Quantitative Evaluation of Edge Detectors Using the Minimum Kernel Variance Criterion

Qiang Ji

Department of Computer Science
University of Nevada

Robert M. Haralick

Department of Electrical Engineering
University of Washington

Abstract

In this paper, we introduce a new criterion for analytically evaluating different edge detectors (both gradient and zero-crossing based methods) without the need of ground-truth information. Our criterion is based on our observation that most edge detectors make a decision of whether a pixel is an edgel or not based on the result of convolution of the image with a kernel. The variance of the convolution output therefore directly affects the performance of an edge detector. We show how to compute the variance of a convolution. We then describe results from comparing four well-known edge detectors using the proposed criterion.

1 Introduction

Despite the enormous amount of literature on edge detectors, there are only a few papers on evaluating and/or comparing the performance of different edge detectors. Such a study is important since it helps us understand the strengths and the weakness of an edge operator as well as its applicability to a particular application. Abdou and Pratt [1] proposed Pratt's figure of Merit criterion for analytically evaluating the edge detectors for synthetic images or real images with ground-truth data. Kitchen and Rosenfeld [4] proposed an edge detection evaluation criterion based on edge coherence and without requiring knowledge of the ideal edge position. Ramesh and Haralick [6] proposed a method for evaluating a facet-based edge detector with known perturbation model for the input image, based on which probabilities of mis-detection and false alarm rates are computed analytically. Wang and Binford [8] analytically evaluated the performance of a step-edge detection method by fitting a surface to the gradient magnitude values. Heath et al [5] recently proposed an empirical method for evaluating the edge detectors for real images based on subjective evaluation of edge images by people without the need of groundtruth information.

In this paper, we describe a new criterion for analytically evaluating different edge detectors (both gradient and zero-crossing based methods) without the need of ground-truth information. The criterion, called *kernel-variance*, evaluates each edge detector based on the variance of its output quantity.

For comparison, we will evaluate the performance of two gradient operators: the Canny edge detector and the integrated gradient operator; and two zero-crossing edge detectors: the facet zero-crossing edge detector (hereafter referred to as the Haralick edge detector) and the Laplacian of Gaussian (LOG) using both synthetic and real images.

2 Performance Evaluation

Most edge detectors, be the gradient-based methods or zero-crossing approaches, require convolving an image with a kernel to compute gradients or zero-crossings. A decision is then made as to whether a pixel is an edgel or not based on the result of the convolution. The performance of an edge detector therefore largely depends on the result of the convolution, which is determined by the kernel. The variance of the convolution output therefore directly affects the performance of the edge detector. A larger variance with the convolution result usually leads to a higher mis-detection and false alarm rate. Based on the above analysis, we adopt the kernel-variance criterion for comparing different edge detectors. An optimal edge detector has a convolution kernel that yields the minimum variance on its output given the same input perturbation. Using this criterion, we studied the performance of four edge detectors. The results of this study are summarized in sections 2.1 and 2.2 respectively.

2.1 Performance comparison of gradient edge detectors

This section discusses the results from a quantitative performance analysis and characterization of the Canny gradient edge detector [2] and the integrated gradient edge detector [9] using the minimum-variance criterion. The convolution kernels for each edge detector are computed using the theories described in [2] and [9] respectively.

Given a $2M \times 2N$ kernel, let y be its output and $w(r,c)$ be the entries of the kernel, then its response to the image $I(r,c)$ is

$$y = \sum_{r=-M}^{M} \sum_{c=-N}^{N} w(r,c) I(r,c) = W'I$$

where W is a $4MN \times 1$ vector whose elements are $w(r,c)$ and I is a vector containing the corresponding intensity values $I(r,c)$. σ_y^2, the variance of y, is

$$\begin{aligned} \sigma_y^2 &= E(W'II'W) \\ &= W'\Sigma_I W \end{aligned}$$

where Σ_I is the covariance matrix of vector I. If elements of I are contaminated by an independently and identically Gaussian distributed noise with zero mean and a standard deviation of σ, then we have $\sigma_y^2 = \sigma^2 W'W$. Figure 1 plots the output variances of the Canny gradient kernel and the integrated gradient kernel versus kernel size. For the Canny kernel, kernel size is related to the smoothing factor s. Increasing kernel size requires increasing s accordingly to avoid a truncated kernel.

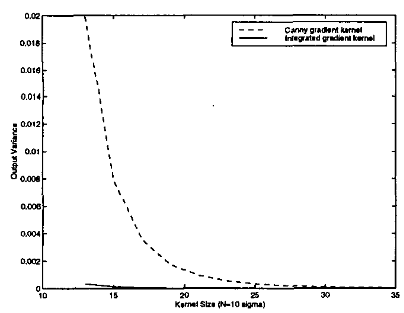


Figure 1: Kernel output variance versus kernel size for iid input perturbation

In reality, the iid perturbation assumption may not hold. Perturbations on each pixel may be correlated, especially for neighboring pixels. To account for the correlation between pixel perturbations, we assume that the iid perturbed image is subsequently smoothed by a Gaussian kernel. The Gaussian kernel introduces correlations to image pixel perturbation. Let the Gaussian kernel be g, the gradient kernel be h, and the iid perturbed image be I, then the process of introducing correlation and subsequent convolution with a gradient kernel can be described as

$$y = h * (I * g)$$

where $*$ represents convolution and y is output. Since convolution is communicative, we have

$$y = (h * g) * I \tag{1}$$

Equation 1 suggests that convolving h with an non-independently perturbed image is the same as convolving the gradient kernel h first with g and then convolving an independently contaminated image with the resulting kernel $h' = h * g$. To study the Canny gradient and integrated gradient operators on non-iid perturbed images, we can convolve each gradient kernel with the Gaussian g and then study the output variance of the resulting kernels. Figure 2 shows the performance of the two kernels for an image smoothed by Gaussian kernels of sizes 3 and 5 respectively.

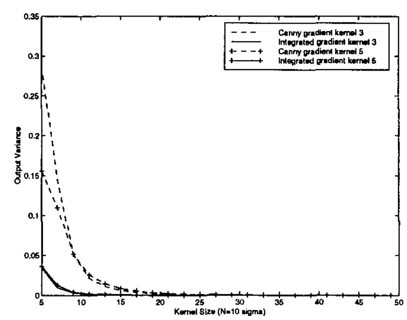


Figure 2: Kernel output variance versus kernel size for correlated input perturbation

Both figures 1 and 2 show that as kernel size increases, the output variance reduces, i.e., larger kernel window yields better estimate. This agrees with our intuition. However, larger window size introduces more locational errors and requires more computation. In reality, a balance must be maintained between the estimate precision, the locational errors, and the computational complexity. From the minimum-variance point of view, both figures 1 and 2 show the integrated gradient operator is superior to the Canny optimal gradient operator. This result agrees with the conclusions drawn by Zuniga and Haralick [9].

To validate this conclusion, we perform further performance evaluation of the two edge detectors using both synthetic and real images. Figure 3 shows the synthetic test image used. Downloaded from the Internet, the SUSAN image [7] is selected because it contains different types of edges such as step edge, roof edge, and ramp edge.

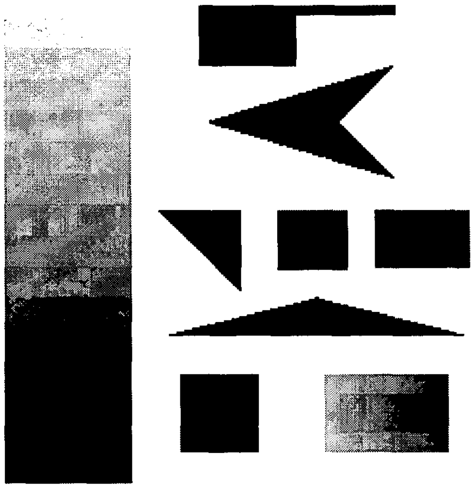


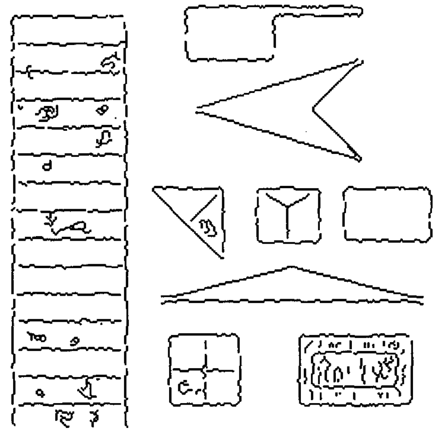
Figure 3: The synthetic test image

The test performed is to evaluate the robustness of the edge detectors under different noise levels. The input test image was perturbed with independently and identically a Gaussian distributed noise with zero mean and a standard deviation of σ. Amount of perturbation is controlled by varying σ. For the integrated gradient operator and Canny edge detector, the low and high thresholds for hysteresis linker are fixed at 1 and 3. The smoothing factor for Canny is set at 0.8. All other parameters are optimally tuned. Figures 4 (a) and (b) show edge detection results for the two edge detectors when noise level is at 5. We also applied the two edge detectors to real images. Figures 5 and 6 give the sample outputs from a real image. We can conclude from both the synthetic and real images that Canny tends to generate more false edges but fewer missing edges than the integrated gradient method. This echos the conclusion drawn from the minimum-variance analysis.

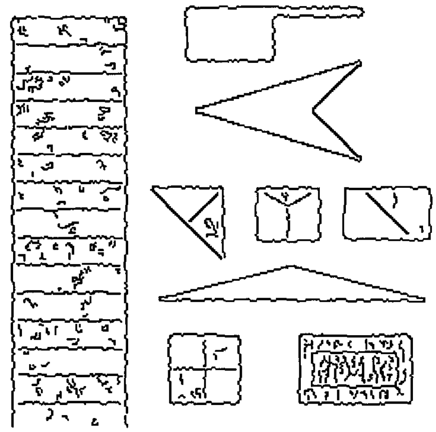
2.2 Haralick facet zero-crossing operator versus Laplacian zero-crossing operator

In this section we study the performance difference between the LOG zero-crossing operator and Haralick's facet zero-crossing operator [3] in terms of the minimum variance criterion we established in the previous section. The Laplacian of each pixel can be approximated using a LOG kernel or can be obtained analytically from the fitted facet coefficients.

Figures 7 and 8 plot the output variances of the LOG kernel and the facet Laplacian kernel versus the kernel sizes, with iid input perturbation and correlated input perturbation respectively. It is clear from the two figures that for iid input perturbation, the Laplacian of a pixel computed using facet parameters has a much lower variance than that obtained using LOG kernel, especially when kernel size is small. However, for real images (image with correlated pixel perturbations), the two techniques generate comparable variance. They also show that if LOG must be used, do not use LOGs with kernel size less than 11 since they may yield very unreliable results. For kernel sizes larger than 25 pixels, the two method yield very comparable results.



a



b

Figure 4: Output of the integrated edge detector (a) and Canny edge detector (b) when noise level is 5

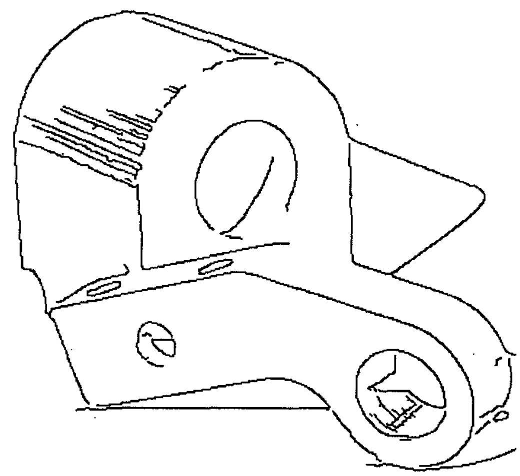


Figure 5: Output of Canny edge detector

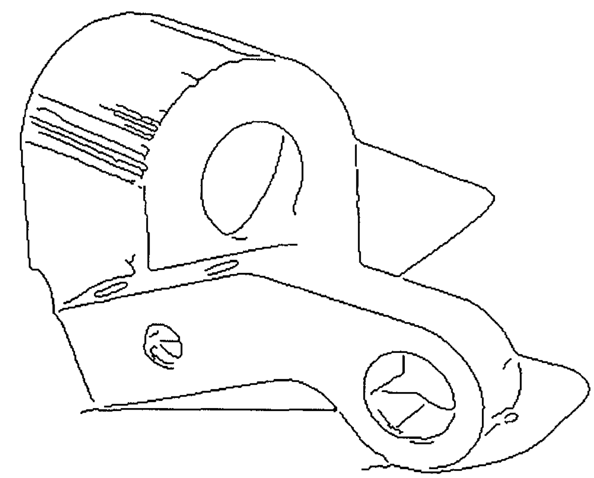


Figure 6: Output of the integrated gradient edge detector

2.3 Edge detectors performance comparison via ROC analysis

To further study the performance of the edge detectors under different parameters settings, we performed an Receiver Operating Characteristics (ROC) analysis of different edge detectors using a synthetic image. The analysis yields the ROC curves of each edge detector. The area under the ROC curve can be used as an index for measuring the performance of the edge detectors. The smaller the area under the ROC curve is, the better performance of the detector is. From figure 9, we can see integrated gradient detector yields the best performance, followed by the Canny and Facet edge detector. This basically echos the conclusions we drew from kernel variance analysis.

3 Summary

We introduced a quantitative measure based on the variance of the edge detector's output to evaluate performance of edge detectors. The proposed criterion is simple and effective without the need of groundtruth information. Furthermore, the underlying theory is also important for the design of a new edge detector or even other feature detectors.

The comparative study based on this measure reveals that the integrated gradient operator coupled with Canny's hysteresis linking procedure can yield better edge detection result than the Canny edge detector. The study also shows the superiority of facet Laplacian kernel to the LOG kernel in terms of the variance with the computed zero-crossings.

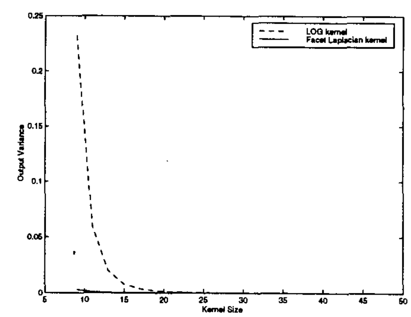


Figure 7: Kernel output variance versus kernel size for LOG and facet Laplacian kernel with iid perturbation

References

[1] K. E. Abdou and W. K. Pratt. Quantitative design and evaluation of enhancement/thresholding edge detectors. *Proc. of IEEE*, 67(5):753–763, 1979.

[2] John Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6), 1986.

[3] R. M. Haralick. Edge and region analysis for digital image data. *Computer Graphics and Image Processing*, 12:60, 1980.

[4] L. Kithcen and A. Rosenfeld. Edge evaluation using local edge coherence. *IEEE Trans. Systems, Man, and Cybernetics*, 1:597–605, 1981.

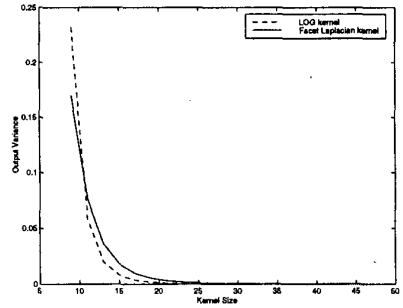


Figure 8: Kernel output variance versus kernel size for LOG and facet Laplacian kernel with correlated input perturbation

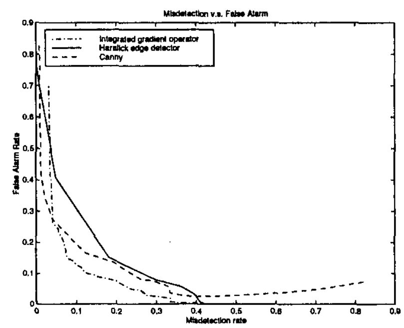


Figure 9: The ROC curves: false alarm versus misdetection

[5] K. W. Boywer M. Health, S. Sarkar. T. Sanocki. A robust visual method for assessing the relative performance of edge-detection algorithms. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(12):1063–1170, 1997.

[6] V. Ramesh and R. M. Haralick. Random perturbation model and performance characterization in computer vision. *Proc. of IEEE conf. computer vision and pattern recognition*, pages 521–527, 1992.

[7] S. M. Smith. Susan-a new approach to low level image processing. *DRA Technical Report TR95SMS1b,Dept. of Engineering Science, Oxford University, Oxford, UK*, 1995.

[8] S. Wang and T. Binford. Local step edge estimation: a new algorithm, statistical model, and performance evaluation. *Proc. of 1993 ARPA IUW*, pages 1063–1170, 1993.

[9] O. A. Zuniga and R. M. Haralick. Integrated directional derivative gradient operator. *IEEE Tran. on Systems, Man, and Cybernetics*, SMC-17(3):508–517, 1987.