

**QUANTITATIVE ANALYSIS OF A  
MOMENT BASED EDGE OPERATOR**

**Paul H. Eichel**

**Edward J. Delp**

**Department of Electrical and Computer Engineering  
and  
Computer, Information, and Control Engineering Program**

**The University of Michigan**

**Ann Arbor, Michigan 48109-1109**

**April 1984**

**CENTER FOR ROBOTICS AND INTEGRATED MANUFACTURING**

**Robot Systems Division**

**COLLEGE OF ENGINEERING**

**THE UNIVERSITY OF MICHIGAN**

**ANN ARBOR, MICHIGAN 48109-1109**

**THE UNIVERSITY OF MICHIGAN  
ENGINEERING LIBRARY**

engn  
UMR1017

**TABLE OF CONTENTS**

<b>1. INTRODUCTION: EDGE DETECTION IN DIGITAL IMAGES .....</b>	<b>3</b>
<b>2. DEFINITION OF THE SAMPLE - VARIANCE OPERATOR .....</b>	<b>3</b>
<b>3. IMAGE EDGE MODEL .....</b>	<b>4</b>
<b>4. PROBABILITY DENSITY FUNCTION OF THE S.V. OPERATOR .....</b>	<b>4</b>
<b>5. APPROXIMATIONS TO THE DENSITY FUNCTION .....</b>	<b>5</b>
<b>6. EDGE DETECTION .....</b>	<b>7</b>
<b>7. PREDICTING THE PRATT FIGURE OF MERIT .....</b>	<b>8</b>
<b>8. CONCLUSION .....</b>	<b>11</b>
<b>9. REFERENCES .....</b>	<b>12</b>
<b>10. APPENDIX .....</b>	<b>13</b>

**ABSTRACT<sup>2</sup>**

Many mathematical operators capable of locating intensity edges in discrete images have been proposed in the literature. We introduce a new operator, based on the sample variance of a group of pixels, which exhibits two unique properties: freedom from a predefined shape and a rigorous stochastic formulation. We demonstrate the utility of the latter property in applying detection theoretic principles to the task of edge detection and in theoretical predictions of experimental performance measures.

---

<sup>2</sup>This work was partially supported by the Center for Robotics and Integrated Manufacturing, Robot Systems Division, College of Engineering, University of Michigan and the Air Force Office of Scientific Research under Grant F49620-82-C-0089.

## 1. INTRODUCTION: EDGE DETECTION IN DIGITAL IMAGES

Detection of intensity edges, that is of local intensity discontinuities at the boundaries of objects, is a low-level process of fundamental importance in image processing. An edge operator is a mathematical operator of small spatial extent which responds in some way to these discontinuities, usually classifying every image pixel as either belonging to an edge or not.

The literature abounds with edge operators [1],[2]. Generally speaking these operators tend to fall in one of three broad classifications: 1) gradient-based; 2) template matching; and 3) parametric models. All of these operators usually employ some global or adaptive threshold on their outputs as a decision device to finally decide whether a given location is on an edge or not.

In this paper, we investigate the use of the sample variance of a local group of pixels as an edge operator. The use of the sample variance in image boundary detection is not new. It has been used for example to detect regions of similar texture, changes in which are defined to be edges[3]. It has also been used as an intensity edge operator [4]. We shall show that given an appropriate image model the Sample - Variance operator has a rather simple probability density function. The problem of edge detection can then be formulated using classical detection theory based on a composite likelihood ratio test. Along with another result which is presented in the Appendix, the optimum threshold for the Pratt Figure of Merit, a common performance measure for edge detectors, can be obtained.

## 2. DEFINITION OF THE SAMPLE - VARIANCE OPERATOR

The Sample - Variance (SV) operator will be defined as the sample variance of a group of  $N$  pixels, i.e.:

$$s^2 = \left\{ \frac{1}{N} \sum_{i=1}^N \left( X_i - \frac{1}{N} \sum_{j=1}^N X_j \right)^2 \right\} \quad (1)$$

where:  $N$  = number of pixels in operator.

$X_i$  = gray level value of the  $i^{\text{th}}$  pixel in pixel group.

We note that an important feature of this definition is that no predefined operator shape is implied for the edge operator. In fact, the operator is a function of an arbitrary group of pixels. These pixels may be chosen in any manner in the spatial dimensions of a single image, through time in a sequence of images, or

across several spectral channels (e.g. Landsat imagery) of image data.

### 3. IMAGE EDGE MODEL

The image model considered in this paper is a two-dimensional discrete parameter random field [5]. Given a probability space  $(\Omega, \mathcal{A}, P)$ , a measurable function,  $f_{\vec{r}}(\omega)$ ;  $\omega \in \Omega$ ;  $\vec{r} \in I^2$ , is defined on this space as:

$$f_{\vec{r}}(\omega) = g_{\vec{r}} + n_{\vec{r}}(\omega) \quad : \quad \Omega \times I^2 \rightarrow R \quad (2)$$

Here,  $g_{\vec{r}}$  is a deterministic discrete image consisting of two or more regions, each of which is of constant gray-level over its extent. The boundaries or edges between these regions have a ramp-like cross section as in Figure 1. The height, width, and direction of a ramp are parameters of a given edge.

To this function  $g_{\vec{r}}$  is added a zero mean i.i.d. Gaussian distributed random field  $n_{\vec{r}}(\omega)$ . For fixed  $\vec{r}$ , the random variable  $f_{\vec{r}}(\omega)$  therefore has the density:

$$p_f(\vec{r}, z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-g_{\vec{r}})^2}{2\sigma^2}} \quad (3)$$

where:  $\sigma^2 =$  noise variance

### 4. PROBABILITY DENSITY FUNCTION OF THE S.V. OPERATOR

Given the image model of Section 2, the problem then is to apply the SV operator and identify which pixels belong to an edge. We take a detection-theoretic approach to this problem by first finding the probability density function of the SV operator (i.e. find the pdf of  $S^2$  in Equation 1). We can then use one of several well-known tests based on the likelihood ratio to obtain a decision threshold.

Under the assumption of Section 2, each  $X_i$  of the SV operator (Equation 1) is a random variable with density function as in Equation 3 i.e. the  $X_i$ 's are i.i.d. Gaussian random variables with variance  $\sigma$  and means:

$$E\{X_i\} = g_{\vec{r}} \quad (4)$$

where  $\vec{r}_i$  is the position vector of the  $i^{\text{th}}$  pixel. Letting:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i, \quad Y_i = \frac{X_i - \bar{X}}{\sigma} \quad (5)$$

we can rewrite Equation 1 as:

$$\frac{NS^2}{\sigma^2} = \sum_{i=1}^N Y_i^2 \quad (6)$$

where:

$$E\{Y_i\} = \frac{1}{\sigma} (E\{X_i\} - \frac{1}{N} \sum_{j=1}^N E\{X_j\}) \quad \text{Var}(Y_i) = 1$$

Thus, the variable  $\frac{NS^2}{\sigma^2}$  of Equation 8 is the sum of squares of normal, unit variance random variables with unequal means. This is known to have a non-central chi-square distribution of  $N-1$  degrees of freedom [6]. One degree of freedom is lost due to the fact that the sample mean is subtracted from each  $X_i$  in Equation 7 instead of the actual mean. The non-centrality parameter,  $\Theta$ , is given by:

$$\Theta = \sum_{j=1}^N (E\{Y_j\})^2 \quad (7)$$

We will denote the density of  $\frac{NS^2}{\sigma^2}$  by  $X_{(\nu, \Theta)}^2$  i.e. chi-square with  $\nu$  degrees of freedom and non-centrality parameter  $\Theta$ .

## 5. APPROXIMATIONS TO THE DENSITY FUNCTION

The variable  $\frac{NS^2}{\sigma^2}$  is distributed as a non-central chi-square. Unfortunately, there is no closed form for the probability density of  $X_{(\nu, \Theta)}^2$  [7]. We therefore investigate suitable approximations for the density of  $X_{(\nu, \Theta)}^2$ .

- (1) First, we approximate  $X_{(\nu, \Theta)}^2$  by a central chi-square variable,  $\rho X_{(\lambda, \delta)}^2$ , where the constants  $\rho$  and  $\lambda$  are chosen so that  $X_{(\nu, \Theta)}^2$  and  $\rho X_{(\lambda, \delta)}^2$  have the same mean and variance [8]:

$$\rho = 1 + \frac{\Theta}{\nu + \Theta} \quad \lambda = \nu + \frac{\Theta^2}{\nu + 2\Theta}$$

So:

$$\frac{X^2_{(\nu, \Theta)}}{1 + \frac{\Theta}{\nu + \Theta}} \text{ is approximately } X^2_{\left(\nu + \frac{\Theta^2}{\nu + 2\Theta}, 0\right)} \quad (8)$$

- (2) Second, we approximate the newly formed  $X'^2$  variable as a normal variable. The standard approximation is [8]:

$$[2X'^2] \approx N((2\lambda - 1), 1) \quad (9)$$

Putting Equations (10) and (11) together, we have:

$$\left[ \frac{2NS^2}{\rho\sigma^2} \right] \approx N((2\lambda - 1), 1)$$

or:

$$S \approx N\left( \left[ (2\lambda - 1) \frac{\rho\sigma^2}{2N} \right], \frac{\rho\sigma^2}{2N} \right) \quad (10)$$

where:

$$\begin{aligned} \rho &= 1 + \frac{\Theta}{\nu + \Theta} & \Theta &= \sum_{j=1}^N (E\{Y_j\})^2 \\ \lambda &= \nu + \frac{\Theta^2}{\nu + 2\Theta} & \nu &= N - 1 \end{aligned}$$

This approximate density for  $S$  is used for all calculations of error probabilities, etc. Numerical justification of this approximation has been tabulated by Patnaik[8], Wilson[9] and others. The approximation is fairly good at  $N=10$  and converges rather rapidly as  $N$  increases.



## 6. EDGE DETECTION

We now address the original problem of identifying edges in the image. We would like to use our knowledge of the operator's density function to obtain a classification threshold based on a statistical test.

From Equation 10 we see that the SV operator has an approximately normal density function whose mean and standard deviation are a function of the noise power,  $\sigma^2$ , the non-centrality parameter,  $\Theta$ , and the number of pixels in the operator,  $N$ .  $\Theta$  in turn depends on the proximity of an edge in the operator window. Examination of Equations 4, 6 and 7 reveals that  $\Theta$  is a measure of the variance of the deterministic image elements,  $g_r$ , in the operator window. Thus, when the operator window is located at a region of the image in which there is no edge,  $\Theta$  will be small. On the other hand, when the window is centered on an edge, the parameter has a high value since both low and high element values,  $g_r$  are present. Further,  $\Theta$  varies smoothly between these two extremes as the ratio of high to low element values changes between one half and either zero or one, i.e. as the window goes from a centered position over the edge to one side or the other.

This is illustrated in Figure 2 for a 3x3 window and a ramp edge model with height to width ratio of 2. The contours in Figure 2 are loci of edges with equal parameter  $\Theta$ , normalized to that of a vertical edge through the center of the window. That is, an edge whose midline is tangent to the 0.5 contour, for example, has a  $\Theta$  value 0.5 times that of a vertical edge through the center pixel. Edges near the center of the window have a high  $\Theta$  value and the parameter falls off with increasing distance from the center. The response is nearly isotropic and very similar to the square-root Sobel (3x3) operator whose magnitude response contours are shown in Figure 3 for comparison.

The mean of the density function of the random variable  $S$  is a function of the parameter  $\Theta$ . Referring to Equation 10, we see that a large value of  $\Theta$  (edge present) gives rise to a density function with a large mean as compared to smaller values of  $\Theta$  (no edge). Thus the density function of  $S$  conditioned on the presence of an edge in the operator window has a different mean from that of  $S$  conditioned on the absence of an edge (Figure 4). Our job is to choose a decision threshold,  $S_i$ , so as to minimize errors.

The choice of  $S_i$  cannot be based on a simple likelihood ratio test because  $\Theta$  can take on a range of values. Let the  $H_0$  hypothesis be the event, denoted  $e_0$ , that no edge is present in the window and  $H_1$  be the event,  $e_1$ , that an edge is present. Under  $H_1$  then,  $\Theta$  may lie in some range  $\theta \in [a,b]$  where  $a$  is the maximum value of  $\Theta$  (vertical centered edge) and  $b$  is appropriately chosen from Figure 2 (e.g.  $b = 0.5 * a$ ). Under  $H_0$ ,  $\theta \in (b,0]$ . In other words the  $H_1$  hypothesis is that an edge falls inside the  $b/a$  contour of Figure 2, the  $H_0$

hypothesis is that the edge falls outside that contour. A composite likelihood ratio test [10] on the variable  $S$  for deciding whether or not an edge is present is then:

$$\Lambda(s) = \frac{\int_{\theta_0 \in \Omega_0} p(s|e_0, \theta_0) p(\theta_0) d\theta_0}{\int_{\theta_1 \in \Omega_1} p(s|e_1, \theta_1) p(\theta_1) d\theta_1} \begin{matrix} d_0 \\ > \\ < \\ d_1 \end{matrix} T \quad (11)$$

where:

$$\Omega_1 = [a, b] \quad \Omega_0 = (b, 0]$$

$$d_1 = \text{decide } e_1 \text{ is true} \quad d_0 = \text{decide } e_0 \text{ is true}$$

In the case  $p(\theta)$  is not known a priori, one simplifying assumption is that edges are uniformly distributed i.e. no particular edge location or orientation is any more likely than any other. One may then calculate  $p(\theta)$  for this case from Figure 2 by means of differential area. That is,  $p(\theta)$  is proportional to the area bounded by the contours corresponding to  $\theta$  and  $\theta + d\theta$ . Equation 11 can then be evaluated numerically to obtain a decision threshold,  $S_t$ , for  $s$ . If  $T = 1$ , the resulting test is maximum likelihood. Other possibilities include MAP ( $T = \frac{P(\text{no-edge})}{P(\text{edge})}$ ), Neyman-Pearson, Bayes Risk, and Minimax tests. The probability of error can also be calculated. An example of this procedure will be given in the next section.

## 7. PREDICTING THE PRATT FIGURE OF MERIT

The Pratt Figure of Merit [11] for edge detectors has been used in the literature as a means for making performance comparisons among edge detectors. It is an experimental tool to quantify the performance of edge detectors in an effort to make such comparisons less subjective. It penalizes for both Type I and Type II errors as well as for displacements of detected edge points from true location. In use, an artificial image is constructed with a single vertical ramp edge of height  $h$ , width two pixels, and with added iid noise of variance  $\sigma^2$ . The edge detector's output from this image is then used to form the figure of merit defined by:

$$F = \frac{1}{I_M} \sum_{i=1}^{I_D} \frac{1}{1 + \alpha l_i^2} \quad (12)$$

where:

- $I_M = \max (I_D, I_I)$
- $I_I =$  number of ideal edge points
- $I_D =$  number of detected edge points
- $l_i =$  displacement of the  $i^{\text{th}}$  detected edge point from the ideal edge.
- $\alpha =$  scaling parameter

In the past, the merit value for an edge detector was found by experimentally changing the magnitude threshold for the detector until the best merit value was obtained. This experimental optimization had to be performed with a new artificial image for each value of signal-to-noise ratio of interest. However, since the SV Edge Operator has a known probability density, we can establish a priori what the Pratt Figure of Merit is as a function of signal-to-noise ratio, and what the optimum threshold on  $S$  should be in order to attain that merit value. The ability to do this is crucially dependent on the following Theorem which is proved in the Appendix.

Theorem: When  $p(s|e_1)$  and  $p(s|e_0)$  are normally distributed, then the threshold for  $S$  that maximizes  $F$  is the threshold  $S_i$  such that:

$$P(d_0, e_1) = P(d_1, e_0) \quad (13)$$

The corresponding optimum figure of merit is :

$$F \approx \frac{P(d_1, e_1)}{P(d_1)} \quad (14)$$

We will use this Theorem along with the material developed in the previous section to predict the Pratt Figure of Merit for a 3x3 SV operator. We first expand the joint probabilities of Equation 13:

$$P(d_0, e_1) = P(e_1) \int_{S_i}^{\infty} \int_{t_1 \in \Omega_1} p(s|e_1, \theta_1) p(\theta_1) d\theta_1 ds$$

$$P(d_1, e_0) = P(e_0) \int_{-\infty}^{S_t} \int_{\theta_0 \in \Omega_0} p(s|e_0, \theta_0) p(\theta_0) d\theta_0 ds \quad (15)$$

Now since the edge is vertical of width two pixels and the operator is a 3x3 square, only three values of  $\Theta$  are possible. The first corresponds to the window centered on the edge:

$$\theta_1 \in \Omega_1 = \{(3h/2\sigma)^2\}$$

The other two possibilities of  $\Theta$  correspond to: 1) the window being centered one pixel to either side of the edge midline; and 2) the edge is completely outside the window:

$$\theta_0 \in \Omega_0 = \{(h/2\sigma)^2, 0\}$$

where  $\frac{h^2}{\sigma^2}$  is defined to be the signal to noise ratio. Furthermore, if the test image is square of size  $M \times M$ , these three values of  $\Theta$  have a priori probabilities in the ratios 1:2:M-2. Consequently, Equations 15 become:

$$\begin{aligned} P(d_0, e_1) &= \frac{1}{M} \int_{S_t}^{\infty} p(s|e_1, \theta_1=(3h/2\sigma)^2) ds \\ &= \frac{1}{M} Q\left[\frac{S_t - \mu_1}{\sigma_1}\right] \end{aligned}$$

$$\begin{aligned} P(d_1, e_0) &= \frac{2}{M} \int_{-\infty}^{S_t} p(s|e_0, \theta_0=(h/2\sigma)^2) ds + \frac{M-2}{M} \int_{-\infty}^{S_t} p(s|e_0, \theta_0=0) ds \\ &= \frac{2}{M} Q\left[\frac{\mu_{2a} - S_t}{\sigma_{2a}}\right] + \frac{M-2}{M} Q\left[\frac{\mu_{2b} - S_t}{\sigma_{2b}}\right] \end{aligned} \quad (16)$$

where  $\mu_1, \sigma_1$  are calculated from Equation 10 for  $\theta_1=(3h/2\sigma)^2$ , etc. Substituting Equation 16 into Equation 13, we have:

$$Q\left[\frac{S_t - \mu_1}{\sigma_1}\right] = 2 Q\left[\frac{\mu_{2a} - S_t}{\sigma_{2a}}\right] + (M-2) Q\left[\frac{\mu_{2b} - S_t}{\sigma_{2b}}\right] \quad (17)$$

which may be solved numerically for the threshold,  $S_t$ . Equation 14 is then used to find the resulting Figure of Merit.

Figures 5 and 6 illustrate these results. They compare the theoretical figure of merit and optimum threshold verses snr (solid lines) to the experimental data (stars). The experimental points were obtained by iteratively changing the threshold until the highest figure of merit was obtained, a procedure universal to the literature. The accuracy of the probability model for the SV operator and the above Theorem clearly allows us to predict the outcome with assurance.

## 8. CONCLUSION

In this paper we have shown that the Sample - Variance operator is useful for detecting intensity edges in discrete images. Under the hypothesis of additive, Gaussian distributed i.i.d. noise, the operator's probability density function may be very accurately approximated by a normal density. This allows the use of a composite likelihood ratio test to determine a decision threshold for the operator. In this way we can implement a variety of classical detection schemes such as Maximum Likelihood and MAP.

We have also proved a theorem concerning predicting the optimum decision threshold for the Pratt Figure of Merit. The predicted optimum threshold for a 3x3 SV operator was found to be virtually identical to the actual value obtained by the usual experimental method. This demonstrates the power of the detection theory approach to finding edges in images.

**9. REFERENCES**

- [1] Kunt, M., "Edge detection: a tutorial review," *Proc. of ICASSP 82. IEEE Inter. Conf. on Acou. Speech & Sig. Proc.* vol.2, 1982.
- [2] Davis, L., "A survey of edge detection techniques," *Computer Graphics and Image Processing.* vol.4, 1975, pp248-270.
- [3] Weszka, J., C. Dyer, and A. Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Trans. on Sys. Man & Cyber.* vol.6, no.4, April 1976, pp269-285.
- [4] Hale, J., "Detection of elementary features in a picture by non-linear local numerical processing," *Proc. Third Int. Joint Conf. on Pattern Recognition.* 1976, pp764-768.
- [5] Wong, E., *Stochastic Processes in Information and Dynamic Systems* McGraw Hill, 1971.
- [6] Hogg, R. and A. Craig, *Introduction to Mathematical Statistics* Macmillan Co., 1971, pp.316-317.
- [7] Tiku, M., "Laguerre series forms of noncentral chi-square and F distributions," *Biometrika* vol.52, 1965, p415.
- [8] Patnaik, P., "The noncentral chi-square and F distributions and their applications," *Biometrika* vol.36, 1949, p202.
- [9] Wilson, E. and M. Hilferty, "The distribution of chi-square," *Proc. Nat. Acad. Sci.* vol.17, no.694, 1931.
- [10] Melsa, J. and D. Cohen, *Decision and Estimation Theory* McGraw-Hill, 1978, pp.21-53.
- [11] Pratt, W. and I. Abdou, "Quantitative design and evaluation of enhancement/thresholding edge detectors," *Proc. of the IEEE* vo.67, no.5, May 1979, pp.753-763.

## 10. APPENDIX

Theorem: When  $p(s|e_0)$  and  $p(s|e_1)$  are normally distributed, then the threshold for  $S$  that maximizes  $F$  is the threshold  $S_t$  such that:

$$P(d_0, e_1) = P(d_1, e_0)$$

This Theorem will be proved in several parts, and only for  $F > 0.5$ . For  $F$  less than this, only a weaker statement can be made.

( i ) Lemma (1):

If  $S_t$  is chosen such that  $P(d_0, e_1) = P(d_1, e_0)$ ,

Then:  $I_M = I_I = I_D$ . (See Equation 12)

proof:

Let  $I_T$  = total number of image points. Then:

$$I_I = I_T P(e_1) = I_T \{ P(d_0, e_1) + P(d_1, e_1) \} \quad (18)$$

$$= I_T \{ P(d_1, e_0) + P(d_1, e_1) \}$$

$$= I_T P(d_1) = I_D$$

$$\therefore I_M = \max(I_I, I_D) = I_I = I_D$$

Let  $F_T \triangleq F$  resulting from  $S_t = T$  chosen as above.

Let  $F_t \triangleq F$  resulting from threshold  $S_t = t$

( ii ) Lemma (2):  $F_t \leq F_T$  ;  $t \geq T$

proof:

$I_D = I_T P(d_1)$  and  $P(d_1)$  is a monotonically non-increasing function of  $t$ .

Furthermore:  $t = T \implies I_D = I_I$  (Lemma (1))

$$\therefore I_D \leq I_I \quad ; \quad t \geq T$$

$$\therefore I_M = I_I \quad ; \quad t \geq T$$

Thus:

$$F_t = \frac{1}{I_t} \sum_{i=1}^{I_D} \frac{1}{1 + \alpha l_i^2} \quad ; t \geq T$$

is a monotonically non-increasing function of  $t$ , since  $I_t$  is a constant. The lemma follows.

(iii) Lemma (3):  $F_t \leq F_T$  ;  $t \leq T$  if  $F_T \geq 0.5$

First, we show that:

$$F_t \approx \frac{P(d_1, e_1)}{P(d_1)} \quad ; t \leq T \quad (19)$$

proof:  $I_M = I_D$  ;  $t \leq T$  (see above). Therefore:

$$F_t = \frac{1}{I_D} \sum_{i=1}^{I_D} \frac{1}{1 + \alpha l_i^2} \quad ; t \leq T$$

Now, if a detected edge point is a real edge point,  $l_i = 0$ .  $\therefore \frac{1}{1 + \alpha l_i^2} =$

1. The detected edge points that are not real edge points are randomly scattered throughout the image (iid noise). So  $l_i$  is large on the average. Therefore:

$$\begin{aligned} \sum_{\text{false edge points}} \frac{1}{1 + \alpha l_i^2} &<< \sum_{\text{real edge points}} \frac{1}{1 + \alpha l_i^2} \\ &= I_T P(d_1, e_1) \end{aligned}$$

Therefore:

$$\begin{aligned} F_t &= \frac{1}{I_D} \left\{ \sum_{\text{real edge points}} \frac{1}{1 + \alpha l_i^2} + \sum_{\text{false edge points}} \frac{1}{1 + \alpha l_i^2} \right\} \\ &\approx \frac{1}{I_T \{P(d_1)\}} \left\{ I_T P(d_1, e_1) \right\} \\ &\approx \frac{P(d_1, e_1)}{P(d_1)} \quad ; t \leq T \end{aligned}$$

Next we prove the lemma by showing that:



$$\frac{dF_t}{dt} > 0 \quad \text{for } t \leq T \quad \text{if } F_T \geq 0.5 \quad (20)$$

proof: Using Equation 19:

$$\begin{aligned} \frac{dF_t}{dt} &= \frac{d}{dt} \left\{ \frac{P(d_1, e_1)}{P(d_1)} \right\} \\ &= \frac{1}{[P(d_1)]^2} \left\{ P'(d_1, e_1)P(d_1) - P'(d_1)P(d_1, e_1) \right\} \end{aligned}$$

So:

$$\frac{dF_t}{dt} > 0 \quad \iff \quad \frac{P'(d_1, e_1)}{P'(d_1)} < \frac{P(d_1, e_1)}{P(d_1)} \quad (21)$$

$$\text{since } P'(d_1) = \frac{d}{dt} P(d_1) < 0$$

$$\begin{aligned} P'(d_1, e_1) &= \frac{d}{dt} \left\{ P(d_1, e_1) \right\} \\ &= \frac{d}{dt} \int_t^{\infty} p(s, e_1) ds = p(s, e_1) \end{aligned}$$

And similarly:  $P'(d_1) = p(s, e_1) + p(s, e_0)$

$$\begin{aligned} \therefore \frac{dF_t}{dt} > 0 &\iff \frac{p(s, e_1)}{p(s, e_1) + p(s, e_0)} < \frac{P(d_1, e_1)}{P(d_1, e_1) + P(d_1, e_0)} \\ &\iff \frac{p(s, e_0)}{p(s, e_1)} > \frac{P(d_1, e_0)}{P(d_1, e_1)} \end{aligned} \quad (22)$$

Now, from Equation 19:

$$F_t = \frac{P(d_1, e_1)}{P(d_1, e_1) + P(d_1, e_0)}$$

$$\therefore \frac{P(d_1, e_0)}{P(d_1, e_1)} = \frac{1}{F_t} - 1$$

$$\Rightarrow \frac{P(d_1, e_0)}{P(d_1, e_1)} < 1 \quad \text{whenever } F_t > 0.5 \quad (23)$$

Examining Equation 22 in light of Equation 23, we see that Equation 20 and the Lemma will be proved if we can but show:

$$\frac{p(s, e_0)}{p(s, e_1)} \geq 1 \quad ; t \leq T \quad (24)$$

But by definition,  $T$  is such that :

$$\begin{aligned} P(d_0, e_1) &= P(d_1, e_0) \\ \text{i.e. } \int_{-\infty}^T p(s, e_1) ds &= \int_T^{\infty} p(s, e_0) ds \\ \text{or. } P(e_1) \int_{-\infty}^T p(s|e_1) ds &= P(e_0) \int_T^{\infty} p(s|e_0) ds \\ \text{or. } P(e_1) Q\left(\frac{\mu_1 - T}{\sigma_1}\right) &= P(e_0) Q\left(\frac{T - \mu_0}{\sigma_0}\right) \end{aligned} \quad (25)$$

Now since  $\sigma_1 > \sigma_0$  always for this operator (see Section 5 and Equation 10), Equation 25 therefore implies:

$$\Rightarrow P(e_1) p(s=T|e_1) < P(e_0) p(s=T|e_0) \quad (26)$$

$$\begin{aligned} \Rightarrow 1 &< \frac{P(e_0) p(s=T|e_0)}{P(e_1) p(s=T|e_1)} \\ \Rightarrow 1 &< \frac{p(s=T, e_0)}{p(s=T, e_1)} \\ \Rightarrow 1 &< \frac{p(s, e_0)}{p(s, e_1)} \quad ; t \leq T \end{aligned} \quad (27)$$

This proves Lemma 3. Lemmas 2 and 3 together prove the Theorem when  $F_T > 0.5$ . Equation 19 gives the resulting Figure of Merit. If  $F_{\max} \leq 0.5$ , only Lemma 2 holds; i.e. the optimum threshold occurs at  $T$  or less.



Figure 1: A ramp edge in cross section. Here, the height is  $h$  and the width is two pixels.

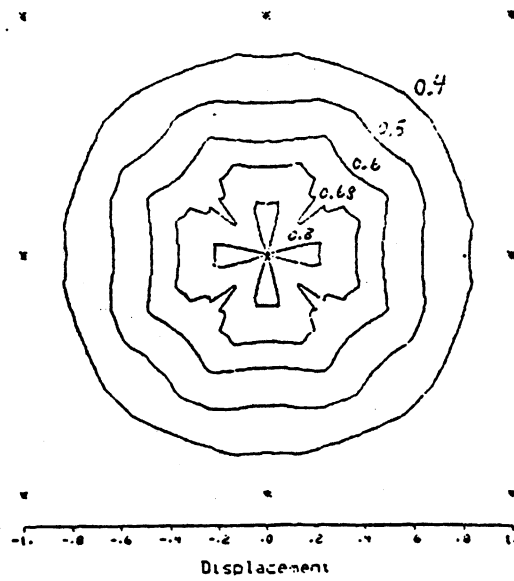


Figure 2: Contours of equal parameter  $\Theta$  for a  $3 \times 3$  SV operator. The displacement is in pixels.

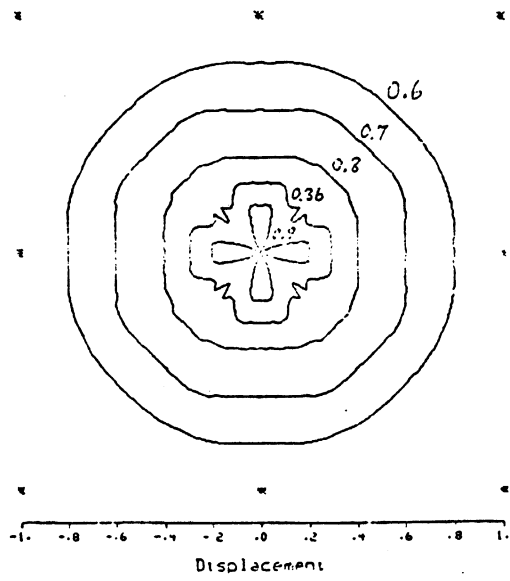


Figure 3: Contours of equal magnitude for the 3x3 Sobel operator.

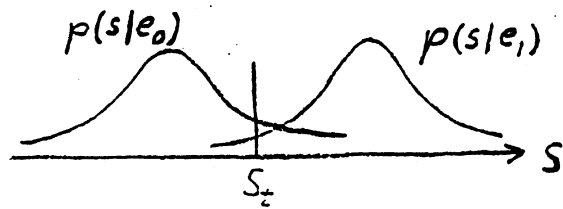


Figure 4: The density function of  $S$  conditioned on the presence or absence of an edge in the operator window.

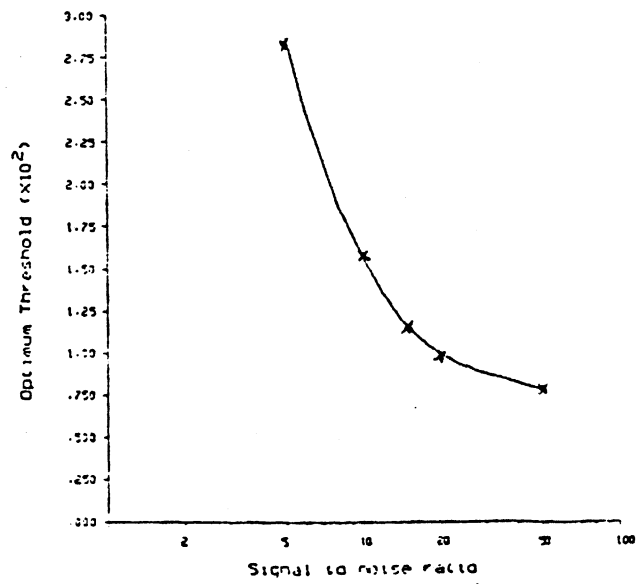


Figure 5: Predicted (curve) verses actual (stars) optimum threshold for the Pratt Figure of Merit.

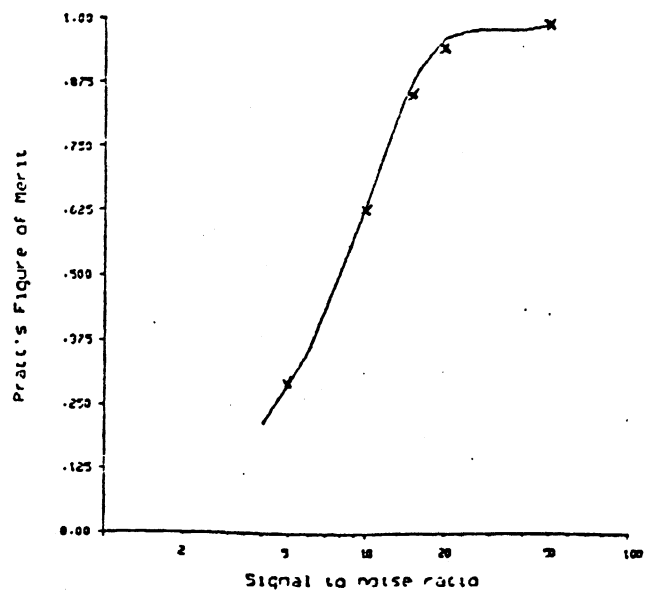


Figure 6: Predicted (curve) verses actual (stars) Pratt Figure of Merit.

UNIVERSITY OF MICHIGAN



3 9015 02653 6069