Context Dependent Edge Detection

Robert M. Haralick, James S. J. Lee*

EE Dept, FT-10, University of Washington, Seattle, WA 98195, Tel:(206)545-4974

*Boeing High Tech Center, P.O.Box 24969, MS 7J-24, Seattle, WA 98124-6269, Tel:(206)865-3075

ABSTRACT

To simulate the edge perception ability of human eyes and detect scene edges from an image, context information must be employed in the edge detection process. To accomplish the optimal use of context, we introduce an edge detection scheme which uses the context of the whole image. The edge context for each pixel is the set of all row monotonically increasing paths through the pixel. The edge detector assigns a pixel that edge state having highest edge probabilty among all the paths.

Experiments indicate the validity of the edge detector. Upon comparing the performance of the context dependent edge detector with the context free second directional derivative zero-crossing edge operator, we find that the context dependent edge detector is superior.

I. INTRODUCTION

Edges in a scene are the consequence of changes in some physical and surface properties, such as illumination (shadows, for example), geometry (orientation or depth) and reflectance. As there is a direct relationship between the edges and physical properties of a scene, much of the scene information can be recovered from an edge image. Thus, edge detection plays a key step in the early processing of a computer vision system.

Image edges occur in places of significant intensity changes on the image. There are many kinds of intensity changes in an image. The usual aim of edge detection is to locate edges belonging to boundaries of objects of interest. While the human eyes perform this task easily, the detection of edges is a complex task to achieve. The difficulties in edge detection are mostly caused by noise, blurring and quantizing effects. This results in a situation in which not all the image edges correspond to scene edges and vice versa. People incorporating world knowledge and contextual information can detect edges selectively. Minor image edges which do not correspond to main scene edges are ignored and major scene edges are detected even though they do not correspond perfectly to image edges.

A variety of edge detection schemes (Marr et al. 1980) have been proposed in the past decade. Most of these operators perform reasonably well on simple noise free images whereas they tend to fail on the images degraded by noise. This is because that as the noise of an image increases, the correspondence between image edges and scene edges becomes weaker and weaker. Thus, an edge detector which can perform well on noisy images is most desirable.

The solution to edge detection on noisy images should not be image smoothing, because image smoothing alone tends to blur edges. The best solution, we believe, is to incorporate world knowledge and edge context information into the edge detection process.

The context approach described here is related to the dynamic programming idea of Montanari (1971) and Martelli (1976) of linking together edge segments. However, the dynamic programming, as they employed it, is basically a postprocessing process whose performance heavily relies on the starting points for linking which are provided by the preprocessing. A context dependent edge detection using relaxation labeling was described by Zucker et. al. (1977). Their scheme is more computational expensive than ours and does not have a true probability interpretation.

II. EDGE CONTEXT

Consider all the row monotonically increasing paths which begin at any border pixel of the image above a selected pixel, go through the selected pixel, and end at some border pixel of the image below the selected pixel. Each such path represents a context for the pixel. Corresponding to each path and the observed pixel values on the path, there is an associated probability of edge state for the given pixel. Among all the paths there is some best 'edge' path which assigns the current pixel as an 'edge' pixel with a probability that is higher than the probability of every other 'edge' path. Similarly, there is some best 'no-edge' path which assigns the current pixel as an 'no-edge' pixel with a probability that is higher than the probability of every other 'no-edge' path.

According to Lee (1985), the edge detection problem can be formulated as a Baysian decision problem. The solution to this problem is: for each pixel position (r,c) of the image assign the edge state ε_{rc} as 'edge' if

$$P(\varepsilon^*_{rc} = \text{'}edge\text{'}|K) > P(\varepsilon^*_{rc} = \text{'}no-edge\text{'}|K) \quad (1)$$

and assign the edge state ε_{rc} as 'no-edge' otherwise. The context information which appears in *equation (1)* is denoted by K which is the facet model representation of each pixel's local neighborhood (Haralick et al. 1980) of the whole image. In line with *equation (1)*, the context dependent scheme assigns a pixel edge state 'edge' if the edge probability of the best 'edge' path is higher than the average no-edge probability of the best 'no-edge' paths and assigns a pixel 'no-edge' otherwise.

Let U_{rc} designate the set of all row monotonically increasing paths which begin at some border pixel of the image above or to the left of pixel (r,c) and terminate at pixel (r,c). The set L_{rc} designates the set of all row monotonically increasing paths which begin at (r,c) and terminate at some border pixel below or to the right of pixel (r,c). These are illustrated in *Figure 1* (a) and (b). Let $N_1(r,c) = \{(r-1,c-1),(r-1,c),(r-1,c+1),(r,c-1)\}$. and $N_2(r,c) = \{(r+1,c-1),(r+1,c),(r+1,c+1),(r,c+1)\}$. The set $U_{rc(pq)}$ where $(p,q) \in N_1(r,c)$ is defined as

$$U_{rc(pq)} = \{T : T \in U_{rc} \; and \; (p,q) \in T\}$$

Similarly, we can define

$$L_{rc(ij)} = \{T : T \in L_{rc} \; and \; (i,j) \in T\}$$

where $(i,j) \in N_2(r,c)$.

Z_{rc} designates the set of all row monotonically increasing paths begining from a border of the image passing through pixel (r,c) and continuing to another border pixel of the image which is just the join of all paths in U_{rc} with the paths in L_{rc}. Similarly, we can define $Z_{rc(pq,ij)}$ as the join of all paths in $U_{rc(pq)}$ with the paths in $L_{rc(ij)}$.

The set U^*_{rc} designates the set of all row monotonically increasing paths which begin at some border pixel of the image at the same row or above pixel (r,c) and terminate at pixel (r,c). The set L^*_{rc} designates the set of all row monotonically increasing paths which begin at pixel (r,c) and terminate at some border pixel at the same row or below the pixel (r,c). This is illustrated in *Figure 1* (c) and (d).

III. DETECTION ALGORITHM

For pixel (r,c), let $\underline{k}_{rc}$ designate its facet parameter representation and $f_{Z_{rc}}(\varepsilon^*_{rc})$ be the probability that its true edge state is ε^*_{rc} given the facet parameters of the best row monotonically increasing path T taking the direction θ^*_{rc0} through (r,c), where θ^*_{rc0} is the direction which maximizes $P(\theta^*_{rc}|\underline{k}_{rc})$. Let $g_{Z_{rc}}$ be

$$g_{Z_{rc}}(\varepsilon^*_{rc},\theta^*_{rc}) = \max_{T \in Z_{rc}} P(\varepsilon^*_{rc},\theta^*_{rc}|\underline{k}_{ij} : (i,j) \in T). \quad (2)$$

Then, $f_{Z_{rc}}(\varepsilon^*_{rc}) = g_{Z_{rc}}(\varepsilon^*_{rc},\theta^*_{rc0})$. It is noted that $f_{Z_{rc}}$ is not a function of θ^*_{rc0} because θ^*_{rc0} is a fixed value which is determined by

$$P(\theta^*_{rc0}|\underline{k}_{rc}) > P(\theta_{rc}|\underline{k}_{rc}) \; \forall \; \theta_{rc} \neq \theta^*_{rc0}.$$

Similarly, we can define $f_{Z_{rc(pq,ij)}}(\varepsilon^*_{rc})$. The average probability $\overline{f}_{Z_{rc}}(\varepsilon^*_{rc})$ is defined by

$$\overline{f}_{Z_{rc}}(\varepsilon^*_{rc}) = \frac{1}{16} \sum_{\substack{(pq)\in N_1(r,c) \\ (ij)\in N_2(r,c)}} f_{Z_{rc(pq,ij)}}(\varepsilon^*_{rc}). \quad (3)$$

The Baysian decision theory based edge detection scheme (1) is now, assigning ε^*_{rc} = 'edge' if

$$f_{Z_{rc}}(\varepsilon^*_{rc} = \text{'}edge\text{'}) > \overline{f}_{Z_{rc}}(\varepsilon^*_{rc} = \text{'}no-edge\text{'})$$

and ε^*_{rc} = 'non-edge', otherwise. To perform edge detection, we have to compute $f_{Z_{rc}}$ and then derive $\overline{f}_{Z_{rc}}$ from $f_{Z_{rc}}$. To compute $f_{Z_{rc}}$, we have to compute $g_{Z_{rc}}$ first. Analogous to the definition of $g_{Z_{rc}}(\varepsilon^*_{rc},\theta^*_{rc})$, $g_{U_{rc}}(\varepsilon^*_{rc},\theta^*_{rc})$

and $g_{L_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc})$ can be defined and we can derive

$$g_{Z_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc}) = \frac{P(\underline{k}_{rc})}{P(\underline{k}_{rc}|\varepsilon^*_{rc}, \theta^*_{rc})} g_{U_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc}) g_{L_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc}). \quad (4)$$

Similarly, we can also define

$$h_{U^*_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc}) = \max_{T \in U^*_{rc}} P(\varepsilon^*_{rc}, \theta^*_{rc}|\underline{k}_{ij} : (i,j) \in T). \quad (5)$$

After some derivations and mathematical manipulations (see Lee, 1985) we have,

$$g_{U_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc}) = \frac{P(\underline{k}_{rc}|\varepsilon^*_{rc}, \theta^*_{rc})}{P(\underline{k}_{rc})} *$$

$$\begin{aligned} \max\{ & \sum_{\theta^*_{r,c-1}, \varepsilon^*_{r,c-1}} g_{U_{r,c-1}}(\varepsilon^*_{r,c-1}, \theta^*_{r,c-1}) * \\ & a(\varepsilon^*_{r,c-1}, \theta^*_{r,c-1}, \varepsilon^*_{rc}, \theta^*_{rc}), \\ & \sum_{\theta^*_{r-1,c-1}, \varepsilon^*_{r-1,c-1}} h_{U^*_{r-1,c-1}}(\varepsilon^*_{r-1,c-1}, \theta^*_{r-1,c-1}) \\ & * a(\varepsilon^*_{r-1,c-1}, \theta^*_{r-1,c-1}, \varepsilon^*_{rc}, \theta^*_{rc}), \\ & \sum_{\theta^*_{r-1,c}, \varepsilon^*_{r-1,c}} h_{U^*_{r-1,c}}(\varepsilon^*_{r-1,c}, \theta^*_{r-1,c}) * \\ & a(\varepsilon^*_{r-1,c}, \theta^*_{r-1,c}, \varepsilon^*_{rc}, \theta^*_{rc}), \\ & \sum_{\theta^*_{r-1,c+1}, \varepsilon^*_{r-1,c+1}} h_{U^*_{r-1,c+1}}(\varepsilon^*_{r-1,c+1}, \theta^*_{r-1,c+1}) \\ & * a(\varepsilon^*_{r-1,c+1}, \theta^*_{r-1,c+1}, \varepsilon^*_{rc}, \theta^*_{rc})\} \end{aligned} \quad , \quad (6)$$

where $a(\varepsilon^*_0, \theta^*_0, \varepsilon^*_1, \theta^*_1)$ is an edge consistency function which has maximal value when the edge direction at the immediate adjacent neighbor agrees with that at the center, and the expected neighbor direction agrees with the true neighbor direction and is defined in functional forms in Lee, 1985. *Equation (6)* says that for each edge label $\varepsilon^*_{rc}, \theta^*_{rc}$, the conditional probability of $\varepsilon^*_{rc}, \theta^*_{rc}$ given $\{\underline{k}_{ij} : (i,j) \in T\}$ of the best path T can be obtained on the basis of the previously computed $g_{U_{r,c-1}}$, and on the previously computed $h_{U^*_{r-1,c-1}}, h_{U^*_{r-1,c}}, h_{U^*_{r-1,c+1}}$ coming from the row above the current row. *Equation (6)* specifies a recursive neighborhood operator which scans the image in a top down left right scan to produce for each pixel (r,c) and for each edge label $\varepsilon^*_{rc}, \theta^*_{rc}$ the probability $g_{U_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc})$ providing we can demonstrate a way to compute $h_{U^*_{rc}}$ which can be derived as,

$$h_{U^*_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc}) = \max\{g_{U_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc}), \frac{P(\underline{k}_{rc}|\varepsilon^*_{rc}, \theta^*_{rc})}{P(\underline{k}_{rc})} * \sum_{\theta^*_{r,c+1}, \varepsilon^*_{r,c+1}} h_{U^*_{r,c+1}}(\varepsilon^*_{r,c+1}, \theta^*_{r,c+1}) a(\varepsilon^*_{r,c+1}, \theta^*_{r,c+1}, \varepsilon^*_{rc}, \theta^*_{rc})\}. \quad (7)$$

Equation (7) states that $h_{U^*_{rc}}$ can be recursively computed from $g_{U^*_{rc}}$ and the previous $h_{U^*_{r,c+1}}$ in a right left scan of a row done after $g_{U^*_{rc}}$ has been computed. To start the recursive calculation (7), we take $h_{U^*_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc}) = g_{U_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc})$ for that column position c which is the rightmost position.

An absolutely mirror image derivation applies to $g_{L_{rc}}$. It can be computed by a bottom up right left scan of the image recursively from $g_{L_{r,c+1}}$, $h_{L^*_{r+1,c-1}}$, $h_{L^*_{r+1,c}}$, and $h_{L^*_{r+1,c+1}}$ which had been computed. A left right scan of row r is then performed to compute $h_{L^*_{rc}}$.

As soon as $g_{L_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc})$ has been computed, it can be combined with $g_{U_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc})$ to compute $g_{Z_{rc}}(\varepsilon^*_{rc}, \theta^*_{rc})$ (see (4)). In practice, the only useful θ^*_{rc} for edge detection is θ^*_{rc0} which maximizes $P(\theta^*_{rc}|\underline{k}_{rc})$. Hence, we determine θ^*_{rc0} first and then only compute $g_{rc}(\varepsilon^*_{rc}, \theta^*_{rc0})$ for both ε^*_{rc}='edge' and 'no-edge'. The two probability terms $f_{Z_{rc}}(\varepsilon^*_{rc})$ and $\bar{f}_{Z_{rc}}(\varepsilon^*_{rc})$ are then readily available. The edge state of each pixel (r,c) can now be labeled by means of the rule of *equation (4)*.

To perform the recursive algorithms (6), (7), etc., we need the probability ratio $\frac{P(\underline{k}_0|\varepsilon^*_0, \theta^*_0)}{P(\underline{k}_0)}$. The detailed derivation of the conditional probabilities is shown in Lee, 1985.

VI. EXPERIMENTAL RESULTS

To understand the performance of the context dependent edge detector, we examine the behavior of the context edge detector on one well structured simulated image and two real images. We then compare the results with the context free second derivative zero-crossing edge operator (Haralick, 1984) to see how and in what degree the context information can improve the operator.

The simulated test image is a noisy bar image. The image size is 100 X 50 pixels. The pixel intensity is 0 for dark bars and 175 for white bars. A 2 X 2 averaging is applied to this image to simulate ideal single pixel width edge lines. A zero mean Gaussian noise with standard deviation 40 is then added to this image. We fit each 5 X 5 neighborhood of the test image by a cubic polynominal and then apply the context dependent operator and second derivative zero crossing edge operator to it. In order to quantitively see the difference in performance of these two schemes, we measure for the whole image the conditional probabilities $P(E'|E^*)$, $P(E^*|E')$, $P(E'|\bar{E}^*)$, and $P(\bar{E}'|E^*)$ where E' and $\bar{E}'$ designate the assigned 'edge' and 'no-edge' pixel sets and E^* and $\bar{E}^*$ designate the true 'edge' and 'no-edge' pixel sets. When determining $P(E'|E^*)$ and $P(E^*|E')$, the adjustable parameters of each edge operator are chosen to equalize these two conditional probabilities ($P(E'|E^*) \approx P(E^*|E')$). The performance in terms of $P(E'|E^*)$ and $P(E^*|E')$ are shown in *Table 1*. When determining $P(E'|\bar{E}^*)$ and

operator	$P(E'\|E^*)$	$P(E^*\|E')$
context dependent	0.8600	0.8600
context free	0.6950	0.6814

Table 1. $P(E'|E^*)$ and $P(E^*|E')$ values of the context dependent edge operator and the context free edge operator.

$P(\bar{E}'|E^*)$, the adjustable parameters of each edge operator are chosen to equalize these two conditional probabilities ($P(E'|\bar{E}^*) \approx P(\bar{E}'|E^*)$). The probability ($P(E'|\bar{E}^*)$ for the context free operator is 0.02 while the probability for the context operator is 0. The results show that the context scheme performs much better than the context free operator.

We now apply the context dependent edge detector to two 3-D range images (128 X 128 pixels) of man-made objects. To these images a zero mean Gaussian noise of standard deviation 30 is added. The noisy images are shown in *Figures 2(a), and 3(a)* . A 5 X 5 cubic polynominal fitting is applied to the noisy images followed by the applications of both the context free second derivative zero-crossing edge operator and the context dependent edge operator. The context bounds we used for both images are ϵ_{min}= 0.1 and ϵ_{max}=10. The results of the context free edge operator are shown in *Figure 2(b), and 3(b).* The results of the context dependent edge operator are shown in *Figure 2(c), and 3(c).* It can be easily verified visually that the edge images of the context dependent edge operator show better connectivity and much less noise than the context free edge operator.

In order to see the performance of the context dependent edge operator under different noise levels. The context dependent and the context free edge operators are applied to images with noises of standard deviations 10, 20, 30, 40, and 50, respectively. The general edge evaluator is employed to measure the performance scores of both edge operators. The edge results which maximize the edge score (see Haralick and Lee, 1988) are shown in *Figure 4.* It is also easy to verify that the context edge operator has better performance over the context free edge operator by a visual evaluation. The edge score curves are shown in *Figure 5.*

V. CONCLUSIONS

We have developed an edge detection scheme from a Bayesian theoretic framework. The edge detection makes use of the edge context of the entire image. For a given pixel the edge context of the whole image related to this pixel is organized as monotonically increasing paths which begin at any border pixels of the image above the selected pixel, pass through the selected pixel, and end at some border pixels of the image below the selected pixel. We will assign a pixel edge state 'edge' if the edge probability of the best 'edge' path is higher than the average probability of the best 'no-edge' paths.

We derived the algorithms of finding the best paths as a recursive scheme. It starts with a top down left right scan of the image followed by a right left scan. It then performs a bottom up right left scan of the image followed by a left right scan.

Experiments were performed to illustrate the validity of the context edge detector and the general edge evaluator. We have compared the performance of the context edge detector with the context free second directional derivative zero-crossing edge operator. The results show that the context edge detector has superior performance.

The approach of using full context and local coherence can be extended to the detection of other local image features.

REFERENCES

1. Robert Haralick, 'Zero-Crossing of Second Directional Derivative Edge Operator', *Society of Photogrammetric Instrumentation Engineers Symposium on Robot Vision*, Washington, D.C., May, 1982.
2. Robert Haralick, 'Digital Step Edges from Zero Crossing of Second Directional Derivatives', *IEEE Trans. PAMI*, Vol. PAMI-6, No. 1, January, 1984, pp. 58-68.
3. Robert Haralick, 'Decision Making in Context', *IEEE Trans. PAMI*, Vol. PAMI-5, No. 4, July 1983, pp. 417-28.
4. Robert Haralick and L. T. Watson, 'A Facet Model for Image Data', *CGIP*, Vol. 15, 1981, pp. 113-29.
5. Robert Haralick and James S.J. Lee, 'Context Dependent Edge Detection and Evaluation', To appear *International Journal of Pattern Recognition*, 1988.
6. James Lee, 'Edges From Image', PH.D. dissertation VPI & SU, August, 1985.
7. David Marr and Ellen Hildreth, 'Theory of Edge Detection', *Proc. Royal Society of London*, B, Vol. 207, 1980, pp. 187-217.
8. A. Martelli, 'An Application of Heuristic Search Methods to Edge and Contour Detection', *Commun. ACM* 19, 2 February 1976, pp. 73-83.
9. U. Montanari, 'On the Optimal Detection of Curves in Noisy Pictures', *Commun. ACM* 14, 5 May 1971, pp. 335-45.
10. Steve W. Zucker and Robert A. Hummel and Azriel Rosenfeld, 'An application of relaxation labeling to line and curve enhancement', *IEEE Trans. Computers*, Vol. C-26, No.4, April 1977, PP. 394-403.

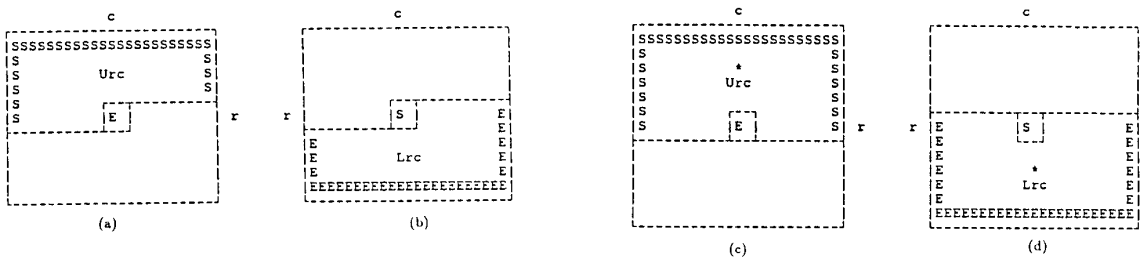

Figure 1. illustrates (a) the set U_{rc} and (b) the set L_{rc}. U_{rc} is the set of all row and column monotonically increasing paths beginning at a border of the image above or to the left of the pixel (r,c) and terminating at pixel (r,c). L_{rc} is the set of all row and column monotonically increasing paths beginning at the pixel (r,c) and terminating at the border of the image below or to the right of the pixel. Similarly, the sets U^*_{rc} and L^*_{rc} are illustrated in (c) and (d).

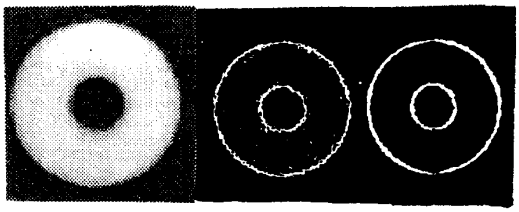

Figure 2. The noisy object image 1 (a), its context free edge image (b), and context dependent edge image(c). The window size for cubic polynomial fitting is 5 X 5.

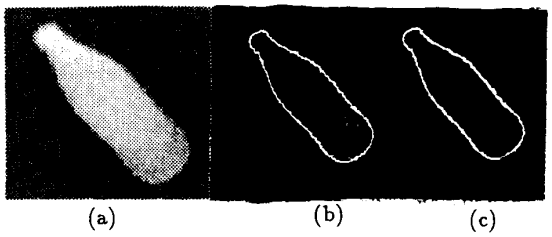

Figure 3. The noisy object image 2 (a), its context free edge image (b), and context dependent edge image(c). The window size for cubic polynomial fitting is 5 X 5.

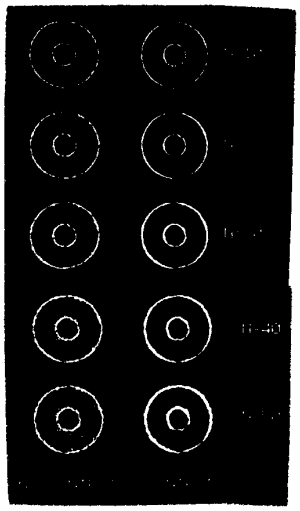

Figure 4. Shows the edge images of the context dependent edge operator and the context free edge operator applied to the image shown in *Figure 12* with noise standard deviations 10, 20, 30, 40, and 50, respectively.

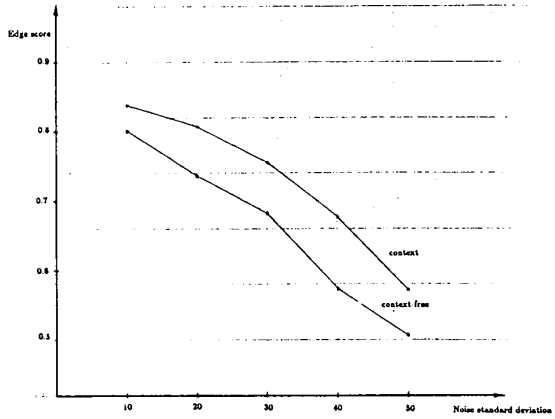

Figure 5. Shows the edge scores of the context dependent edge operator and the context free edge operator applied to the image shown in *Figure 12* with noise standard deviations 10, 20, 30, 40, and 50, respectively.