SUSAN – A New Approach to Low Level Image Processing

Technical Report TR95SMS1c
(Shorter versions have now been published in IJCV [66] and ICPR96 [64],
and the relevant patent is [60]. To reference this research, cite [66])

1995

S.M. Smith
Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB),
Department of Clinical Neurology, Oxford University, Oxford, UK
(Previously in Computer Vision and Image Processing Group, DRA Chertsey, DERA, UK)
steve@fmrib.ox.ac.uk
www.fmrib.ox.ac.uk/~steve

J.M. Brady
Department of Engineering Science, Oxford University, Oxford, UK

Abstract

This paper describes a new approach to low level image processing; in particular, edge and corner detection and structure preserving noise reduction.

Non-linear filtering is used to define which parts of the image are closely related to each individual pixel; each pixel has associated with it a local image region which is of similar brightness to that pixel. The new feature detectors are based on the minimization of this local image region, and the noise reduction method uses this region as the smoothing neighbourhood. The resulting methods are accurate, noise resistant and fast.

Details of the new feature detectors and of the new noise reduction method are described, along with test results.

Keywords: edge detection, feature detection, univalue areas, noise reduction, smoothing.

1 Introduction

This paper describes an entirely new approach to low level image processing, specifically, edge detection (one dimensional feature detection), "corner" detection (two dimensional feature detection, including, therefore, corners, junctions, etc.) and structure preserving noise reduction. The new approach represents a significant departure from feature extraction and noise reduction methods previously developed.

The paper begins with an explanation of the SUSAN feature detection principle, and continues with details of the applications of it, including reviews of relevant past research and results of testing the applications. Because the research is based on fundamentally non-linear filtering, the approach to theoretical justification is necessarily different from that traditionally applied, for example, to "optimal" edge filtering.

2 The SUSAN Principle for Feature Detection

The SUSAN principle is now introduced, from which the research described in this paper is derived. Consider Figure 1, showing a dark rectangle on a white background. A circular mask (having a centre pixel which shall be known as the "nucleus") is shown at five image positions. If the brightness of each pixel within a mask is compared with the brightness of that mask's nucleus then an area of the mask can be defined which has the same (or similar) brightness as the nucleus. This area of the mask shall be known as the "USAN", an acronym standing for "Univalue Segment Assimilating Nucleus". In Figure 2 each mask from Figure 1 is depicted with its USAN shown in white.

This concept of each image point having associated with it a local area of similar brightness is the basis for the SUSAN principle. The local area or USAN contains much information about the structure of the image. It is effectively region finding on a small scale. From the size, centroid and second moments of the USAN two dimensional features and edges can be detected. This approach to feature detection has many differences to the well known methods, the most obvious being that no image derivatives are used and that no noise reduction is needed.

The area of an USAN conveys the most important information about the structure of the image in the region around any point in question. As can be seen from Figures 1 and 2, the USAN area is at a maximum when the nucleus lies in a flat region of the image surface, it falls to half of this maximum very near a straight edge, and falls even further when inside a corner. It is this property of the USAN's area which is used as the main determinant of the presence of edges and two dimensional features. Consider now Figure 3, where a small part of a test image has been processed to give USAN area as output.[1] Each point in the input image is used as the nucleus of a small circular mask, and the associated USAN is found. The area of the USAN is used in the three dimensional plot shown. The USAN area falls as an edge is approached (reaching a minimum at the exact position of the edge), and near corners it falls further, giving local minima in USAN area at the exact positions of image corners. Figure 4 shows a small part of a real noisy image, and the resulting output from USAN area processing. (The variation in brightness within the "flat regions" is of the order of 15 – out of 256 – greyscale levels.) Again there is edge and corner enhancement, with the noise having no visible effect on the final plot.

Consideration of the above arguments and observation of the examples and results shown in Figures 1, 2, 3 and 4 lead directly to formulation of the SUSAN principle:

> An image processed to give as output inverted USAN area has edges and two dimensional features strongly enhanced, with the two dimensional features more strongly enhanced than edges.

This gives rise to the acronym SUSAN (Smallest Univalue Segment Assimilating Nucleus). Mathematical analyses of the principle are given after the algorithms have been described in detail.

The fact that SUSAN edge and corner enhancement uses no image derivatives explains why the performance in the presence of noise is good. The integrating effect of the principle, together with its non-linear response, give strong noise rejection. This can be understood simply if an input signal with identically

[1] Note that the scale on the vertical axis is inverted.

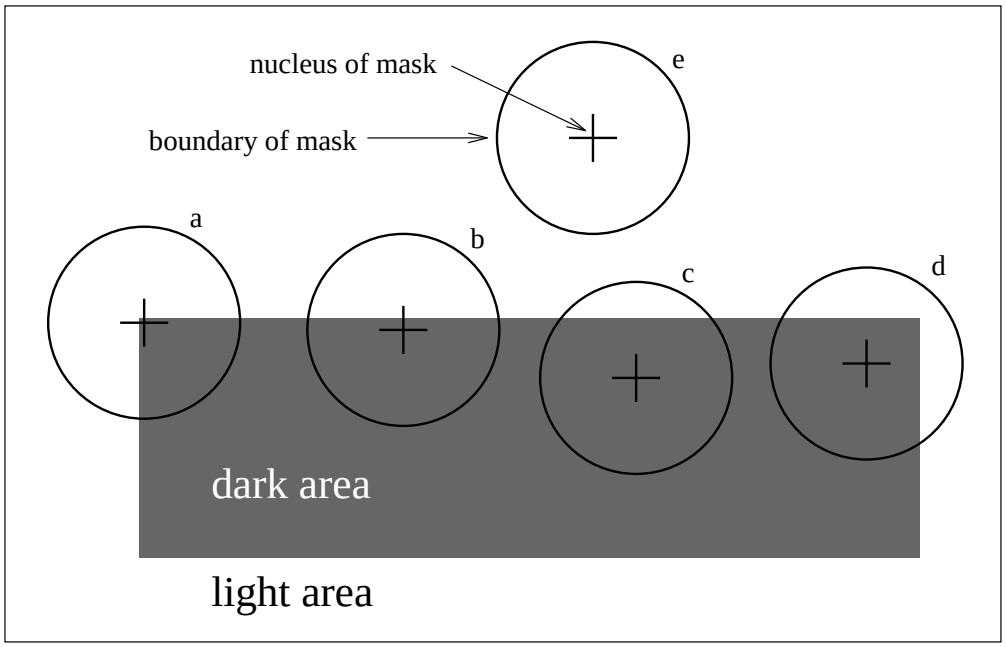

Figure 1: Four circular masks at different places on a simple image.

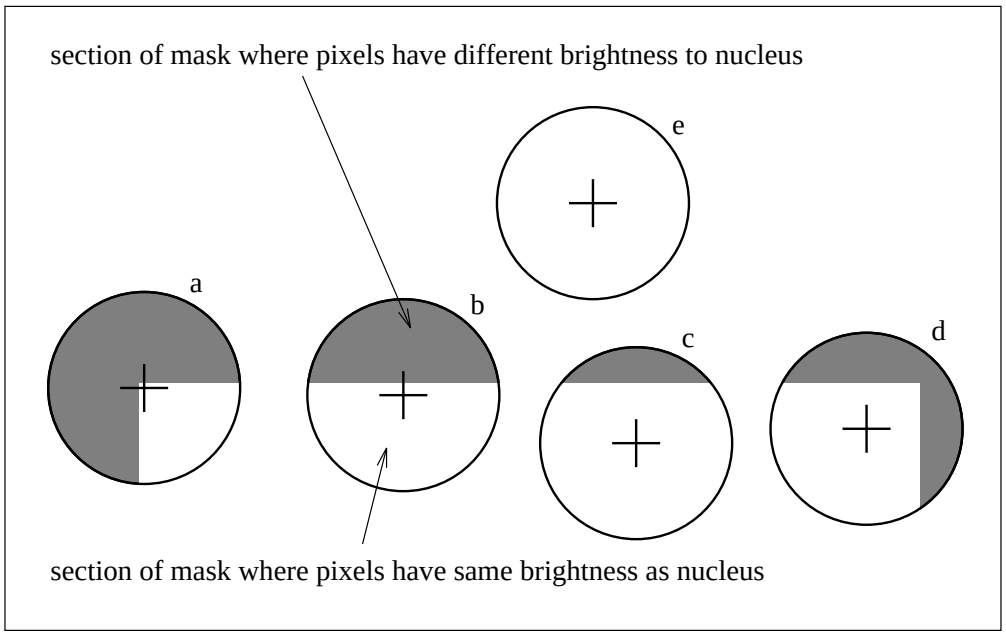

Figure 2: Four circular masks with similarity colouring; USANs are shown as the white parts of the masks.

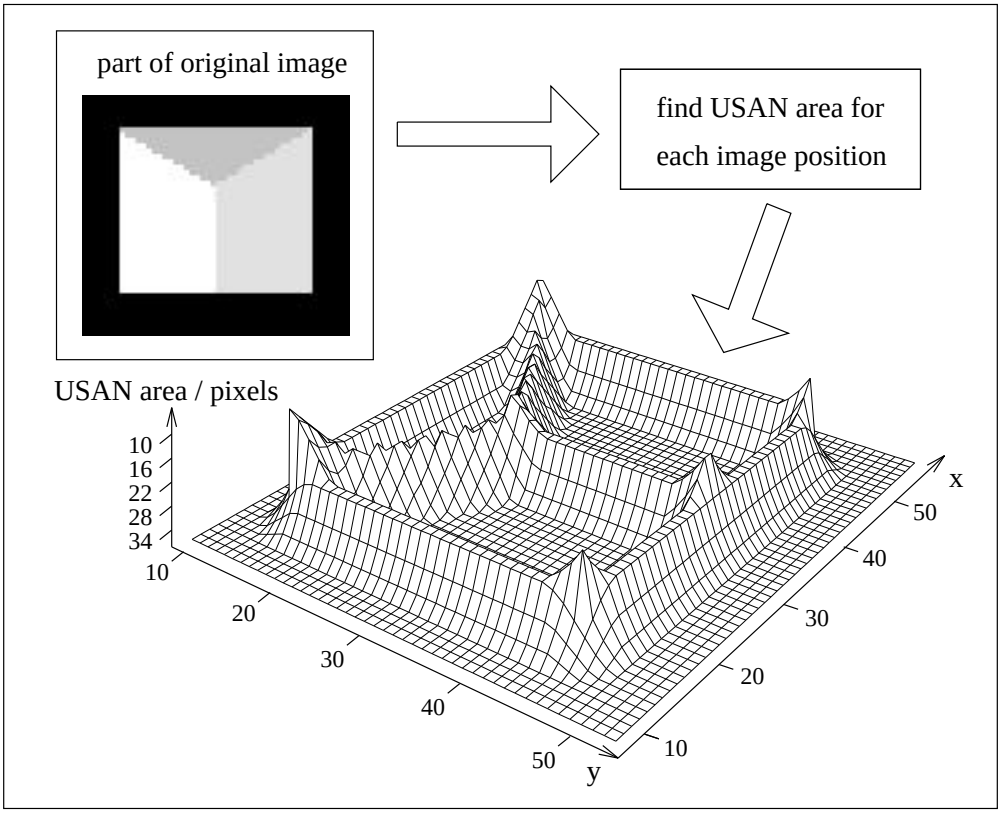

Figure 3: A three dimensional plot of USAN area given a small part of a test image, showing edge and corner enhancement.

independently distributed Gaussian noise is considered. As long as the noise is small enough for the USAN function (see Figure 5) to contain each "similar" value, the noise is ignored. The integration of individual values in the calculation of areas further reduces the effect of noise. Another strength of the SUSAN edge detector is that the use of controlling parameters is much simpler and less arbitrary (and therefore easier to automate) than with most other edge detection algorithms.

The SUSAN noise reduction algorithm is related to the SUSAN principle in that the USAN is used to choose the best local smoothing neighbourhood.

3 Criteria for Determining the Quality of Feature Detectors

In this section the desired qualities of feature detectors are explained. The criteria eventually given have a slightly different emphasis from those previously used.

In [9] three criteria for edge detection are given. These have been used in similar form in a good deal of vision research. They are:

1. Good detection. There should be a minimum number of false negatives and false positives.[2]

[2]Canny develops Criterion 1 to mean that the edge enhancing filter should maximize the signal-to-noise ratio. This makes sense in the case of a linear filter which has features thresholded at a later stage than the enhancement, and Canny formulates a signal-to-noise functional which expresses this requirement. However, the SUSAN principle is fundamentally non-linear. (This can be simplistically interpreted as performing thresholding at an earlier stage.) In fact, it is clear from Figure 4 that the term "enhancement" is a little weak, given the obvious signal-to-noise ratio. (The plot of the operator output shows no visible noise.) Thus the criterion is most appropriate as it stands.

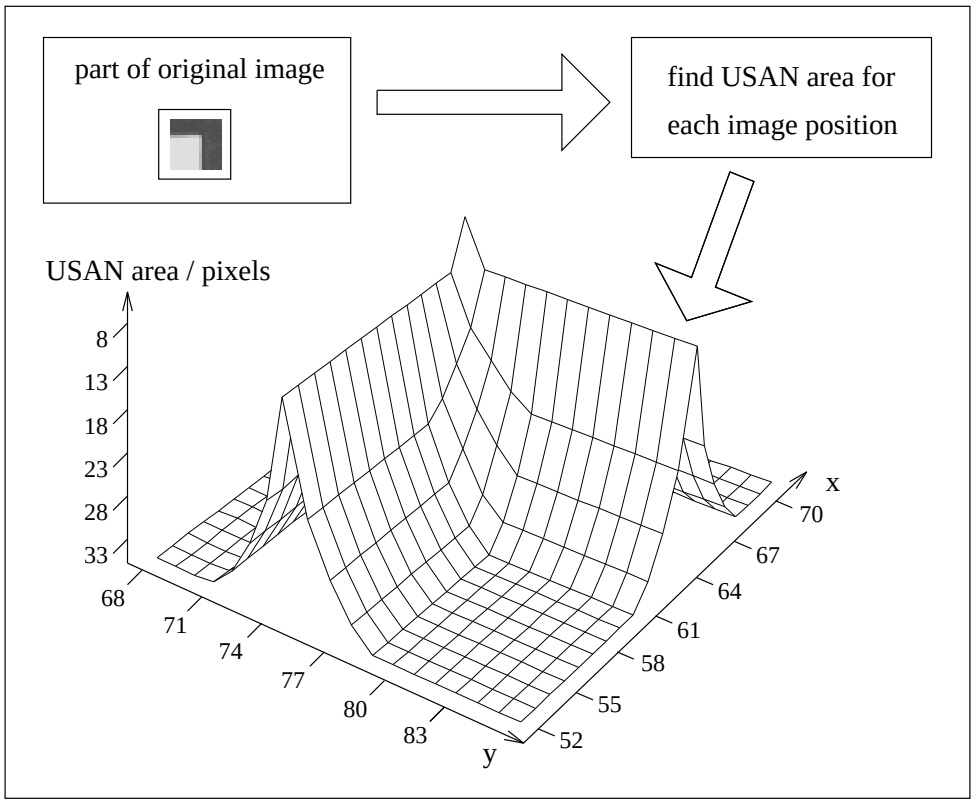

Figure 4: A three dimensional plot of USAN area given a small part of a real noisy image, showing edge and corner enhancement.

2. Good localization. The edge location must be reported as close as possible to the correct position.

3. Only one response to a single edge.

These criteria are equally appropriate for two dimensional feature detection, except for the complication that arises from using the term "correct position" in Criterion 2. There are many mathematical descriptions of the two dimensional image structure which define a "corner" or more general feature, leading to a lack of one agreed definition of the exact location of a feature. Thus in the case of two dimensional features, measuring fine localization error is not appropriate. However, it will be seen later that the error in localizing two dimensional features is often several pixels with some existing algorithms, greater even than the expected variation in positions determined by hand.

The purpose of developing new feature detectors was that they should be appropriate to being used as part of a *real time* system using *real* image sequences. Therefore it has become apparent that for the purpose of this work, one final criterion is important;

4. Speed. The algorithm should be fast enough to be usable in the final image processing system.

Obviously this criterion must not be allowed to dominate the development of a feature detector to an extent that the quality of the results is adversely affected. However a fast algorithm performing well in the other criteria is more desirable than a slow one.

Canny has formulated his criteria mathematically (in [8] and [9]). He uses the criteria functionals to derive "optimized" edge filters for each image type. Criteria 1 and 2 are clearly the most important ones in terms of "optimizing" the filter. Later analysis shows how the SUSAN brightness comparison function has

been optimized to give the lowest number of false negatives and false positives. In the case of localization, the spatial domain part of the SUSAN detectors does not introduce any new concepts, and, following Canny and others, either a Gaussian or square distribution may be used (see later). Criterion 3 has not been of relevance in this work, as multiple responses have not been a problem at all. Finally, as will be seen, the SUSAN feature detectors are extremely computationally efficient.

In [71] a similar set of criteria are defined, not for optimizing filters, but for comparing the outputs of different edge detectors. There has been very little work on the area of objective quantitative tests for feature detectors. In [71] four quantitative criteria are defined, those being proportional to the number of false negatives, the number of false positives, the number of multiple detections and the number of incorrectly localized pixels. A linear combination of these measures is then used to obtain one quantity, the "failure measure" or "FM". The weights used to create this weighted sum are found according to a synthesis of some simple minimization rules with the constraint that the resulting edges found using this sum correspond to the best edges judged by eye. The weights will also depend on the proposed use of the reported edges.

The authors are of the opinion that this method of calculating the quality of an edge filter is not very meaningful, as the measures defined in [71] are found before post-processing takes place. Non-maximum suppression (a common enough operation) will normally eliminate multiple detections of a single edge completely. Binary thinning (particularly when continuously taking the initial response into account, as described in [63]) will greatly reduce the number of false negatives and false positives. All that remains is the localization error. Unfortunately, in the scheme described, the importance given to this error is between three and ten times less than that given to the number of false negatives or false positives.

However, in the absence of a better comparative method in the literature, quantitative tests have been carried out using the FM scheme, comparing the output from the first stage of the SUSAN edge detector with the four filters tried in [71], namely, Sobel, Prewitt, Roberts and the "three-point energy model"[3] (see [67], [50], [51] and [72]). A test image identical to the image described was created (i.e. a vertical step edge with added Gaussian noise) giving a signal to noise ratio of 14.79. The performance of the initial stage of the SUSAN detector was evaluated for the four criteria defined; the result was that the SUSAN edge detector gave FMs which were better than those obtained with the other four detectors, independent of the set of weights used. Note that the algorithms compared here with the first stage of SUSAN represent the first (or enhancement) stage of edge detection, rather than complete edge detectors, as this is appropriate to the comparative method developed in [71].

It is worth noting here that in the absence of multiple features the SUSAN principle bypasses the "uncertainty principle" which applies to most feature detectors (and most obviously the Gaussian based ones) with respect to Canny's first and second criteria. This well understood problem means that the better the detection quality (including noise suppression) then the worse the localization of the detected feature. In the case of the SUSAN feature detectors, however, it is clear that the localization of the features is independent of the mask size, as long as no other features are found within the mask's region. Thus there is no conflict here between the two most "important" criteria.

4 The SUSAN Edge Detector

The details of the SUSAN edge finding algorithm are given, followed by an analysis of the algorithm's validity. Finally, examples of the output of the edge detector are presented and discussed. Firstly, however, a brief review of existing approaches is given. (For a longer review of edge finding, see [63].)

[3] This edge detector combines the outputs of two filters to attempt to find both step edges and ridge/roof edges. It basically performs a simple test to find maxima in the first or second derivatives.

4.1 Review

There has been an abundance of work on different approaches to the detection of one dimensional features in images. The wide interest is due to the large number of vision applications which use edges[4] and lines as primitives, to achieve higher level goals.

Some of the earliest methods of enhancing edges in images used small convolution masks to approximate the first derivative of the image brightness function, thus enhancing edges; e.g., see [50] and [67]. These filters give very little control over smoothing and edge localization.

In [34] Marr and Hildreth proposed the use of zero crossings of the Laplacian of a Gaussian (LoG). Contours produced using the LoG filter have the property, convenient for some purposes, of being closed. However, connectivity at junctions is poor, and corners are rounded. Also, the use of non-directional derivatives means that the edge response parallel to an edge is always measured (as well as the expected response perpendicular to it), reducing the signal to noise ratio. The use of directional first and second derivatives improves on this. Finally, the LoG filter gives no indication of edge direction, which may be needed by higher level processes.

In [8] Canny described what has since become one of the most widely used edge finding algorithms. The first step taken is the definition of criteria which an edge detector must satisfy (see Section 3). These criteria are then developed quantitatively into a total error cost function. Variational calculus is applied to this cost function to find an "optimal" linear operator for convolution with the image. The optimal filter is shown to be a very close approximation to the first derivative of a Gaussian.[5] Non-maximum suppression in a direction perpendicular to the edge is applied, to retain maxima in the image gradient. Finally, weak edges are removed using thresholding. The thresholding is applied with hysteresis. Edge contours are processed as complete units; two thresholds are defined, and if a contour being tracked has gradient magnitude above the higher threshold then it is still "allowed" to be marked as an edge at those parts where the strength falls below this threshold, as long as it does not go below the lower value. This reduces streaking in the output edges.

The Gaussian convolution can be performed quickly because it is separable and a close approximation to it can be implemented recursively. However, the hysteresis stage slows the overall algorithm down considerably. While the Canny edge finder gives stable results, edge connectivity at junctions is poor,[6] and corners are rounded, as with the LoG filter. The scale of the Gaussian determines the amount of noise reduction; the larger the Gaussian the larger the smoothing effect. However, as expected, the larger the scale of the Gaussian, the less accurate is the localization of the edge.

Canny also investigated the synthesis of results found at different scales; in some cases the synthesis improved the final output, and in some cases it was no better than direct superposition of the results from different scales.

Finally, Canny investigated the use of "directional operators". Here several masks of different orientation are used with the Gaussian scale larger along the direction parallel to the edge than the scale perpendicular to it. This improves both the localization and the reliability of detection of straight edges; the idea does not

[4]Much research in this field has assumed that the only one dimensional features of interest are step edges. However, there exist many other types of feature. These include lines (ridges in the image surface), ramp ends and roof edges; see Figure 9 for examples of these. There are three main reasons for the concentration on step edges. The first is that they are the most common type of one dimensional change. The second is that edges containing a step component are the most well localized one dimensional features, that is, they are formed by a "first order" change. The third reason for working only with step edges is that some proposed edge finders (such as Canny's) are easily extended to finding other types of change once the theory for step edges has been completed. Thus many detectors have been developed using rigorous derivations of optimal algorithms using various criteria based on the model of the ideal step edge.

[5]The problem with using image derivatives is that differentiation enhances noise as well as edge structure, so most edge detectors include a noise reduction stage. Thus the use of the derivative of a Gaussian enables differentiation to take place at the same time as the smoothing; this is allowable, as the two processes commute (exactly in the continuous case, and approximately in the discrete case). The problem of noise enhancement is even worse when differentiation is performed twice.

[6]Some higher level algorithms use Canny's method *because* of this characteristic, as they work better with simple unconnected edges. However, achieving full connectivity at junctions is clearly a worthwhile goal as it correctly represents the scene. In [32] Li et. al. suggest heuristic extensions to the Canny algorithm to enable the joining of open contour ends with nearby contours. This however produces some false edge extensions.

work well on edges of high curvature.

In [14] and [56] similar analytical approaches to that of Canny are taken, resulting in efficient algorithms which have exact recursive implementations. These algorithms give results which are very similar to Canny's.

In [22], Haralick proposes the use of zero crossings of the second directional derivative of the image brightness function. This is theoretically the same as using maxima in the first directional derivatives, and in one dimension is the same as the LoG filter. The zero crossings are found by fitting two dimensional functions to the image brightness surface and then analyzing the resulting fit (see the following section). The functions used are "discrete orthogonal polynomials of up to degree three". There is a problem with "phantom edges" created by the second directional derivative at "staircase structures". Connectivity at junctions is poor. In [18] Fleck describes the use of the second directional derivative for edge finding, with various extensions to the basic use of zero crossings. The problem of phantom edges is reduced with a test using the first and third derivatives. Image noise is reduced using topological sums; smoothing is achieved at potential edge points by measuring the local support they have for their "edge point type" classification. The overall algorithm (and in particular the noise reduction part) is computationally expensive.

The approach of surface fitting defines a two dimensional function which should be "flexible" enough to provide a good approximation to the image brightness function. The best fit of the function to the image is then found at each image position, and the parameters of the fit are used either to estimate image derivatives directly or to find edges using alternative definitions. The problem with the mathematical model approach is that deviations from the surface model cannot (by definition) be accommodated by varying the model's parameters. Work taking this approach includes [28], [42] and [59].

In [43] Noble uses mathematical morphology to find image structure. Several different morphological operations are described; these are used to enhance edges and find two dimensional features. The "erode-dilate" operator is similar to a first order derivative, and the "open-close" operator is similar to a second order derivative. Good quality edges are found by tracking both sides of each edge and then stitching these "half boundaries" together. Connectivity at junctions is good, although spurious short "tails" sometimes appear at structures such as "T" junctions. The algorithm, including edge tracking, is fairly computationally expensive.

In [72] and [70] Venkatesh, Owens et. al. describe the approach of using "local energy" (in the frequency domain) to find features. The local energy is found from quadrature pairs of image functions, such as the image function and its Hilbert transform. Fourier transforms of the image were initially used to find the required functions, at large computational cost. In a more "realistic" implementation also described, processing is performed using a function pair very similar to a first and a second derivative; this is shown to be equivalent to the method proposed originally. This approach has the advantage of being able to find a wider variety of edge types than that typically found using antisymmetric linear filters, however, this is at the expense of optimizing the signal to noise ratio. (In [9] Canny makes a similar point, when discussing the symmetric and antisymmetric components of a filter.) Other work taking this approach includes [48] and [53].

In [6] Blake and Zisserman used a weak membrane model in the framework of statistical regularization. The elements of the cost functional (which is to be globally minimized) are the smoothness of the estimated surface, the quality of the fit of the estimated surface to the original data and the "line processes", i.e., the discontinuities in the membrane. This method does not pick up fine complicated detail very well, and is computationally expensive. This approach has been developed further, for example, in [21], [44] and [68]. Other methods of using "global" support for edge detection include [45] and [23].

The edge detector described in this paper uses a completely new definition of edges, and in doing so, solves many of the problems which existing algorithms have failed to overcome.

4.2 The SUSAN Edge Detector in Detail

The edge detection algorithm described here follows the usual method of taking an image and, using a predetermined window centred on each pixel in the image, applying a locally acting set of rules to give an edge response. This response is then processed to give as the output a set of edges.

The SUSAN edge finder has been implemented using circular masks (sometimes known as windows or kernels) to give isotropic responses.[7] Digital approximations to circles have been used, either with constant weighting within them or with Gaussian weighting – this is discussed further later. The usual radius is 3.4 pixels (giving a mask of 37 pixels), and the smallest mask considered is the traditional three by three mask. The 37 pixel circular mask is used in all feature detection experiments unless otherwise stated.

The mask is placed at each point in the image and, for each point, the brightness of each pixel within the mask is compared with that of the nucleus (the centre point). Originally a simple equation determined this comparison – see Figure 5(a);

$$c(\vec{r},\vec{r_0}) = \begin{cases} 1 & \text{if } |I(\vec{r}) - I(\vec{r_0})| \leq t \\ 0 & \text{if } |I(\vec{r}) - I(\vec{r_0})| > t, \end{cases} \tag{1}$$

where $\vec{r_0}$ is the position of the nucleus in the two dimensional image, $\vec{r}$ is the position of any other point

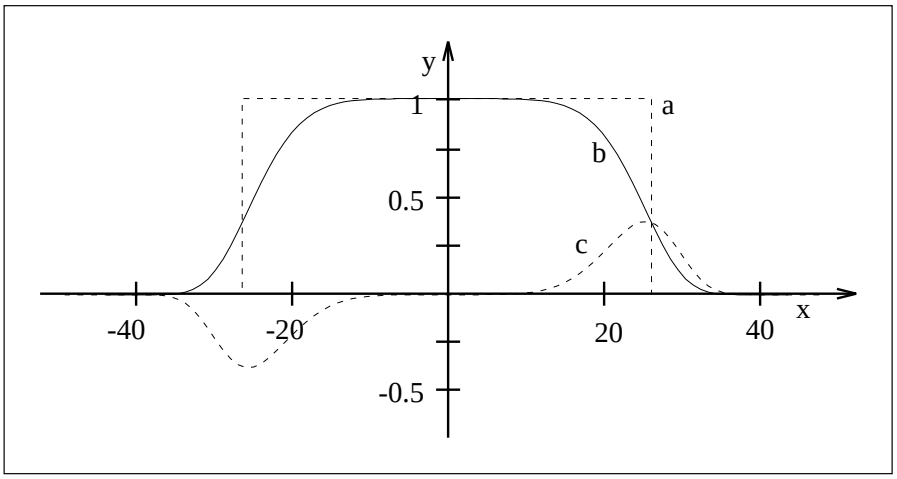

Figure 5: a) The original similarity function (y axis, no units) versus pixel brightness difference (x axis, in greylevels). For this example the pixel brightness difference "threshold" is set at ± 27 greylevels. b) The more stable function now used. c) The boundary detector B (see later text).

within the mask, $I(\vec{r})$ is the brightness of any pixel, t is the brightness difference threshold and c is the output of the comparison.[8] This comparison is done for each pixel within the mask, and a running total, n, of the outputs (c) is made;

$$n(\vec{r_0}) = \sum_{\vec{r}} c(\vec{r},\vec{r_0}). \tag{2}$$

This total n is just the number of pixels in the USAN, i.e. it gives the USAN's area. As described earlier this total is eventually minimized. The parameter t determines the minimum contrast of features which will be detected and also the maximum amount of noise which will be ignored. Further discussion later will show why performance is not dependent on any "fine-tuning" of the value of t.

Next, n is compared with a fixed threshold g (the "geometric threshold"[9]), which is set to $3n_{\max}/4$, where $n_{\max}$ is the maximum value which n can take. The initial edge response is then created by using the

[7] In [9] the circular Gaussian mask is developed into a non-circular one once edge direction has been found. This reduces the contribution of noise to the edge signal. This step could be mirrored in the SUSAN algorithm, but there is no need, as no signal as such is being filtered, and the noise reduction is already large due to integration over the mask.

[8] The software implementation uses a value of 100 rather than 1 for the maximum similarity, so that integer arithmetic may be used rather than floating point, for speed of computation.

[9] Note that the *only* threshold introduced in this paper which is *not* optimally established by careful analysis is t, the brightness difference threshold. This is the parameter which controls the sensitivity of the feature detection algorithms. Feature detection applications almost always require that sensitivity can be controlled by variation of one or more parameters, preferably only one.

following rule:

$$R(\vec{r_0}) = \begin{cases} g - n(\vec{r_0}) & \text{if } n(\vec{r_0}) < g \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

where $R(\vec{r_0})$ is the initial edge response. This is clearly a simple formulation of the SUSAN principle, i.e., the smaller the USAN area, the larger the edge response. When non-maximum suppression has been performed the edge enhancement is complete.

When finding edges in the absence of noise, there would be no need for the geometric threshold at all. However, in order to give optimal noise rejection g is set to $3n_{\max}/4$. This value is calculated from analysis of the expectation value of the response in the presence of noise only – see later. The use of g should not result in incorrect dismissal of correct edges for the following reasons. If a step edge (of general curvature) is considered, it can be seen that n will always be less than (or equal to) $n_{\max}/2$ on at least one side of the edge. In the case of a curved edge, this will correspond to the boundary of the region which is convex at the step edge. Thus valid edges should not be rejected. If the edge is not an ideal step edge but has a smoother profile then n will have even lower minima so that there is even less danger of edges being wrongly rejected.

The algorithm as described gives quite good results, but a much more stable and sensible equation to use for c in place of Equation 1 is

$$c(\vec{r}, \vec{r_0}) = e^{-(\frac{I(\vec{r}) - I(\vec{r_0})}{t})^6}. \tag{4}$$

This equation is plotted in Figure 5(b). The form of Equation 4 was chosen to give a "smoother" version of Equation 1. This allows a pixel's brightness to vary slightly without having too large an effect on c, even if it is near the threshold position. The exact form for Equation 4, i.e., the use of the sixth power, can be shown to be the theoretical optimum; see later for analytic comparison of shapes varying from one extreme (Gaussian) to the other (square function, as originally used). This form gives a balance between good stability about the threshold and the function originally required (namely to count pixels that have similar brightness to the nucleus as "in" the univalue surface and to count pixels with dissimilar brightness as "out" of the surface). The equation is implemented as a look up table for speed. (The threshold t determines, of course, the minimum contrast of edges which will be picked up, and provides an easy way of controlling this, even automatically, if necessary.)

Computation of the edge direction is necessary for a variety of reasons. Firstly, if non-maximum suppression is to be performed the edge direction must be found. It is also necessary if the edge is to be localized to sub-pixel accuracy. Finally, applications using the final edges often use the edge direction for each edge point as well as its position and strength. In the case of most existing edge detectors, edge direction is found as part of the edge enhancement. As the SUSAN principle does not require edge direction to be found for enhancement to take place, a reliable method of finding it from the USAN has been developed. This method is now described.

The direction of an edge associated with an image point which has a non zero edge strength is found by analyzing the USAN in one of two ways, depending on the type of edge point which is being examined. For examples of the two types of edge points, see Figure 6. It can be seen that points (a) and (b) have USAN shapes which would be expected for an ideal step edge. In this case (which shall be known as the "*inter*-pixel edge case") the vector between the centre of gravity $\vec{\overline{r}}$ of the USAN and the nucleus of the mask is perpendicular to the local edge direction. The centre of gravity is found thus;

$$\vec{\overline{r}}(\vec{r_0}) = \frac{\sum_{\vec{r}} \vec{r}\, c(\vec{r}, \vec{r_0})}{\sum_{\vec{r}} c(\vec{r}, \vec{r_0})}. \tag{5}$$

This simple rule allows the edge direction to be found for this type of edge point.

The point shown in case (c) lies on a thin band which has a brightness roughly half way between the brightnesses of the two regions which generate the edge. This occurs when the real edge projects very close to the centre of a pixel rather than in between pixels (or when the edge is not a sharp step edge in the first place) and when the edge contrast is high. In this case, (which shall be known as the "*intra*-pixel edge case")

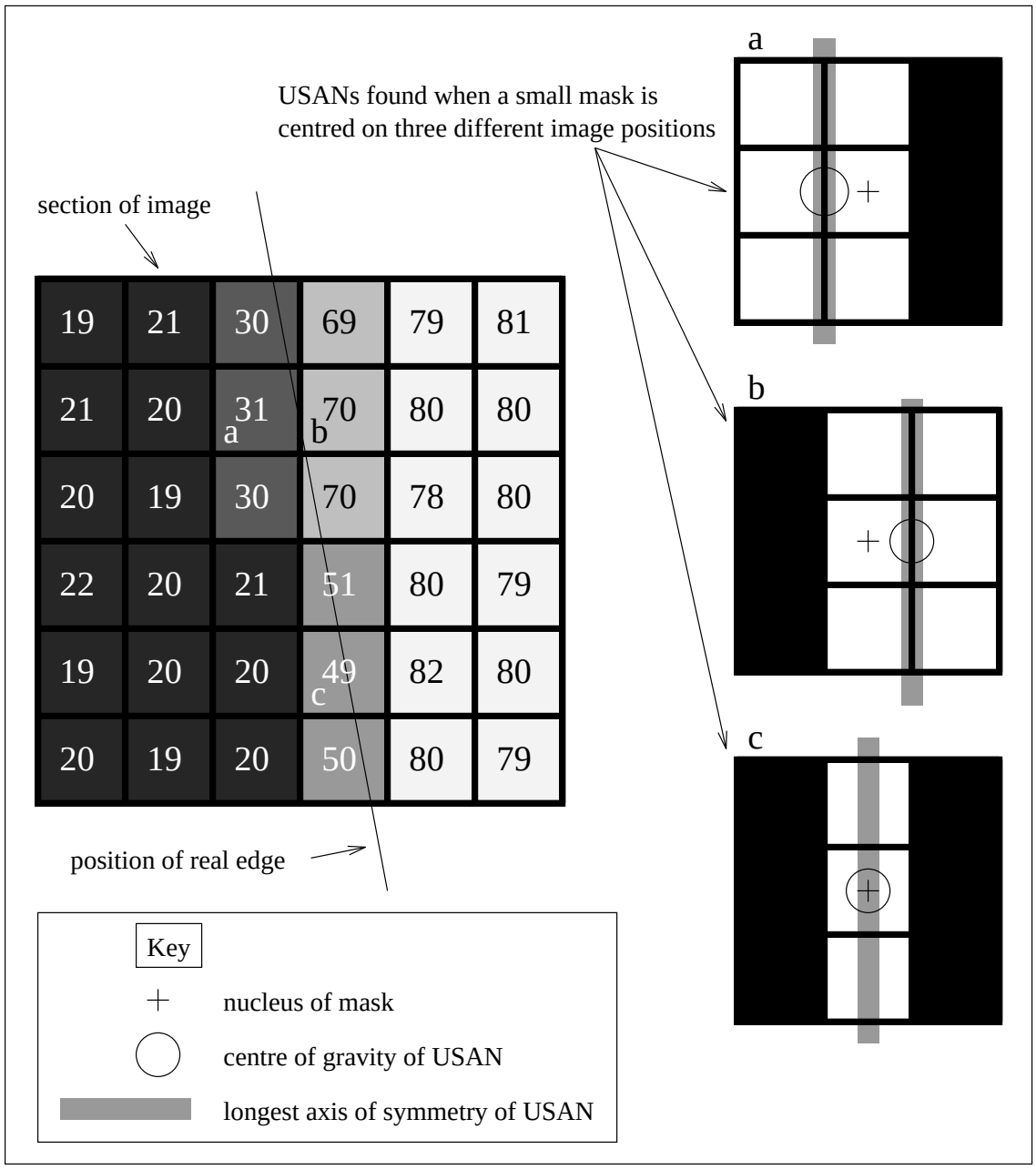

Figure 6: The two main edge types; a typical straight edge in a section of a real image, with brightness indicated by the numerical text as well as the shading of the pixels. The USANs for three points of interest are shown as the white regions of a small (3 by 3) mask. Points (a) and (b) are standard edge points, lying definitely on one side of the edge or the other. Point (c) lies on a band of brightness half way between the brightnesses of the two regions generating the edge. It therefore has a differently shaped USAN, and a centre of gravity coinciding with the nucleus.

the USAN formed is a thin line in the direction of the edge, as can be seen in Figure 6. The edge direction is thus calculated by finding the longest axis of symmetry. This is estimated by taking the sums

$$\overline{(x-x_0)^2}(\vec{r_0}) = \sum_{\vec{r}} (x-x_0)^2 \; c(\vec{r},\vec{r_0}), \tag{6}$$

$$\overline{(y-y_0)^2}(\vec{r_0}) = \sum_{\vec{r}} (y-y_0)^2 \; c(\vec{r},\vec{r_0}) \tag{7}$$

and

$$\overline{(x-x_0)(y-y_0)}(\vec{r_0}) = \sum_{\vec{r}} (x-x_0)(y-y_0) \; c(\vec{r},\vec{r_0}). \tag{8}$$

(No normalization of the sums is necessary with the second moment calculations.) The ratio of $\overline{(y-y_0)^2}$ to $\overline{(x-x_0)^2}$ is used to determine the orientation of the edge; the sign of $\overline{(x-x_0)(y-y_0)}$ is used to determine whether a diagonal edge has positive or negative gradient. Thus in the case of edge points like (c) the edge direction is again found in a simple manner.

The remaining question is how to automatically determine which case fits any image point. Firstly, if the USAN area (in pixels) is smaller than the mask diameter (in pixels) then the intra-pixel edge case is assumed. Figure 6 shows clearly the logic behind this. If the USAN area is larger than this threshold, then the centre of gravity of the USAN is found, and used to calculate the edge direction according to the inter-pixel edge case. If however, the centre of gravity is found to lie less than one pixel away from the nucleus then it is likely that the intra-pixel edge case more accurately describes the situation. (This can arise for example if the intermediate brightness band is more than one pixel wide, when larger mask sizes are used.)

The edge direction can be found to varying accuracy using this method, depending on the intended application. If the final output desired is simply a binarized edge image it may be enough to simply categorize an edge element as being "vertical" or "horizontal".

An interesting point about the SUSAN edge detector is shown by the USAN areas in Figure 6. It can be simply seen that as an edge becomes blurred, the area of the USAN at the centre of the edge will decrease. Thus we have the interesting phenomenon that the response to an edge will increase as the edge is smoothed or blurred. This is most unusual for an edge detector, and is not an undesirable effect.

Finally, therefore, the edge response image is suppressed so that non-maxima (in the direction perpendicular to the edge) are prevented from being reported as edge points. Following this, the "strength thinned" image can be "binary thinned". This means that standard thinning processes can be used to ensure that the edges obey required rules about number-of-neighbour connectivity, so that remaining noise points can be removed, and so that edge points incorrectly removed by the non-maximum suppression can be replaced. A set of rules for binary thinning has been implemented (still making use of the strengths in the non-suppressed edge response image) which work well to give a good final binarized edge image. These rules are described in [63].

If the position of an edge is required to accuracy greater than that given by using whole pixels (the accuracy which would be achieved by using the theory presented so far), it may be calculated in the following way. For each edge point, the edge direction is found and the edge is thinned in a direction perpendicular to this. The remaining edge points then have a 3 point quadratic curve fit (perpendicular to the edge) to the initial edge response, and the turning point of this fit (which should be less than half a pixel's distance from the centre of the thinned edge point) is taken as the exact position of the edge.

With respect to the scale-space behaviour of the SUSAN edge detector, scale-space graphs showing edge localization against mask size (e.g., plotting a single horizontal line from the edge image against mask size, in the manner of Witkin [75]) give vertical lines. (Most edge detectors do not give scale-invariant edge positions, thus producing curves in scale-space graphs.) This is obviously a desirable feature, as it means that accuracy does not depend on mask size. This is to be expected; the minimum USAN area when approaching an edge occurs on top of the edge regardless of the mask size.

In summary then, the algorithm performs the following steps at each image pixel:

1. Place a circular mask around the pixel in question (the nucleus).
2. Using Equation 4 calculate the number of pixels within the circular mask which have similar brightness to the nucleus. (These pixels define the USAN.)
3. Using Equation 3 subtract the USAN size from the geometric threshold to produce an edge strength image.
4. Use moment calculations applied to the USAN to find the edge direction.
5. Apply non-maximum suppression, thinning and sub-pixel estimation, if required.

4.3 Analysis of the SUSAN Edge Detector

A simple derivation is now given which shows the theoretical coincidence of the exact SUSAN edge position and the zero crossing of the second derivative of the image function.

If we assume a one dimensional input signal which is monotonically increasing then it can be simply shown that a minimum in the USAN area (in this case USAN length) will be equivalent to the edge definition which places the edge at the position of inflection of the curve, that is, where the second derivative of the image function is zero. See Figure 7 for an example image function and three one dimensional mask positions.

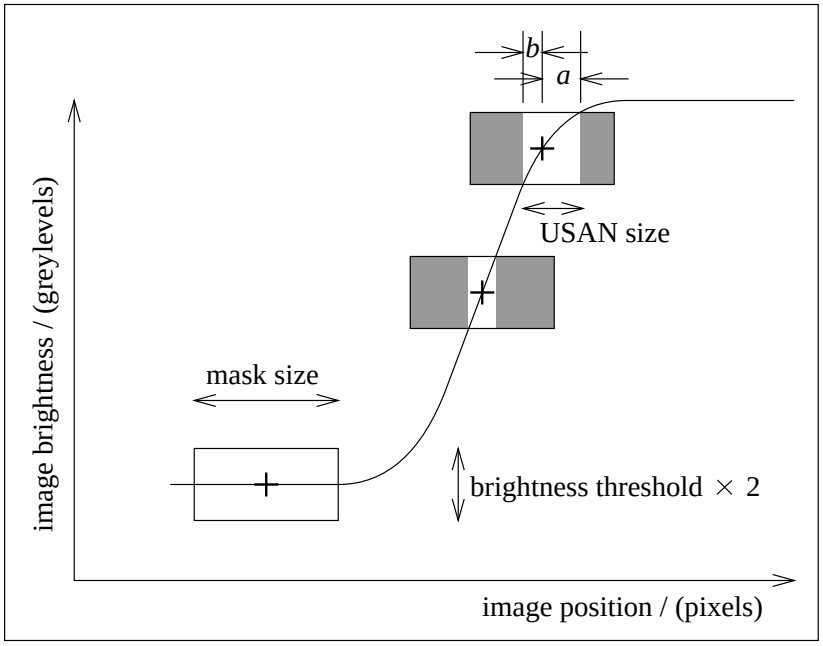

Figure 7: A monotonically increasing one dimensional input signal and three mask positions showing USAN size varying with signal gradient. The USAN is shown as the white portion of the mask. The mask centred on the signal inflection has the smallest USAN.

If we have general image position x, image function $I(x)$ and brightness threshold t, then a and b (the "right" and "left" widths of the USAN – see Figure 7), for any specific image position x_0, are, by definition, given by

$$I(x_0 + a(x_0)) = I(x_0) + t \tag{9}$$

and

$$I(x_0 - b(x_0)) = I(x_0) - t. \tag{10}$$

The SUSAN principle is formulated in the following equation, where $n(x_0)$ is the USAN size at x_0 (the size being a length in this one dimensional case);

$$\frac{dn(x_0)}{dx_0} = \frac{d(a(x_0) + b(x_0))}{dx_0} = 0, \tag{11}$$

assuming that a check is made for n being a minimum and not a maximum. This gives

$$\frac{d}{dx_0}(x(I(x_0) + t) - x(I(x_0) - t)) = 0, \tag{12}$$

giving

$$\left(\frac{dx}{dI}(I(x_0) + t)\ \frac{dI}{dx_0}(x_0)\right) - \left(\frac{dx}{dI}(I(x_0) - t)\ \frac{dI}{dx_0}(x_0)\right) = 0, \tag{13}$$

that is,

$$\frac{dI}{dx}(x_0 + a(x_0)) - \frac{dI}{dx}(x_0 - b(x_0)) = 0, \tag{14}$$

unless the image function is constant, an uninteresting case. This is equivalent, in the continuous limit (that is, as t tends to zero), to

$$\frac{d^2I}{dx^2}(x_0) = 0, \tag{15}$$

and we have equivalence between the two definitions of exact edge position.

This mathematical equivalence in no way means that the SUSAN edge finding method is performing the same function as a derivative based edge detector. No derivatives are taken; indeed, no direction of maximum gradient is identified (in order for differentiation to be performed) in the first place. However, the preceding argument shows why the SUSAN edge detector gives edges to good sub-pixel accuracy even with edges that are not perfect step edges.

It can be seen from the preceding discussion that the initial response of the SUSAN edge detector to a step edge will be increased as the edge is smoothed. The response is also broadened.

As mentioned earlier, the form of the brightness comparison function, the heart of USAN determination (Equation 4), can be shown to achieve the optimal balance between stability and enhancement. This is equivalent to satisfying the criterion that there should be a minimum number of false negatives and false positives. This criterion is formulated in the following expression:

$$F(d, t, s) = \frac{\sqrt{\mathrm{var}(R_S)} + \sqrt{\mathrm{var}(R_N)}}{< R_S > - < R_N >}, \tag{16}$$

where F is proportional to the number of false positives and false negatives that will be reported, s is image noise standard deviation, R_N is the SUSAN response strength with no edge present and R_S is the SUSAN response strength when the mask is centred on an edge of strength d. Thus this expression simply divides the expected total "noise" by the expected "signal" in the most logical way.

The value of F will depend on the relative values of d, t and s. Thus fixing t at 1 and varying the other two variables will cater for all eventualities. If this expression is averaged over all likely values of d and s, the overall quality can be evaluated. The (wide) ranges of d and s chosen were $1 \rightarrow 10$ and $0 \rightarrow 1/\sqrt{2}$ respectively; the reason for choosing the the upper limit on s will shortly become apparent.

All of the calculations within the SUSAN feature detectors are based around the use of the equation

$$\Delta = I(\vec{r}) - I(\vec{r_0}), \tag{17}$$

where I is defined as before. If the image noise is Gaussian with standard deviation s then the noise on Δ will also be Gaussian with standard deviation $s_\Delta = \sqrt{2}s$. Thus F is evaluated over the interval $s_\Delta = 0 \rightarrow 1$.

To evaluate F given d, t and s, it is necessary to find the expectation values and variances of R_N and R_S, using

$$< R_N >= 1 - < N >, \tag{18}$$

$$< R_S >= 1 - \frac{1}{2} < N > - \frac{1}{2} < S >, \tag{19}$$

$$\mathrm{var}(R_N) = \mathrm{var}(N), \tag{20}$$

and

$$\mathrm{var}(R_S) = \frac{1}{4}\mathrm{var}(N) + \frac{1}{4}\mathrm{var}(S), \tag{21}$$

where

$$N = e^{-(\frac{\Delta}{t})^J} \tag{22}$$

and

$$S = e^{-(\frac{\Delta - d}{t})^J}, \tag{23}$$

with J being the even number determining the shape of the brightness comparison function; it is 2 for a Gaussian and infinity for a square function. (N and S are responses due to individual pixels with no edge present and across an edge respectively, using the brightness comparison function. The form of R_S arises from the fact that in the presence of an edge, half of the response will be due to noise only, and half of the response will be due to the edge as well as noise.) The expectation values and variances are computed using the integrals:

$$< X >= \frac{1}{s_\Delta \sqrt{2\pi}} \int_{-\infty}^{\infty} X(\Delta)\, e^{-\frac{1}{2}\left(\frac{\Delta}{s_\Delta}\right)^2} d\Delta \tag{24}$$

and

$$\mathrm{var}(X) = \frac{1}{s_\Delta \sqrt{2\pi}} \int_{-\infty}^{\infty} X^2(\Delta)\, e^{-\frac{1}{2}\left(\frac{\Delta}{s_\Delta}\right)^2} d\Delta - < X >^2, \tag{25}$$

where the integral includes the Gaussian model of noise. The integrals are calculated over the intervals of d and s_Δ specified and the mean result for F is thus found. The resulting plot of F against J is shown in Figure 8. It is clear that the optimal value for J is 6. This gives the form of the function used in the SUSAN filters.

The optimal value for g, the geometric threshold, is found by calculating the mean expectation value for N over the same realistic range of image noise as before. With no noise present, $< N >$ is 1; its mean value in the presence of noise is calculated (using the integral shown in Equation 24) to be close to 0.75. Therefore the value of g which is likely to provide the correct amount of noise rejection is 3/4 of the maximum possible USAN size.

Finally, the spatial part of the filter is of interest. In [9] the "difference of boxes" edge detector is compared unfavourably with a "derivative of Gaussian" edge detector, due to the sharp cutoff in the spatial profile of the "box" filter. Because the Gaussian based filter has smooth edges, image noise is suppressed more effectively than using a box (or square) shaped filter. Therefore it may seem that a similar approach to the design of the mask used in SUSAN feature detection would be preferable. Indeed, in the brightness domain, it was found worthwhile to use a smooth brightness comparison function (as shown in Equation 4), rather than having thresholding done with a sharp cutoff. In the case of the spatial extent of the mask, the equivalent process would be achieved by using

$$n(\vec{r_0}) = \sum_{\vec{r}} e^{\frac{-(\vec{r} - \vec{r_0})^2}{2\sigma^2}} c(\vec{r}, \vec{r_0}) \tag{26}$$

in place of Equation 2, where σ is a suitable scaling factor. This would give a circular mask with pixels contributing less to the calculations as their distance from the nucleus increased – a Gaussian profile. An edge detector based on this kind of mask has been fully implemented and tested.

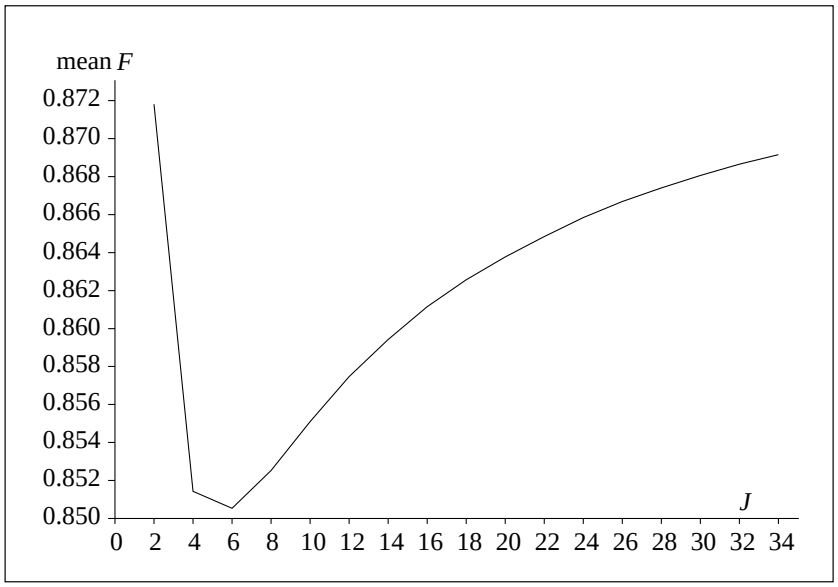

Figure 8: A plot of an objective formulation of expected false negatives and positives against the J factor in the brightness comparison function.

The difference in practice between the Gaussian spatial SUSAN detector and a sharp cutoff SUSAN detector turns out to be minimal; there is very little difference in the final reliability or localization of reported edges. The explanation for this possibly surprising result is that, unlike the differential edge detector, no image derivatives are used, so that the problem of noise is vastly reduced in the first place. Also, a linear filter is performing a different task to the initial SUSAN response. In the former, a filter has a Gaussian profile in a "brightness versus distance" plot. However, the latter uses either a sharp circular mask or a Gaussian profile in the spatial domain and smoothed thresholding in the brightness domain. Thus "half" of the algorithm already has a smooth cutoff action. From the results, it seems that it is this part (the brightness domain) which is more important. Thus it does not appear necessary to use a Gaussian spatial mask, although it is easy to do so if desired.

4.4 Testing of the SUSAN Edge Detector

A quantitative test of the initial response of the SUSAN detector compared with four other edge enhancement algorithms was described in Section 3. To give a more meaningful explanation of the tests described in [71] than the short one given earlier would require an undeserved amount of space. Suffice it to say that the initial response given by SUSAN was better than the best results of the four detectors used in these tests, using all six suggested "failure measures".

A one dimensional test of the SUSAN response to various edge types has been carried out using the input signal shown in Figure 9. The results are good. All the edge types are correctly reported apart from the roof edge. (Note that even the roof edge produces a very small local maximum in the response.) The ridge edges are picked up correctly, and produce only one response, as desired. Thus the SUSAN edge detector is shown to report lines correctly and not just edges. With the two pixel width ridge it is of course a matter of definition whether one ridge or two edges should be reported. The response is symmetrical about each edge type, which results in sub-pixel accuracy giving excellent results.

The roof edge obviously does not conform to the (fairly lenient) model of an edge which the SUSAN detector is based on. In [9] Canny states that

> "A roof edge detector has not been incorporated into the implementation of the edge detector