Estimating Local Image SNR From Gradient Angular Dispersion

Peter H. Gregson*
Department of Electrical Engineering,
Technical University of Nova Scotia,
Halifax, Nova Scotia, B3J-2X4
e-mail address: gregson@tuns.ca

Abstract

Knowledge of the signal to noise ratio (SNR) of an image is necessary if appropriate thresholds are to be selected for use in automated feature detection algorithms [2]. Estimation of SNR has proven computationally expensive previously because algorithms have been extensions of one dimensional approaches.

Analysis reveals that the weighted angular dispersion of intensity gradients is a good indicator of SNR, and is independent of mean illumination. The new SNR estimation algorithm determines angular dispersion from the x and y gradient components about each pixel. SNR is then determined from a look-up table. The algorithm makes explicit use of both the magnitudes and directions of intensity gradients by computing the magnitude-weighted angular dispersion of the gradients over a local neighbourhood.

The algorithm is shown to be both simple to implement and to perform well, especially for values of SNR above about 0.5.

Keywords: edge detection, gradient, edge consistency, gradient angular dispersion, signal-to-noise ratio, edge direction, gradient direction, thresholding,

1 Introduction

Feature detection frequently consists of first convolving the image with one or more filters to enhance the feature of interest (if present) and then comparing the feature-enhanced image to a threshold. Ideally, thresholds are based on the signal to noise ratio (SNR) defined as the ratio of the filter's response to a noise-free image and its response to the noise alone [2,3,4]. In edge detection, in which the feature of interest is an intensity gradient of high magnitude, determination of adequate thresholds has proved problematical because the gradient-based filters used for edge enhancement are generally simplistic two-dimensional extensions of one-dimensional operators. Many approaches use a threshold based on an estimate of the noise statistics made over a patch of the image known to be devoid of edges, and assume that these statistics are both representative and stationary [6,5]. This is clearly unrealistic for all but the simplest scenes ('blocks world', machine parts, etc.) since very frequently regions in an image are distinguishable by both intensity and texture differences. Non-regular texture can be considered to be noise at the spatial scale of the edge detector, even though it provides strong cues for object identification when a suitable non-linear detector is used.

Another approach is to estimate the noise level in the direction perpendicular to the putative edge gradient. At each pixel, oriented edge enhancement filters $F_q(x,y), 0 \leq q < 2\pi/\Delta$ spaced at intervals of Δ in orientation, and corresponding noise estimation filters $H_q(x,y)$ must be used. Each $H_q(x,y)$ is designed to have minimal response $R_{q,H}$ to an edge which elicits a maximum response $R_{q,F}$ from $F_q(x,y)$, for all q [2]. The noise level is deemed to be approximated by that noise estimate for which $R_{q,F}/R_{q,H}$ is maximum, since the edge is then assumed to be perpendicular to the orientation of the noise estimator. This strategy suffers from the necessity of determining both edge response and noise level estimates at a large number of orientations, leading to significant computational burden. Another problem arises from the 'pick the biggest' strategy employed because this extremization in-

*This research was supported by the National Sciences and Engineering Research Council, grant no. 0041796

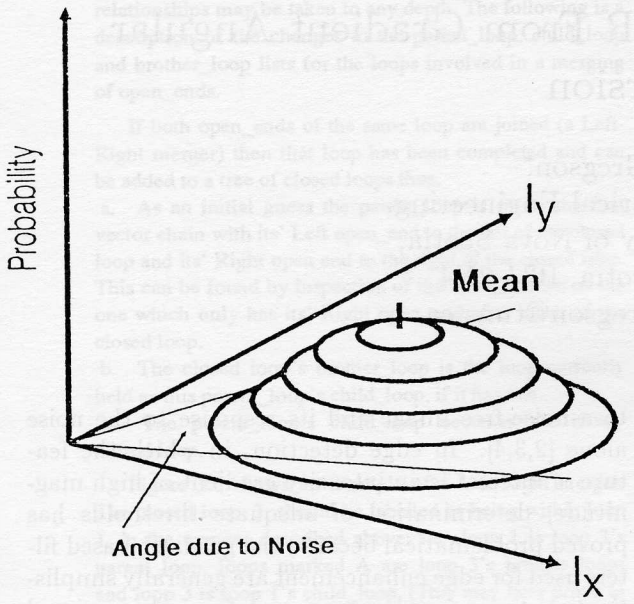

Figure 1: The angle subtended at the origin of gradient probability space is a function of the location of the mean and size of the distribution.

creases estimation error.

In the current approach, gradient-based filters are used to estimate the gradient magnitude and direction at each pixel. Since the gradient estimates are linear combinations of the pixel values, image noise will give rise to noise in the estimates. This noise is manifested as variations in both gradient magnitude and gradient direction. Consider a neighbourhood N of sufficient size to contain m^2 non-overlapping gradient estimation filters, centered on the pixel of interest and containing a uniform intensity gradient corrupted by additive, normally distributed noise. The x and y components of the gradient estimates have normal probability density functions (p.d.f.) with identical non-zero means and identical variances $k2\sigma^2$ due to the image noise variance and the noise gains k of the estimation filters [8]. The joint p.d.f. is then bivariate normal with non-zero mean. The intersection of a threshold (a level surface) with this p.d.f. describes a circle with radius dependent on the value of the threshold and $k^2\sigma^2$ of the bivariate p.d.f. The center of the circle is positioned at the bivariate mean (Figure 1), and all points in the circle represent gradients with probability of occurence greater than the threshold.

As the magnitude of the mean gradient increases, the circle moves farther from the origin. The angle ϕ subtended at the origin by the circle then decreases, provided that the origin lies outside the circle. If the threshold is selected such that the cumulative probability inside the circle is constant, and the image noise magnitude σ is increased, the radius of the circle increases and ϕ increases. The threshold value used in the present work results in ϕ being twice the weighted angular dispersion of the gradient vectors, where the weight is the vector magnitude.

Since the magnitude of the mean gradient is a measure of the filter's response to a gradient and circle radius is a measure of noise response, the ratio of these quantities is a function of the SNR.

2 Algorithm Development

The angle ϕ is unknown since the p.d.f. of the gradient components in N cannot be determined from gradient measurements as N contains only one outcome of the noise process at each pixel. However, a reasonable estimate of ϕ may be made from measurements in N if there is a sufficient number of pixels in N and if the noise is independent and identically distributed (i.i.d.). The SNR which is most likely to have given rise to the observed ϕ is then the best estimate of the SNR in N. Analysis of the distribution of gradients in N suggests an algorithm for estimating ϕ and hence the SNR.

2.1 Gradient Distribution

The two kernels S_X, S_Y of the Sobel operator are used to estimate g_X and g_Y, the x and y components of the gradient vector $\mathbf{G}_i$ at pixel i. (Figure 2). Sobel kernels are used because they result in minimum errors in estimating both magnitude and angle of the gradient [7].

The estimated gradients are eight times the actual gradients due to the gain of the filter, but this is a systematic error and so is correctable. The kernels are applied to every third pixel in the x and y directions in N so as to prevent overlap of the regions of support of adjacent kernels. This is necessary to prevent false correlations in the noise responses (arising from shared pixels) which would be indistinguishable from signal response, yielding erroneously high SNR estimates.

Assuming that the image I consists of a noise free image $\hat{I}$ corrupted with additive, i.i.d. $N(0, \sigma^2)$ noise ζ,

$$I(x, y) = \hat{I}(x, y) + \zeta(x, y) \quad (1)$$

-1	0	1
-2	0	2
-1	0	1

1	2	1
0	0	0
-1	-2	-1

Figure 2: Sobel kernels used for estimating the x and y components of the intensity gradient. The left kernel S_X is used to estimate x components, the right kernel S_Y for y components.

it follows that the means $M_{X,i}, M_{Y,i}$ of the x and y gradient components taken over neighbourhood N_i, centered on pixel i, are

$$\begin{aligned} M_{X,i} &= \frac{1}{m^2}\sum g_X \\ &\approx E_{N_i}\{g_X\} \\ M_{Y,i} &= \frac{1}{m^2}\sum g_Y \\ &\approx E_{N_i}\{g_Y\} \end{aligned} \quad (2)$$

for large N_i, and where the estimated gradient components g_X, g_Y in N_i are generated from m^2 non-overlapping kernels. Expectation over N_i is denoted by $E_{N_i}\{*\}$. Further, since the filters form linear combinations of image pixels, the variances $\sigma^2_{X,i}$ and $\sigma^2_{Y,i}$ of the gradient components in N_i are [8]

$$\begin{aligned} \sigma^2_{X,i} &= \sigma^2_{Y,i} \\ &= \sigma^2 \sum S_X^2 \\ &= 12\sigma^2 \end{aligned} \quad (3)$$

where the summation is taken over all weights in the Sobel kernel used to estimate the x component of the gradient.

The joint density function for g_X and g_Y is then

$$f_{G_{X,i},G_{Y,i}}(g_X, g_Y) = \frac{1}{24\pi\sigma^2} e^{-((g_X - M_{X,i})^2 + (g_Y - M_{Y,i})^2)/24\sigma^2} \quad (4)$$

This may be written in polar coordinates

$$f_{R_i,\Theta_i}(r,\theta) = \frac{1}{24\pi\sigma^2} r e^{-(r^2 + M^2_{R,i} - 2rM_{R,i}\cos(\theta - \theta_{R,i}))/24\sigma^2} \quad (5)$$

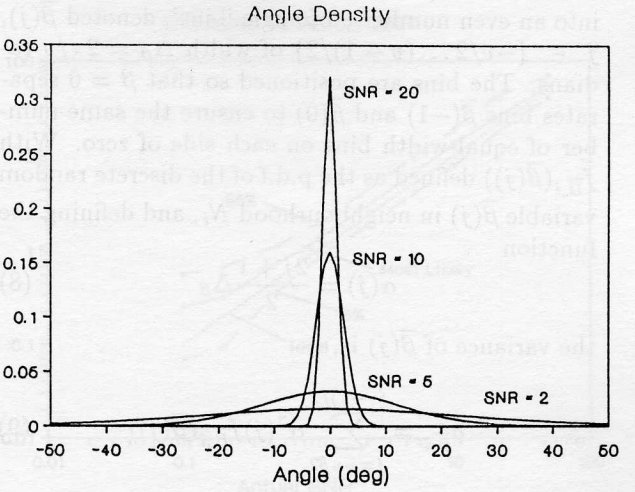

Figure 3: p.d.f. of gradient angle for several SNRs.

where

$$M_{R,i} = E\left\{\frac{1}{m^2}\left[\left(\sum M_{X,i}\right)^2 + \left(\sum M_{Y,i}\right)^2\right]^{1/2}\right\}$$

$$\theta_{R,i} = E\left\{\arctan\frac{\sum M_{Y,i}}{\sum M_{X,i}}\right\} \quad (6)$$

are the expected mean gradient magnitude and expected mean gradient direction over region N_i. The coordinate system is rotated by $\theta_{R,i}$, and a new angle random variable β is defined as $\beta = \theta - \theta_{R,i}$. Completing the square in the exponent, and integrating with respect to r yields the p.d.f. of β in N_i

$$\begin{aligned} f_{B,i}(\beta) = &\frac{1}{2\pi} e^{-M^2_{R,i}/24\sigma^2} \\ &+ \frac{M_{R,i}\cos\beta}{4\sigma\sqrt{6\pi}} e^{-M^2_{R,i} sin^2\beta/24\sigma^2} \\ &\left[1 + \mathrm{erf}\left(\frac{M_{R,i}\cos\beta}{2\sqrt{6}\sigma}\right)\right] \end{aligned} \quad (7)$$

The p.d.f. of β is only a function of the SNR, $M_{R,i}/\sigma$, as expected. At very low SNR, $f_{B,i}(\beta)$ is dominated by the first term in (7), resulting in a uniform distribution for an SNR of zero. At high SNR, the second term dominates, resulting in a well-localized distribution about the direction of the resultant (Figure 3).

2.2 Estimating Angular Deviation

In order to implement the algorithm, β is quantized into an even number v of angle 'bins', denoted $\overline{\beta}(j)$, $j \in \{-v/2 \ldots (v-1)/2\}$ of width $\Delta_\beta = 2\pi/v$ radians. The bins are positioned so that $\beta = 0$ separates bins $\overline{\beta}(-1)$ and $\overline{\beta}(0)$ to ensure the same number of equal-width bins on each side of zero. With $f_{\overline{B},i}(\overline{\beta}(j))$ defined as the p.d.f of the discrete random variable $\overline{\beta}(j)$ in neighbourhood N_i, and defining the function

$$\alpha(j) = \frac{2j+1}{2}\Delta_\beta \tag{8}$$

the variance of $\overline{\beta}(j)$ is,

$$\sigma^2_{\overline{\beta},i} = \sum_{j=-v/2}^{(v-1)/2} \alpha^2(j) f_{\overline{B},i}(\overline{\beta}(j)) \tag{9}$$

But

$$f_{\overline{B},i}(\overline{\beta}(j)) = E\left\{ \frac{\sum_{\overline{\beta}(j)} |\mathbf{G}|}{\sum |\mathbf{G}|} \right\} \tag{10}$$

where the numerator summation is taken over the magnitudes of all gradients in N_i having directions quantized to $\overline{\beta}(j)$, and the denominator is the sum of all gradient magnitudes in N_i. Thus the term inside the expectation in (10) is the total gradient magnitude response for all gradients at angle $\overline{\beta}(j)$ in N_i normalized by the total magnitude response, for a single experiment. By taking the expectation, the value of $f_{\overline{B},i}(\overline{\beta}(j))$ for each $\overline{\beta}(j)$ is found, since the expectation implies that an infinite number of such experiments are performed.

The expectation of a ratio in equation (10) may be simplified with a Taylor's series expansion. Expanding $E\{a/b\}$ about the point a_o/b_o, where $a_o = E\{a\}$ and $b_o = E\{b\}$, using terms of up to second order,

$$\begin{aligned} E\left\{\frac{a}{b}\right\} \approx & \frac{a_o}{b_o} \\ & - E\left\{ \frac{1}{b_o^2}(a-a_o)(b-b_o) \right\} \\ & + E\left\{ \frac{a_o}{b_o^3}(b-bo)^2 \right\} \end{aligned} \tag{11}$$

it is seen that the expectation of a ratio is well approximated by the ratio of expectations provided that the covariance of the numerator with the denominator, and the denominator coefficient of variation, are both suitably small. These terms are largest in (10) when the SNR is zero since this minimizes denominator. Detailed analysis indicates that even in this case only the first term of the expansion is necessary [4] since the denominator is positive and always greater than the numerator. Equation(10) is then

$$f_{\overline{B},i}(\overline{\beta}(j)) \approx \frac{E\left\{\sum_{\overline{\beta}(j)} |\mathbf{G}|\right\}}{E\left\{\sum |\mathbf{G}|\right\}} \tag{12}$$

Invoking the central limit theorem, the numerator and denominator expectations may be dropped if N_i is sufficiently large [8]. Also,

$$\sum_\gamma |\mathbf{G}| = \left| \sum_\gamma \mathbf{G} \right| \tag{13}$$

and the resultant of the gradient vectors is

$$\left|\sum \mathbf{G}\right| = \sum_{j=-v/2}^{(v-1)/2} \sum_{\overline{\beta}(j)} |\mathbf{G}| \cos\left(\alpha(j)\right) \tag{14}$$

where the inner summation is taken over all gradient magnitudes in N_i having quantized gradient directions of $\overline{\beta}(j)$. Using the first two terms in the power series approximation of $\cos\gamma$,

$$\cos\gamma \approx 1 - \frac{\gamma^2}{2} \tag{15}$$

it then follows that the angular deviation may be approximated by the angular dispersion s_i in N_i,

$$\begin{aligned} s_i &= \left[2 \sum_{j=-v/2}^{(v-1)/2} \left(1 - \cos\left(\alpha(j)\right)\right) \frac{\left|\sum_{\overline{\beta}(j)} \mathbf{G}\right|}{\sum |\mathbf{G}|} \right]^{\frac{1}{2}} \\ &= \left[2 - 2\frac{|\sum \mathbf{G}|}{\sum |\mathbf{G}|} \right]^{\frac{1}{2}} \\ &\approx \sigma_{\overline{\beta},i} \end{aligned} \tag{16}$$

Equation(16) provides a means of estimating the weighted angular deviation in N_i. From this, the estimated SNR may be determined numerically from equations (7) and (9) (Figure 4).

A previously stated, the angle ϕ_i is a measure of the 'angular width' of the p.d.f. of β. A suitable definition for ϕ is

$$\phi_i = s_i \tag{17}$$

the weighted angular dispersion of the gradient vectors in N_i.

3 Experimental Results

An experiment was conducted to evaluate the performance of the algorithm at estimating image SNR

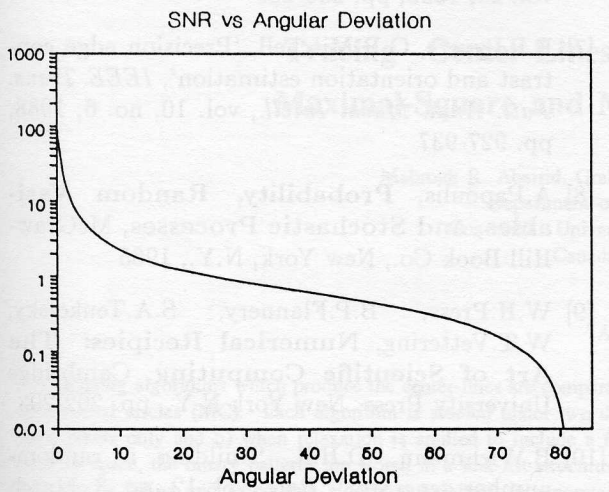

Figure 4: Relationship between weighted angular deviation (in degrees) and SNR.

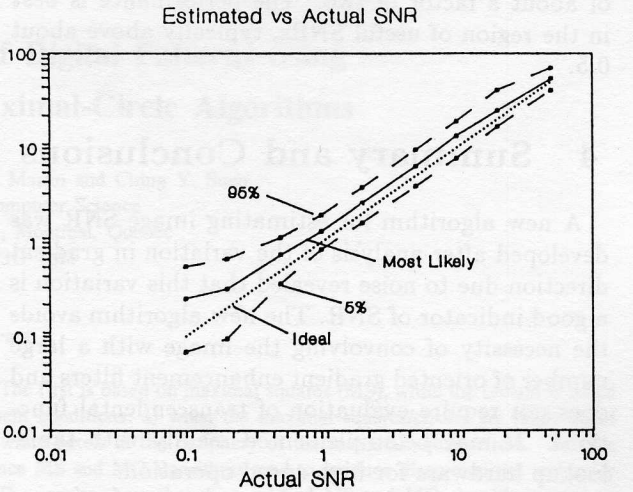

Figure 5: Estimated versus actual SNR. Solid curve indicates most likely estimate, dashed curves indicate 5-th and 95-th percentiles. The dotted curve indicates actual SNR.

at different values of actual SNR. Images of intensity ramps with uniformly distributed, random orientations between zero and 45 degrees, with prescribed slopes were simulated. Additive, independent, identically distributed, normal random noise was added to each image. The mean was zero and the standard deviation was controlled to achieve the desired signal to noise ratio. The random noise source had a sequence length of about 7×10^{12} random deviates [10]. Conversion to a normal distribution used a theoretically correct algorithm [9] instead of an approximation based on the central limit theorem.

Tha angular deviation was estimated over a 9×9 pixel neighbourhood N. This permitted the use of nine non-overlapping Sobel kernels for gradient estimation. The neighbourhood was limited to nine gradient estimates because the spatial support of any noise estimation technique must be as small as possible if it is to be useful in practical computer vision systems. For tasks in which noise estimation is an issue, the visual world is complex and rich (natural scenes, biological and medical images, etc.) and so assumptions of statistical stationarity and sparseness of edges are inappropriate. Large spatial supports imply the former assumption and require the second one.

Conversion from the estimate of ϕ to the estimate of SNR was accomplished by linear interpolation from a table of values determined by numerical evaluation of equations (7) and (9) (Figure 4). The approximation used for the error function is accurate to 1.5×10^{-7} [1], and the numerical integration involved in (9) was performed in steps of one degree. The curve is based on 65 SNR values following an approximately exponential sequence from 0.01 to 100.

The algorithm was tested with 10,000 iterations at each of nine SNR values. The most likely SNR estimate along with the 5-th percentile and 95-th percentile estimates were found for each actual SNR (Figure 5).

It can be seen from the figure that the estimate is very close to the ideal estimate at all but the lowest SNRs. The SNR estimate is high at low SNR because only nine gradients are used to estimate the gradient deviation. With a random noise field only, the resultant of the nine gradients used will have a non-zero resultant which the algorithm cannot distinguish from a signal. The estimated SNR is greater than zero even though the actual SNR is zero in this case. As the signal (intensity slope) is increased, the ratio of signal-based gradient resultant to noise-based gradient resultant increases, thereby swamping this effect. The 5-th and 95-th percentiles diverge at low SNR because the gradient resultant component due to noise is uncorrelated with the signal resultant component, resulting in larger apparent signal variation at low SNR than at high SNR.

The 5-th to 95-th percentile points at any actual SNR above about 0.5 coorespond to an SNR range of about a factor of two. The performance is best in the region of useful SNRs, typically above about 0.5.

4 Summary and Conclusions

A new algorithm for estimating image SNR was developed after analysis of the variation in gradient direction due to noise revealed that this variation is a good indicator of SNR. The new algorithm avoids the necessity of convolving the image with a large number of oriented gradient enhancement filters and does not require evaluation of transcendental functions. It may be implemented readily with table-lookup hardware for high-speed operation.

The new algorithm provides a rapid means of estimating local image SNR, permitting adaptive edge detection thresholds. As this is necessary to make most effective use of an imaging system's limited dynamic range, the algorithm should prove useful for practical computer vision systems.

Areas for further work include performance of the algorithm with curved edges and corners, and the possibility of reducing the size of N by permitting overlapping gradient estimation kernels.

References

[1] M.Abramowitz, I.A.Stegun, **Handbook of Mathematical Functions**, Dover Publications, Inc., New York, N.Y., 1964, pp. 299 §7.1.26

[2] J.Canny, 'A computational approach to edge detection', *IEEE Trans. Patt. Anal. Mach. Intell.*, PAMI-8 no. 6, 1986, pp. 679-697

[3] W.Frei, C.C.Chen, 'Fast boundary detection: a generalization and a new algorithm', *IEEE Trans. Computers*, vol. 26, no. 2, 1977, pp. 988-998

[4] P.II.Gregson, 'Using angular dispersion of gradient direction for detecting edge ribbons', to be submitted to *IEEE Trans. Patt. Anal. Mach. Intell.*

[5] W.E.L.Grimson, T.Pavlidis, 'Discontinuity detection for visual surface reconstruction', *Comp. Vis., Graphics, Im. Proc.*, vol. 30, 1985, pp. 316-330

[6] J.-S.Lee, 'Digital image smoothing and the sigma filter', *Comp. Vis., Graphics, Im. Proc.*, vol. 24, 1983, pp. 255-269

[7] E.P.Lyvers, O.R.Mitchell, 'Precision edge contrast and orientation estimation', *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 10, no. 6, 1988, pp. 927-937

[8] A.Papoulis, **Probability, Random Variables, and Stochastic Processes**, McGraw-Hill Book Co., New York, N.Y., 1965

[9] W.H.Press, B.P.Flannery, S.A.Teukolsky, W.T.Vettering, **Numerical Recipies: The Art of Scientific Computing**, Cambridge University Press, New York N.Y., pp. 202-203

[10] B.Wichmann, D.Hill, 'Building a random-number generator', *Byte*, vol. 12, no. 3, March 1987, pp. 127-128