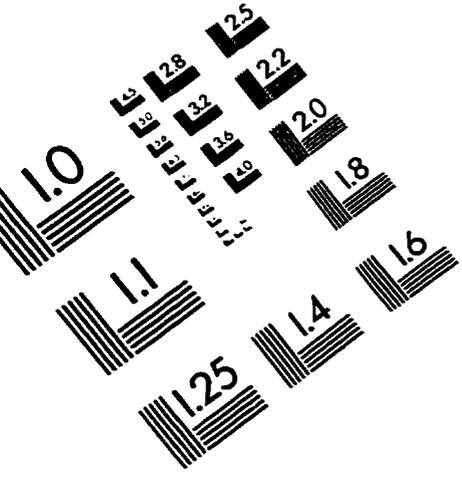
- [25] M. K. Hu. Visual pattern recognition by moment invariants. *IRE Transactions On Information Theory*, 8:179–187, 1962.
- [26] J. P. Ianiello. Time delay estimation via cross-correlation in the presence of large estimation errors. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 30(6):998–1003, 1982.
- [27] A. K. Jain. *Fundamentals Of Digital Image Processing*. Prentice–Hall, 1989.
- [28] B. Javidi and J. Wang. Limitation of the classical definition of the correlation signal-to-noise ratio in optical pattern recognition with disjoint signal and scene noise. *Applied Optics*, 31(32):6826–6829, 1992.
- [29] B. Javidi and J. Wang. Design of filters to detect a noisy target in nonoverlapping background noise. *Journal of the Optical Society of America A*, 11(10):2604–2612, 1994.
- [30] B. Javidi and J. Wang. Optimum filter for detection of a target in nonoverlapping scene noise. *Applied Optics*, 33(20):4454–4458, 1994.
- [31] B. Javidi, G. Zhang, F. Parchekani, and P. Refregier. Performance of minimum-mean-square-error filters for spatially nonoverlapping target and input-scene noise. *Applied Optics*, 33(35):8197–8209, 1994.
- [32] R. R. Kallman. Construction of low noise optical correlation filters. *Applied Optics*, 25(7):1032–1033, 1986.
- [33] D. Kersten. Predictability and redundancy of natural images. *Journal of the Optical Society of America A*, 4(12):2395–2400, 1987.

- [34] B. V. K. V. Kumar. Tutorial survey of composite filter designs for optical correlators. *Applied Optics*, 31(23):4773–4801, 1992.
- [35] B. V. K. V. Kumar and Z. Bahri. Phase-only filters with improved signal to noise ratio. *Applied Optics*, 28(2):250–257, 1989.
- [36] B. V. K. V. Kumar and C. Carroll. Loss of optimality in cross correlators. *Journal of the Optical Society of America A*, 1(4):392–397, 1984.
- [37] B. V. K. V. Kumar and L. Hassebrook. Performance measures for correlation filters. *Applied Optics*, 29(20):2997–3006, 1990.
- [38] B. V. K. V. Kumar and A. Mahalanobis. Alternate interpretation for minimum variance synthetic discriminant function. *Applied Optics*, 25(15):2484–2485, 1986.
- [39] B. V. K. V. Kumar and E. Pochapsky. Signal-to-noise ratio considerations in modified matched spatial filters. *Journal of the Optical Society of America A*, 3(6):777–786, 1986.
- [40] H. L. Kyoon, K. B. Eom, and R. L. Kashyap. Character recognition based on attribute-dependent programmed grammar. *IEEE Transactions On Pattern Analysis And Machine Intelligence*. 14(11):1122–1128, 1992.
- [41] A. B. Vander Lugt. Signal detection by complex matched spatial filtering. *IEEE Transactions on Information Theory*, 10:139–145, 1964.
- [42] A. Mahalanobis, B. V. K. V. Kumar, and D. Casasent. Minimum average correlation energy filters. *Applied Optics*, 26(17):3633–3640, 1987.

- [43] H. Mostafavi. Optimal window functions for image correlation in the presence of geometric distortion. *IEEE Transactions On Acoustics, Speech, And Signal Processing*, 27(2):163–169, 1979.
- [44] H. Mostafavi and F. W. Smith. Image correlation with geometric distortion part ii: Effect on local accuracy. *IEEE Transactions On Aerospace And Electronic Systems*, 14(3):494–500, 1978.
- [45] H. Mostafavi and F. W. Smith. Image correlation with geometric distortion part i: Acquisition performance. *IEEE Transactions On Aerospace And Electronic Systems*, 14(3):487–493, 1978.
- [46] M. Nadler and E. P. Smith. *Pattern Recognition Engineering*. John Wiley & Sons, 1993.
- [47] H. Nishida and S. Mori. Algebraic description of curve structure. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 14(5):416–533, 1992.
- [48] H. Nishida and S. Mori. An algebraic approach to automatic construction of structural models. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 15(12):1298–1311, 1993.
- [49] A. V. Oppenheim and J. S. Lim. The importance of phase in signals. *Proceedings Of The IEEE*, 69(5):529–541, 1981.
- [50] A. Papoulis. *Probability, Random Variables, and Stochastic Processes 3rd Ed*. McGraw Hill, 1991.
- [51] K. R. Rao and J. Ben-Arie. Multiple template expansion using the expansion filter. *IEEE Transactions On Circuits And Systems For Video Technology*, 4(5):490–503, 1994.

- [52] K. R. Rao and J. Ben-Arie. Nonorthogonal image expansion related to optimal template matching in complex images. *CVGIP: Graphical Models and Image Processing*, 56(2):149–160, 1994.
- [53] B. S. Reddy and B. N. Chatterji. An fft-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions On Image Processing*, 5(8):1266–1271, 1996.
- [54] Y. Sheng and H. H. Arsenault. Experiments on pattern recognition using invariant fourier-mellin descriptors. *Journal of the Optical Society of America A*, 3(6):771–776, 1986.
- [55] Y. Sheng and J. Duvernoy. Circular-fourier-radial-mellin transform descriptors for pattern recognition. *Journal of the Optical Society of America A*, 3(6):885–888, 1986.
- [56] M. R. Teague. Image analysis via the general theory of moments. *Journal of the Optical Society of America*, 70(8):920–930, 1980.
- [57] G. L. Turin. An introduction to matched filters. *IRE Transactions On Information Theory*, 311–329, 1960.
- [58] P. A. Wetzell, C. Poprik, and P. Bascom. An eye tracking system for analysis of pilots' scan paths. Technical report, United States Air Force Armstrong Laboratory, 1997.
- [59] L. R. Young and D. Sheena. Survey of eye movement recording methods. *Behavior Research Methods and Instrumentation*, 7:397–429, 1975.

# IMAGE EVALUATION TEST TARGET (QA-3)



**APPLIED IMAGE . Inc**  
 1553 East Main Street  
 Rochester, NY 14609 USA  
 Phone: 716/482-0300  
 Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved