Applying Non-stationary Noise Estimation to Achieve Contrast Invariant Edge Detection

Paul Wyatt and Hiroaki Nakai

Multimedia Laboratory, Toshiba Corporate RDC, 1 Komukai-Toshiba-cho,
Saiwai-ku, Kawasaki 212-8582, Japan
wyatt@eel.rdc.toshiba.co.jp, hiroaki.nakai@toshiba.co.jp

Abstract. To recognize or identify objects it is desirable to use features which are minimally affected by changes in lighting and non-stationary noise. This requires accurate estimation of both signal and noise.

In response to this challenge, this paper proposes a method for estimation of non-stationary isotropic noise based on steering filters to directions perpendicular and parallel to the local signal. From the filter responses in this direction equations for signal and noise are obtained which lead to an edge detection method dependent solely upon local signal-to-noise ratio. The proposed method is compared to various common edge detection methods from the literature, on synthetic and real images. Quantitative improvement is demonstrated on synthetic images and qualitative improvement on real images.

1 Introduction

The extraction of edges and curves is of considerable interest to the vision community, as is evident from the large, diverse literature on the subject, e.g. [1, 2, 3, 4, 5, 6]. From the literature, perhaps three key ideas have emerged. Firstly, orienting filters according to some notion of local optimality: often defined as the direction in which the least squares energy is maximized [1]. Secondly, the idea that on an edge the responses to filters at different scales must be maximally in phase [1, 3] : essentially meaning that as an edge is an odd function, at an edge responses to odd filters will be maximal and even will be zero. The third idea is that structures should be associated with filter scale [4, 6, 7].

However, two related problems remain difficult: estimation of contrast change and non-stationary noise. Unaccounted for, both can lead to poor repeatability and instability. Contrast correction has been most successfully attempted using the Retinex transform [8]. However, it can err in dark regions. Possibly more promising is estimation of noise. If local signal s and noise σ_n are estimated, signal-to-noise ratio (SNR) can be established. This leads toward an edge measure such as $\left(1 - \frac{\sigma_n}{s}\right)$. Assuming both signal and noise share the same relation to contrast or illumination, this implies detected structures would be contrast invariant. The assumption is not unreasonable as edges occur where phase is congruent [1, 3] leading us to infer that signal power is significantly greater than noise power for the case of uncorrelated, isotropic noise. Consequently, even in

P.J. Narayanan et al. (Eds.): ACCV 2006, LNCS 3852, pp. 742–751, 2006.

poor contrast areas an edge requires that the ratio between signal and noise is significant.

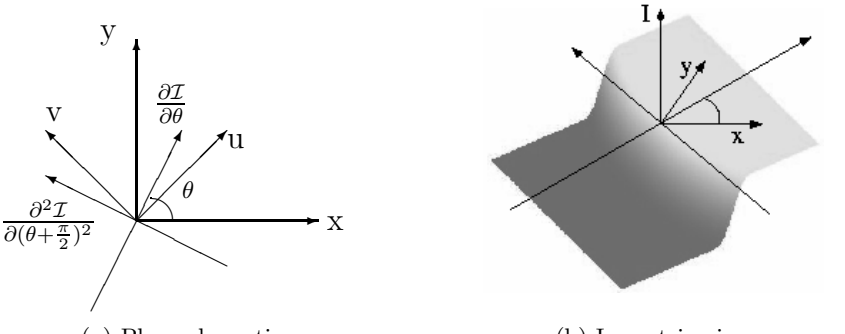
Approaches to noise estimation are often statistically based [3, 9, 10]. They assume that a high pass filtered image contains solely noise coefficients [3] and from an assumption on the expected noise distribution estimate noise variance [3, 11]. Other approaches have attempted to account for structure, for example through anisotropic evolution of the intensity [6]. However, this approach still requires an initial noise estimate. More recent alternatives have attempted to suppress structure, estimating noise from the remainder [12].

This paper focuses on estimation of non-stationary, uncorrelated isotropic noise as part of the structure detection process. Although the focus is on edge detection, the method is general and could equally well be applied to other types of feature, e.g. corners. We make two contributions. Firstly, an integrated model for simultaneously estimating the local noise and structure to obtain an edge measure dependent solely upon the SNR. This improves stability to contrast change. Secondly, we show how scale affects the problem of noise estimation and its applicability to distinguishing a step edge from shadowing.

2 A Combined Edge and Noise Model

This section first focuses on the edge and noise model, considering a single scale. It then proceeds to consider the differences for edges at different scales; i.e. with different degrees of blurring.

Figure 1 contains two diagrams showing an abstraction of a generic edge, a step change in intensity $\mathcal{I}$, positioned at angle θ to image axes x and y. Axes u and v are aligned at $\frac{\pi}{4}$ to x and y. $\frac{\partial \mathcal{I}}{\partial \theta}$ and $\frac{\partial^2 \mathcal{I}}{\partial(\theta+\frac{\pi}{2})^2}$ denote two of the partial derivatives of $\mathcal{I}(x, y)$ perpendicular (orthogonal) and parallel to the edge, our third coordinate system. It is assumed that non-stationary Gaussian noise $\mathcal{N}(0, \sigma_n^2)$ is present. It has been shown, see [1, 3], that a one dimensional edge is comprised



(a) Plan schematic (b) Isometric view

Fig. 1. Schematic for a general curve-like point. The derivatives are derivatives of the intensity function perpendicular and parallel to the local edge structure.

solely of odd sinusoids: a sum which is theoretically in phase at the location of the step. This sum must be large relative to σ_n to facilitate detection. Conversely, the sum of co-sinusoids will be small. These properties can be stated mathematically:

$$\left|\frac{\partial \mathcal{I}}{\partial \theta}\right| \gg \sigma_n \ , \left|\frac{\partial^2 \mathcal{I}}{\partial \theta^2}\right| \sim \left|\mathcal{N}(0, \sigma_n^2)\right| \ . \tag{1}$$

Moving perpendicularly away from the edge (axis θ), equations 1 do not hold. The second derivative $\frac{\partial^2 \mathcal{I}}{\partial \theta^2}$ will increase, making this pair unsuitable for noise estimation. However, equations 1 consider behaviour solely in spatial direction θ. Derivatives in directions θ and $\theta + \frac{\pi}{2}$ are related via curvature κ:

$$\kappa = \frac{\frac{\partial \mathcal{I}}{\partial \theta}\frac{\partial^2 \mathcal{I}}{\partial(\theta+\frac{\pi}{2})^2} - \frac{\partial \mathcal{I}}{\partial(\theta+\frac{\pi}{2})}\frac{\partial^2 \mathcal{I}}{\partial \theta^2}}{\left[\left(\frac{\partial \mathcal{I}}{\partial \theta}\right)^2 + \left(\frac{\partial \mathcal{I}}{\partial(\theta+\frac{\pi}{2})}\right)^2\right]^{\frac{3}{2}}} \approx \frac{\frac{\partial \mathcal{I}}{\partial \theta}\frac{\partial^2 \mathcal{I}}{\partial(\theta+\frac{\pi}{2})^2}}{\left[\left(\frac{\partial \mathcal{I}}{\partial \theta}\right)^2 + \left(\frac{\partial \mathcal{I}}{\partial(\theta+\frac{\pi}{2})}\right)^2\right]^{\frac{3}{2}}} \tag{2}$$

as $\frac{\partial^2 \mathcal{I}}{\partial \theta^2} \sim 0$. As $\frac{\partial^2 \mathcal{I}}{\partial(\theta+\frac{\pi}{2})^2}$ varies, the local structure changes from a straight line to more tightly curved structures (eventually corner-like). Assuming curvature varies smoothly then $\frac{\partial \mathcal{I}}{\partial(\theta+\frac{\pi}{2})}$ should be relatively small. For the case where $\kappa = 0$, a straight edge, $\frac{\partial^2 \mathcal{I}}{\partial(\theta+\frac{\pi}{2})^2} = 0$. In this case, if $\frac{\partial \mathcal{I}}{\partial(\theta+\frac{\pi}{2})} \neq 0$, it must be responding to some disturbance. *We make the assumption that the measurement* $\frac{\partial \mathcal{I}}{\partial(\theta+\frac{\pi}{2})}$ *must be related to the local noise* σ_n. Generally, for any smoothly curving structure, it can reasonably be assumed that this will be true. With equation 2 this observation completes our edge model:

$$\frac{\partial^2 \mathcal{I}}{\partial(\theta + \frac{\pi}{2})^2} \propto \kappa \ , \text{and} \tag{3}$$

$$\frac{\partial \mathcal{I}}{\partial(\theta + \frac{\pi}{2})} \sim \mathcal{N}(0, \sigma_n^2) \ . \tag{4}$$

Although equation 1 is directly suitable for implementation, equation 4 is not. It yields, not σ_n, but one sample from a noise distribution at each point. An estimate of σ_n, $\hat{\sigma_n}$ can be made using samples from a small region about each point. Using δ to denote the extent of this area, about point $x = x_i, y = y_i$,

$$\hat{\sigma}_n \approx \sqrt{\frac{1}{4\delta^2}\int_{x_i-\delta}^{x_i+\delta}\int_{y_i-\delta}^{y_i+\delta}\left(\frac{\partial \mathcal{I}}{\partial(\theta + \frac{\pi}{2})}\right)^2 dydx} \ . \tag{5}$$

Note that $E[\frac{\partial \mathcal{I}}{\partial(\theta+\frac{\pi}{2})}]$ is expected to be zero and is therefore not required for the estimation of $\hat{\sigma}_n$ in equation 5. δ should be set in proportion to the spatial extent of the derivatives filter. Practically a value of twice the largest filter dimension in pixels is used. Furthermore, note that $\hat{\sigma}_n^2$ is scaled by the filter used in obtaining $\frac{\partial \mathcal{I}}{\partial(\theta+\frac{\pi}{2})}$. $\hat{\sigma}^2 \sim \sigma_{true}^2 \sum_i f_i^2$ where f_i are the filter coefficients.

This becomes important if more than one filter scale is used. Using the estimate of signal energy $\frac{\partial \mathcal{I}}{\partial \theta}$ and $\hat{\sigma_n}$, the edge detection measure $\Pr(S)$ is then simply

$$\Pr(S) = \left(\frac{\left\lfloor \left| \frac{\partial \mathcal{I}}{\partial \theta} \right| - \alpha \sigma_n \right\rfloor}{\left| \frac{\partial \mathcal{I}}{\partial \theta} \right|} \right) , \tag{6}$$

where $\lfloor f() \rfloor$ bounds the function, $f()$, from below with zero. As the energy response function is statistical in nature, it suppresses a fraction of all noise responses. For example, setting $\alpha = 1.6449$ will suppress 90% of noise. Practically, α can be adjusted according to whether it is preferable to suppress noise or obtain every potential structure. Results in this paper use $\alpha = 2.5$.

2.1 Natural Scale and Edge Detection

As the model we have defined assumes that fixing filters at one scale is sufficient for edge detection, we briefly justify this with respect to scale invariance. Scale invariance, see [4], shows that a filter has a size or scale by which it may be parameterized and normalisation by this parameter makes the response independent of scale. For instance, the one dimensional family of Gaussians $\frac{\partial^n \mathcal{G}(x)}{\partial x^n}$, $n = 0, 1, 2$, can be parameterized by standard deviation γ. Scale invariance is achieved through multiplication by γ^n.

For an ideal edge, with no blur, the response of $\gamma \partial \mathcal{G}$ will theoretically be constant until γ increases such as to cause interaction with a second edge. After this the response decreases [4]. In practice, for $\gamma < 2$, the filter approximation tends to yield a quickly rising step which plateaus between $1.5 \leq \gamma \leq 3$. Consequently, for an ideal edge using a filter with $\gamma \geq 2$ should yield constant results. For shadowed edges, with pre-blur γ_s, the response of $\gamma \partial \mathcal{G}$, $\gamma \leq \gamma_s$, grows linearly plateauing at $\gamma = \gamma_s$. Consequently, we can detect (and remove) shadowed edges by comparing the ratio of coefficients at two or more different scales. This part of the contrast problem is not further examined in this paper.

3 Implementation

Having detailed a model for an edge in non-stationary noise along with the requisite equations for estimating these properties we now detail the approximations made between the theoretical model and its practical implementation. For context, we first state the complete algorithm.

1. Convolve the image with a Gaussian, of deviation $\gamma_1 = 2$. Calculate $\partial \mathcal{I}$ in directions x, u, y and v.
2. Estimate the global noise, σ_g, using an Expectation Maximization (EM) on the magnitudes of $\partial \mathcal{I}$ in directions x and y to obtain weights (ω_1, ω_2) and variances (σ_1^2, σ_2^2) for a Gaussian Mixture Model (GMM) of signal and noise. $(\mu_1 = \mu_2 = 0.)$
3. At each point: estimate θ and use it to select signal and noise samples from amongst the four derivatives.

4. At each point: evaluate equations 5 and 6.
5. If a binary edge map is desired, threshold using an estimate of the existing fraction of edges, $\min(\omega_1, \omega_2)$, from the EM algorithm.

Considering the filters, normal image blur normally has deviation less than 2 pixels: the Gaussian smoothing filter's is set to the same. It is truncated at 2 deviations. Derivatives are simple central differences, with the diagonal directions scaled appropriately by $\frac{1}{\sqrt{2}}$. The separation of filter responses into noise and signal is then achieved using the least squares estimate of θ: $\theta = \arctan\left(\frac{\partial \mathcal{I}}{\partial y} / \frac{\partial \mathcal{I}}{\partial x}\right)$ [13]. For $\frac{\pi}{8} \leq \theta \leq \frac{3\pi}{8}$ and for $\frac{5\pi}{8} \leq \theta \leq \frac{7\pi}{8}$ derivatives along axes u and v are used, otherwise along x and y. With respect to the choice of filters, first derivatives in four directions is simple and suffices. Although equations 1,3 and 4 are specified in terms of the edge co-ordinate system θ and $\theta + \frac{\pi}{2}$, the question of how to estimate this is not simple. Although other methods for steering filters exist, e.g. [2], they can be computationally expensive in practice and the complex steering mechanism can induce errors for small (7x7 pixels) filters. However, using only four filters, responses away from the axes are affected by image quantisation. For example, edges oriented at $\theta = \frac{(2n+1)\pi}{8}, n = 0, 1, 2$ fall directly between the filter directions. Noise in these directions will be over-estimated leading to decreased stability. If speed is less important than accuracy, more complex steering techniques could be used [13, 2].

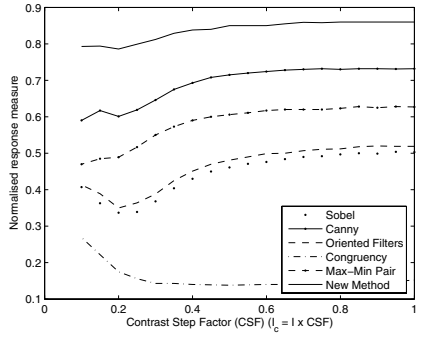
Equation 6 is calculated as stated. Two practical changes are made to 5. Firstly, the integral over the image grid is replaced by convolution with a Gaussian. This second Gaussian's deviation γ_2 is set at twice that of the Gaussian used for the first image convolution: $\gamma_2 = 2\gamma_1 = 4$. Secondly, if $\hat{\sigma_n} \leq \frac{\sigma_g}{20}\ \hat{\sigma}_n$ is set equal to the global noise.

Finally, estimation of σ_g models wavelet coefficients as being comprised, at each scale, by a two state GMM [11]. One state has large variance and denotes structure, the other small variance and denotes noise. Both have zero mean. This idea is also used in wavelet based denoising [10]. The parameters for this model are fitted using an EM algorithm, yielding σ_g and weights, ω_1, ω_2, for the fractional split between edges and noise. The threshold for equation 6 is chosen to obtain this fraction of image points as edges. The EM algorithm is selected simply as it is one method appropriate for fitting a model to unlabeled data.

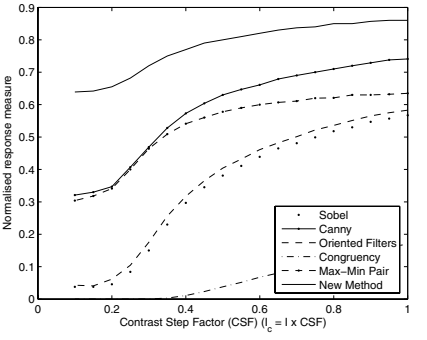
4 Experimental Results

Our method has been evaluated on a wide range of images; real and synthetic. As ground truth is difficult to establish for real images, the performance of the noise estimation is established on synthetic images. Sample results from many tested real images are given.

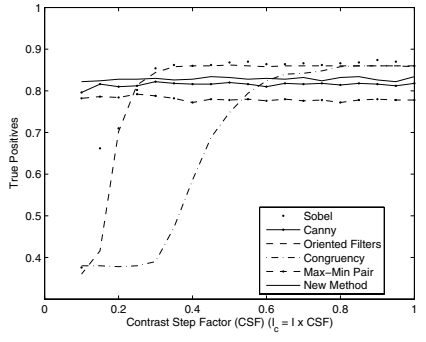
In our first test, synthetic images containing a mixture of curved and straight lines were created. To these, noise with a deviation equal to $\frac{1}{4}$ the edge size is added. Note that the image's left and right hand halves are exact duplicates. Then, a contrast step is applied to one half of the image. This yields images



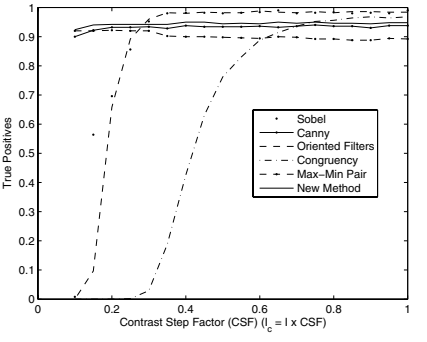
(a) Mean response for whole image



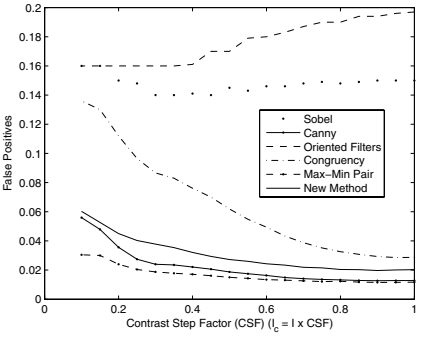
(d) Mean response for affected region



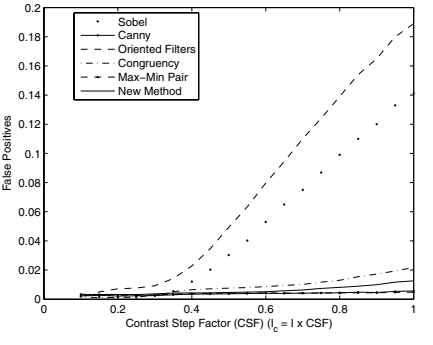
(b) True Positives for whole image



(e) True Positives for affected region



(c) False Positives for whole image



(f) False Positives for affected region

Fig. 2. Evaluation of common measures on Synthetic Test Images containing various oriented and curved structures

with varying contrast but constant SNR: suitable to test whether a method is stable with respect to contrast change. As the images are synthetic, ground truth is known. Various measures are evaluated; True positives: number pixels correctly classified; False positives: number incorrectly classified and, the edge detection energy (equation 6) at an edge. The tests were repeated to remove any

statistical bias from a particular set of generated noise. Stability was evaluated for Sobel, Canny, oriented bandpass filters, phase congruency, pseudo-steering through taking the maximum and minimum of four oriented filters to be signal and noise respectively (Max-Min) and the proposed edge detection methods. Results are shown in figure 2. As can be seen from figure 2, (a) and (d), the proposed method produces a more stable edge energy measure with respect to contrast change than the tested alternatives. Although it experiences some disturbance, for contrast step factors (CSF) ≤ 0.4, this is smaller than for the alternatives. True positives for most methods remain stable down to a CSF of 0.1. In this respect, no method is clearly better. A final point is that increases in false positives in 2(c) reflects the fact that global noise will be underestimated in the presence of significant contrast changes across the image. The proposed method is also affected, despite estimating noise locally, as it will make a percentage of errors in regions where there are no *real* edges: steering the estimation to the direction perpendicular to the strongest response (which in these regions is noise) leads to under-estimation of noise.

Our second set of tests repeats the experiment on synthetic images with real images to which noise and contrast steps were applied. Two sample images from amongst these are shown in figure 3;from the 'Bad Etting' sequence, available at iw1www.ira.uka.de/image_sequences, and the 'Graffiti' data set, available at www.inrialpes.fr/lear/people/Mikolajczyk/). The only difference from the first test, is that after adding noise, the images were requantised to 8 bits. The edge energy functions for the Canny edge function and the proposed method are shown for these images in figures 4 and 5. The Canny edge function is used as the comparison simply as it is probably the most widely used and available

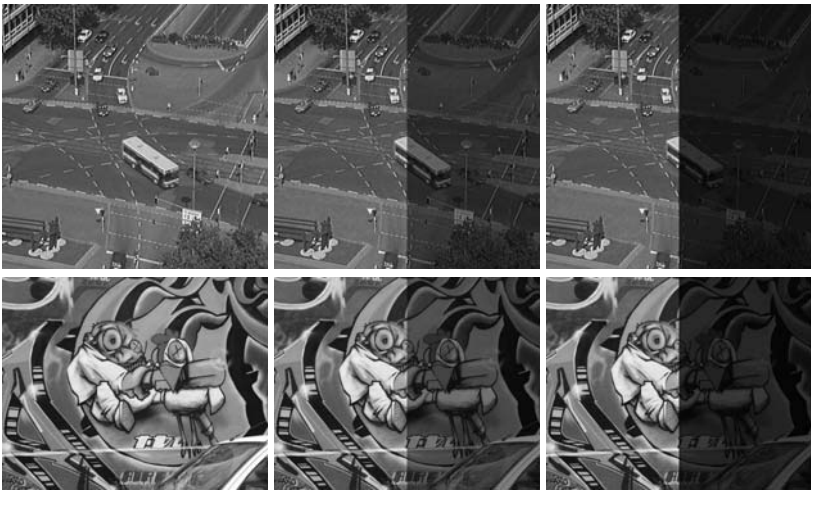
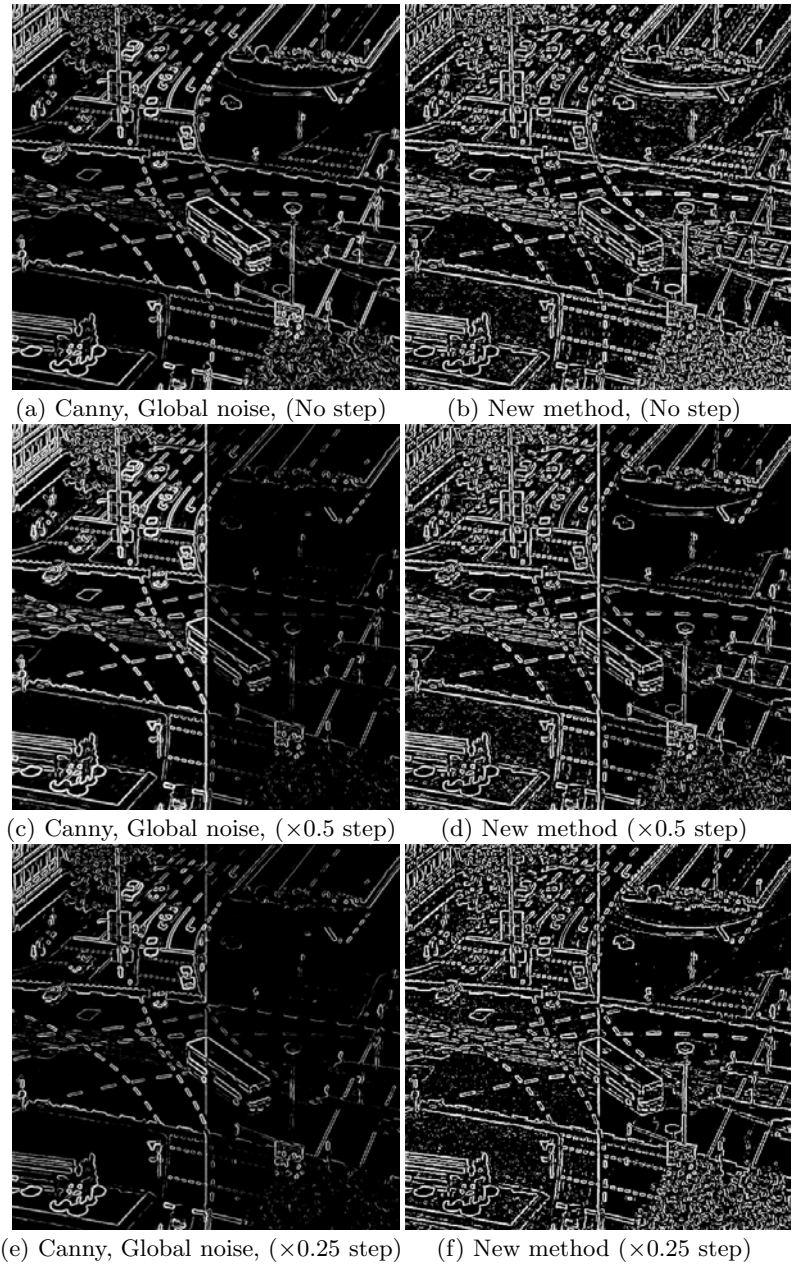


Fig. 3. The image pairs, with contrast steps, used for the tests shown in fig. 4 and 5

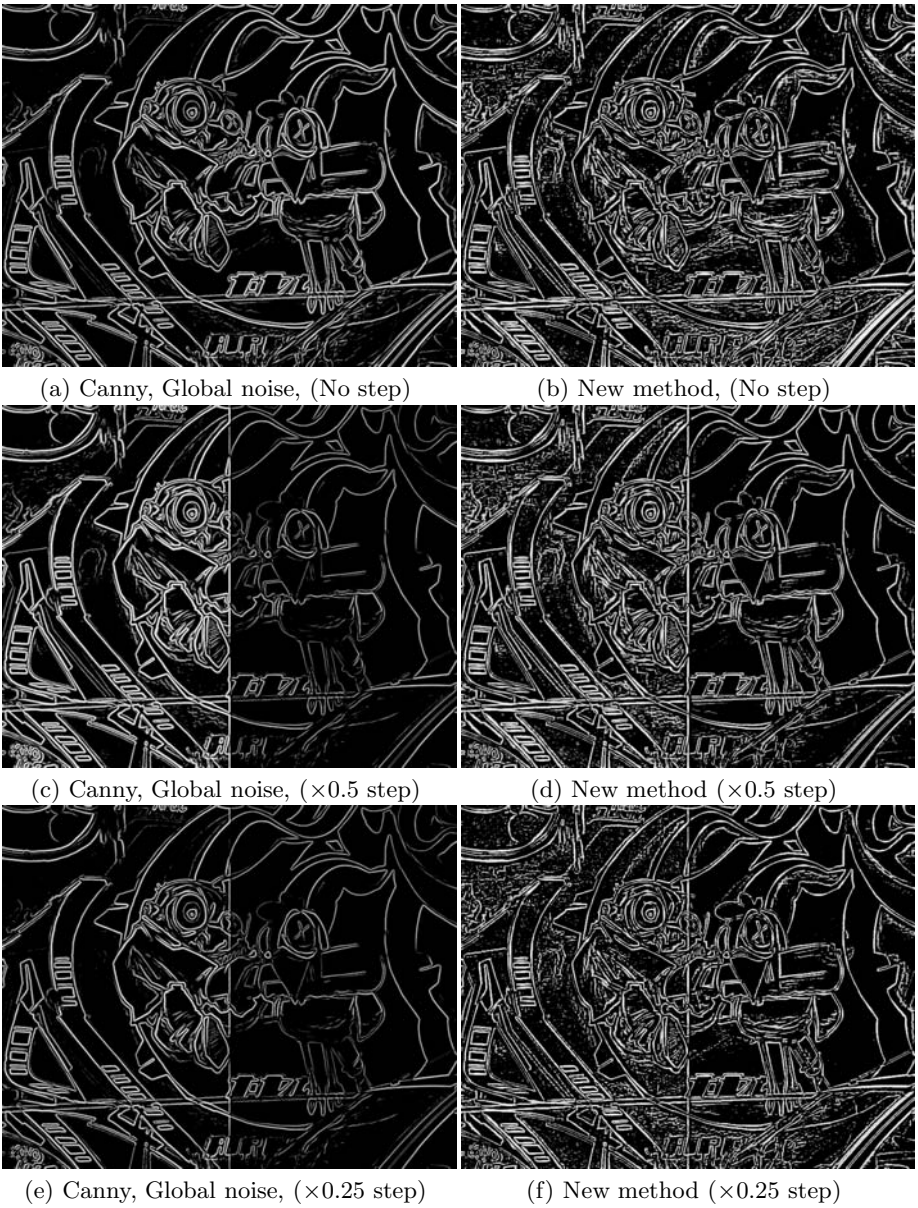


(a) Canny, Global noise, (No step) (b) New method, (No step)

(c) Canny, Global noise, (×0.5 step) (d) New method (×0.5 step)

(e) Canny, Global noise, (×0.25 step) (f) New method (×0.25 step)

Fig. 4. Test set 1: Comparison of Edge detection methods. The two columns show the new method versus the standard Canny edge function, for the original image and with contrast steps of ×0.5 and ×0.25 applied. Images additionally had 2% Gaussian isotropic noise added, *prior* to the contrast step being applied. Note that the images show the edge response energies, and are *NOT* binary edge images.



(a) Canny, Global noise, (No step) (b) New method, (No step)

(c) Canny, Global noise, (×0.5 step) (d) New method (×0.5 step)

(e) Canny, Global noise, (×0.25 step) (f) New method (×0.25 step)

Fig. 5. Test set 2: Edge detection methods for original image and with varied contrast and noise. The two columns show detection the new method versus the standard Canny edge function, for the original image and with contrast steps of ×0.5 and ×0.25 applied. These images have additionally had 2% Gaussian isotropic noise added, *prior* to the contrast step being applied. Note that the images show the edge response energies, and are *NOT* binary edge images.

method. As can be seen, the energies obtained from the proposed method are noticeably more stable with respect to contrast change. A final point is that weak structure can vanish with requantisation: for $CSF = \frac{1}{\delta I}$, a step of less than $\frac{1}{2}\delta I$ becomes constant. Noise magnifies this effect. Generally, from testing on various sequences, e.g. traffic, people, the method appears stable and reliable. Improvement relative to a standard method, e.g. Canny, depends upon whether contrast changes are present and whether the image quantisation has left the structure intact.

5 Summary and Conclusions

We have presented a method for incorporating local noise estimation into edge detection, thereby improving resilience to illumination change. Testing on synthetic and real data demonstrated improvement over previous methods.

References

1. Canny, J.: A computational approach to edge detection. IEEE Transactions PAMI **8** (1986) 679–698
2. Felsberg, M., Sommer, G.: The Monogenic Signal. IEEE Transactions Signal Processing **49(12)** (2001) 3136–3144
3. Kovesi, P.: Image Features from Phase Congruency. Videre: Journal of Computer Vision Research **1(3)** (1999) 1–27
4. Lindeberg, T.: Edge Detection and Ridge Detection with Automatic Scale Selection. IJCV **30(2)** (1998) 117–153
5. Pellegrino, F., Vanzella, W., Torre, V.: Edge Detection Revisited. IEEE: Systems, Man and Cybernetics **34(3)** (2004)
6. Perona, P., Malik, J.: Scale-space and Edge Detection Using Anisotropic Diffusion. IEEE Transactions PAMI **12(7)** (1990) 629–639
7. Elder, J., Zucker, S.: Local Scale Control for Edge Detection and Blur Estimation. IEEE Transactions PAMI **20(7)** (1998) 699–716
8. Adjeroh, D.: On Ratio Based Color-Indexing. IEEE Transactions Image Processing **10(1)** (2001) 36–48
9. Olsen, S.: Noise variance estimation in images. Graphic Models and Image Processing **55(4)** (1993) 319–323
10. Starck, J., Murtaugh, F.: Automatic Noise Estimation from the Multiresolution Support. Publ. of the Astronomical Soc. of the Pacific **110** (1998) 193–199
11. Crouse, M., Nowak, R., Baraniuk, R.: Wavelet-Based Statistical Signal Processing Using Hidden Markov Models. IEEE Transactions Signal Processing **46(4)** (1998) 886–902
12. Corner, B., Narayanan, R., Reichenbach, S.: Noise estimation in remote sensing imagery using data masking. J. Remote Sensing **24(4)** (2003) 689–702
13. Freeman, W., Adelson, E.: The design and use of steerable filters. IEEE Transactions PAMI **13(9)** (1991) 891–906