The robustness of some scale-spaces

Richard Harvey, Alison Bosson, J. Andrew Bangham
School of Information Systems, University of East Anglia,
Norwich, NR4 7TJ, UK.
Tel: +44 1603 593257; fax: +44 1603 593345
{rwh,ab,bosson}@sys.uea.ac.uk

Abstract

There are several methods for forming a scale-space and they may be classified as being based on diffusion or morphology. However it is rare for the methods to be compared. Here we outline a method for such a comparison based on robustness and give results for linear diffusion, the most widely studied method, and a sieve (a new morphological method). We find that the standard diffusion-based systems are not as insensitive to noise and occlusion as one might wish.

1 Introduction

Systems that preserve scale-space causality are usually associated with Gaussian filters [1,2] and diffusion [3] in which the image forms the initial conditions for a discretisation of the continuous diffusion equation,

$$\nabla \cdot (c \nabla f) = \frac{\partial f}{\partial s}. \tag{1}$$

If the conduction coefficent, c, is a constant this becomes the linear diffusion equation, $c\nabla^2 f = f_s$ which may be implemented by convolving the image with the Green's function of the diffusion equation: a Gaussian filter. Of course, care is needed when discretising (1) but, if it is done correctly [4], a scale-space with discrete space and continuous scale may be formed[1] with separable filters ($\gamma = 0$ in [4]) as,

$$f^{(s)}(x,y) = \sum_{m=-\infty}^{\infty} T(m;s) \sum_{n=-\infty}^{\infty} T(n;s) f(x-m, y-n), \tag{2}$$

$f(x,y)$ is the pixel value at position (x,y) and $f^{(s)}(x,y)$ is the pixel value after smoothing to scale s. $T(n;s)$ is the discrete approximation to the Gaussian kernel $T(n;s) = e^{-s} I_n(s)$ and $I_n(s)$ is the modified Bessel function of the first kind.

Such scale-space systems have several problems:

1. The edges of sharp-edged objects become blurred at large scales. This leads to poor localisation and a necessity to track edges back through scale-space.

[1] Or with a discrete scale parameter if preferred.

2. They introduce extrema, albeit in rather pathological cases.

3. They are not scale calibrated. An image filtered to scale s contains features at many scales.

4. Multiple convolutions may require significant amounts of computation using real arithmetic.

Problem 1 may be reduced by reverting to (1) and allowing the conduction coefficient, c, to be a function of $\|\nabla f\|$ thus $c = g(\|\nabla f\|)$. If this function is carefully chosen (several have been suggested) then the effect is to allow diffusion in low contrast regions but not at sharp-edges. Unfortunately anisotropic diffusion requires even more computation than linear diffusion and problem 3 is exacerbated – sharp-edged small scale objects persist to large scales.

Problem 2 is illustrated in Figure 1. On the left is a three-dimensional intensity plot of an image. A disc and a square connected by thin isthmus form its main feature. After diffusion filtering (centre) the single maximum associated with the two discs and isthmus has now become two maxima. This example, due to Pizer [5], is well known and led to a new definition of scale-space causality namely the "non-enhancement of existing regional extrema" principle which is satisfied by diffusion systems. The right-hand plot in Figure 1 shows the result of applying an M-sieve. The smaller area disc is removed and the large area feature is unaffected.

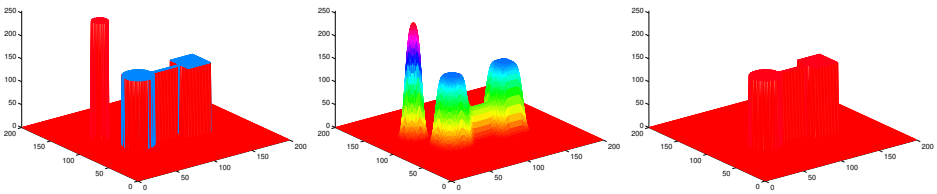

Figure 1: Original image (left), Gaussian filtered image (centre), sieved image (right)

Figure 1 also illustrates problem 3. In the Gaussian processor the intensity of the output at scale s is proportional to the scale and intensity of the original feature. So to recover the parameters of the original signal implies either deconvolution or scale-space tracking. Both of these are difficult tasks and, as far as image interpretation goes, having peaks representing extrema from objects of a variety of scales is inconvenient. It is possible to design morphological filters that produce at scale s only objects of size s.

2 Morphological processors

Mathematical morphology is the analysis of signals, particularly images, by shape and has been developed by Serra [6–8] from work by Matheron [9] but it also derives

from work of others (Blum [10] for example). It is widely used in commercial image processing packages.

Of particular importance are Alternating Sequential Filters with greyscale structuring elements that match features [7, 11]. They are said to be a powerful analogue of linear matched filters and there are strategies for designing optimal filters with flat, rigid, two dimensional flat structuring elements [12]. More recently, alternating sequential filters, that do not impose a rigid geometry on objects, and connected sets have been described [13–17].

A separate stream of development has been that of rank filters, including median, root median, recursive median and, more generally, stack filters. Such filters have been developed primarily to remove random noise from signals and images, although there are suggestions for using them for tasks such as shape recognition [18]. There is a close relationship between rank filters and morphological filters [19]. Rank (median) filters are usually self-dual and robustly reject random noise. A recent development has been the sieve and its variants [20–25] which use the connected set operators found in Alternating Sequential Filters with ordering operators associated with rank filtering.

3 Sieves and their properties

The basis of the sieve is the representation of an image as a graph [26, 27] $G = (V, E)$. The set of edges E describes the adjacency of the pixels (which are the vertices V). For a one-dimensional image the graph is just a list [28] but for a multidimensional image the graph defines the neighbourhood of a particular pixel. For example, in Figure 2 which shows a four-connected two-dimensional image the graph would be $G = (V, E)$ with $V = \{1, 2, 3, \ldots, 16\}$ and $G = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{1, 5\}, \ldots\}$. In Figure 2 the numbers are labels not intensities.

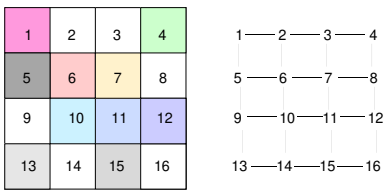

Figure 2: An image represented as a graph.

The algorithm proceeds by defining a region, $C_r(G, x)$ over the graph that encloses the pixel (vertex) x,

$$C_r(G, x) = \{\xi \in C_r(G) | x \in \xi\} \tag{3}$$

where $C_r(G)$ is the set of connected subsets of G with r elements. Thus $C_r(G, x)$ is the set of connected subsets of r elements that contain x. In Figure 2, for example,

$C_2(G,6) = \{\{5,6\},\{2,6\},\{6,10\},\{6,7\}\}$. For each integer $r \geq 1$ the operators ψ_r, γ_r, $\mathcal{M}^r$, $\mathcal{N}^r : \boldsymbol{Z}^V \to \boldsymbol{Z}^V$ are defined as

$$\psi_s f(x) = \max_{\xi \in C_s(G,x)} \min_{u \in \xi} f(u), \qquad \gamma_s f(x) = \min_{\xi \in C_s(G,x)} \max_{u \in \xi} f(u), \tag{4}$$

$\mathcal{M}^r$ is a greyscale opening followed by a closing defined over a region of size r and $\mathcal{N}^r$ is a greyscale closing followed by an opening over the same region. $\mathcal{M}^r = \gamma_r \psi_r$, $\mathcal{N}^r = \psi_r \gamma_r$.

The types of sieve known as M- or N-sieve are formed by repeated operation of the $\mathcal{M}$ or $\mathcal{N}$ operators. They are also known as connected alternating sequential filters. An M-sieve of f is the sequence $(f^{(s)})_{s=1}^{\infty}$ given by

$$f^{(1)} = \mathcal{M}^1 f, \quad f^{(s+1)} = \mathcal{M}^{s+1} f^{(s)}, \quad s \geq 1 \tag{5}$$

The N-sieve is defined similarly. The output of an area sieve is usually taken to be the set of *granule functions*

$$d^{(s)} = f^{(s)} - f^{(s+1)} \text{ for each integer } \quad s \geq 1 \tag{6}$$

These form the scale selection surface and non-zero connected regions within granule functions are called granules. Each granule has sharp edges and, at a particular scale, all granules have the same area. In this sense the sieve is *scale calibrated*, a characteristic illustrated in Figure 1, that addresses Problem 3 in Section 1.

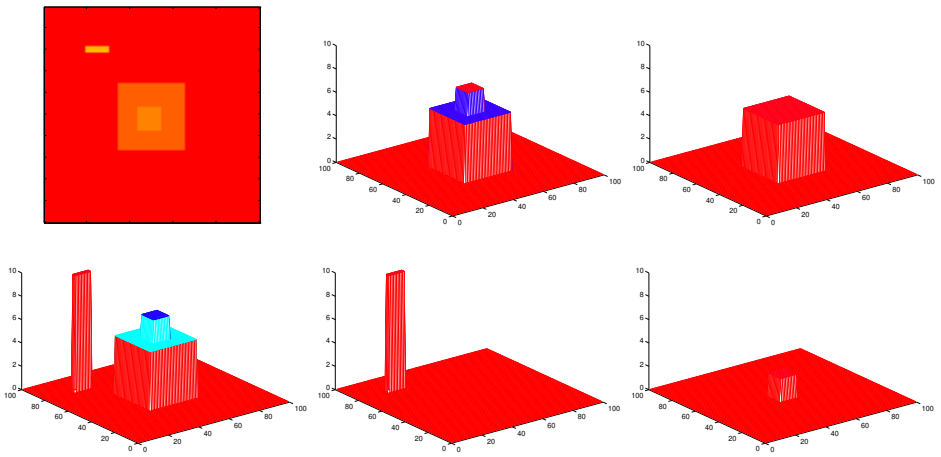

Figure 3: Operation of the area sieve. Original image (left column). Filtered to scale 33 (top centre). Filtered to scale 121 (top right). Granules are shown in bottom centre and right.

An example of the operation of the area sieve is shown in Figure 3. The figures in the first column show the original 100 by 100 pixel image and its three dimensional intensity plot. The image has one maxima of area 33 pixels on the left-hand side and another of area 121 pixels in the centre. The middle column of Figure 3 shows the result of applying an M-sieve to scale 33 (filtering to scales

smaller than 33 pixels produces no change). In the filter output, shown on the top row, the 33 pixel extremum has been removed from the image and is shown as a granule in the granularity domain below. Filtering to increasing scales produces no change up to scale 121 (right column), when the centre extremum is removed. Again this is shown as a granule below. At scale 900 (not shown) the new extremum in the centre of the image is removed leaving a zero intensity image.

The M operation depicted in Figure 3 is the same as the removal of maxima followed by the removal of minima. In an N-sieve these operations are reversed. With this in mind it is possible to define an m-sieve in which the extrema are processed in the order in which they occur as the graph is parsed. In one-dimension such an algorithm is equivalent to the recursive median operator. The m-sieve has properties similar to M- and N-sieves and so not all sieves are alternating sequential filters and not all alternating sequential filters are sieves.

It might appear that the sieving operation could be computationally expensive. In practice, however, the algorithms referred to here, in which the image graph is rewritten as flat-zones are merged, produce scale-spaces quicker than linear diffusion. Typically a 512 by 512 image is processed in seconds [21, 29]. This addresses Problem 4 in Section 1.

Sieves applied in two-dimensions preserve scale-space causality and this has been proved formally (Theorem 6.36 in [21]) together with a number of other properties of area sieves (Theorem 6.49 in [21]). This addresses Problem 2 in Section 1.

The properties described so far would appear to be desirable. However, it is well known that morphological operators such as erosion and dilation are sensitive to noise, more complicated morphological filters are less so, and Gaussian filters, having the form of matched filters, are fairly insensitive to noise. Whether sieves, and other filters in their class, have practical application depends on whether they reject image corruptions robustly.

4 Robustness

Noise interacts with CCD images and photographs in a complicated manner [30, 31]. Furthermore the underlying noiseless image distribution is also extremely complicated. For this reason it is usual to study the performance of filters using rather stylised models of images and noise. Here we shall consider the underlying image to be a disc or a square which we shall assume is corrupted with either Gaussian noise (a model of noise in the imaging system) or impulsive noise (a model for specular reflection or pixel drop-out).

Figure 4 shows examples of the target images. The target image is a disc or square of amplitude 144 in the centre of a 100 by 100 pixel image with background 112. To this, is added either uncorrelated Gaussian noise ($\mu = 0, \sigma = 24$), or alternatively, pixels are replaced with a random value in the range $(0, 256)$ with a noise density of 0.2. The resulting image is clipped into the range (0,255).

The performance of sieves is compared to Gaussian processors in which the scale-space was generated using separable filters where the image at scale s is

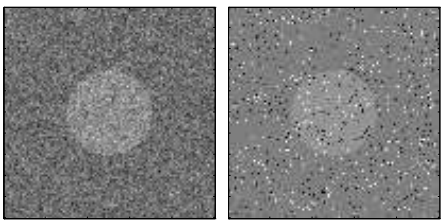

Figure 4: Examples of noise-corrupted images: Gaussian noise (top) and Impulsive noise (bottom)

computed as in (2). The scale-selection surface ([4] p. 323)

$$d = s^2 \left(f_{xx}^{(s)} + f_{yy}^{(s)} \right)^2 \tag{7}$$

is used to locate the scale-space estimate of the target parameters. For the sieve the scale-selection surface is formed from the granule functions. In the absence of noise both scale-selection surfaces will have a peak corresponding to the target.

An iterative search is conducted over all scales to find this maximum. The linear scale-space processor is scale-calibrated by normalising s so that the peak in (7) corresponds to the true area of the disc. Improved precision position estimates are obtained through least-squares interpolation around the scale-selection surface peak. Since the area sieve operates by "slicing off" peaks and troughs, the extrema may be large and flat so the centroid of these is taken as the position estimate.

		cc	diff	m	M	N	O
Gaussian	σ_x	0.212	0.643	0.280	0.241	0.261	2.882
noise	σ_y	0.198	0.629	0.243	0.244	0.251	1.141
	σ_s	30.30	32.4	55.0	58.9	48.7	105.75
Impulsive	σ_x	0.273	0.925	0.0425	0.0464	0.0400	0.076
noise	σ_y	0.330	0.790	0.0416	0.0382	0.0448	0.712
	σ_s	35.85	59.9	3.91	3.90	4.30	8.894

Table 1: Standard deviations of 150 estimates of the position (x, y) and scale, s of a disc in Gaussian and impulsive noise for normalised cross-correlation (cc), linear diffusion (diff), m- M- N- and O-sieve.

Table 1 shows that, in Gaussian noise, the diffusion system, M-, N-, and m-sieves approach the performance of the benchmark normalised cross-correlation system in which a series of increasing scale target templates are used. Opening- or O-sieves, in which one repeatedly applies increasing scale greyscale openings, have worse performance (it is known that openings are sensitive to noise). The result of multiscale closing is not shown since the closing operator cannot resolve a positive disc under these conditions[2]. Studies with the equivalent one-dimensional

[2]The closing removes minima, leaving maxima intact. Since, in this simple experiment, we search for maxima, this is a problem.

filters [32] showed that the linear diffusion system is more sensitive to Gaussian noise than the sieve. In two-dimensions this is not the case, probably because the diffusion filter matches the shape of the target. For square or rectangular targets the diffusion system does not perform so well.

Table 1 also shows the results of the impulsive noise tests. The m-, M-, N- and O-sieves can produce better scale and position estimates than those from the diffusion system. This is an important property since in real vision systems impulses, glint and occlusion are commonplace. Again, the asymmetry of multiscale closings means that the experiment would have to be specially tuned to handle positive targets.

		Intensity of disc									
		130		144		150		160		170	
		D	S	D	S	D	S	D	S	D	S
G	σ_x	1.05	1.19	0.64	0.28	0.42	0.21	0.35	0.13	0.29	0.09
	σ_y	0.86	3.65	0.63	0.24	0.42	0.19	0.35	0.11	0.27	0.10
	σ_s	69.0	144.4	32.4	55.0	31.3	40.1	23.8	25.7	17.4	16.9
I	σ_x	0.99	0.05	0.92	0.04	0.44	0.05	0.40	0.04	0.34	0.04
	σ_y	1.02	0.05	0.79	0.04	0.47	0.05	0.37	0.04	0.34	0.04
	σ_s	67.7	4.22	57.9	3.91	30.3	4.31	26.5	4.56	21.6	4.13

Table 2: Standard deviations of estimates in Gaussian (G) and impulsive (I) noise for target discs of varying amplitude for linear diffusion (D) and m–sieve (S)

The performance of sieves in replacement noise is not surprising since it is known that, in one dimension, $\mathcal{M}$ and $\mathcal{N}$ filters behave similarly to median filters which have good performance in impulsive noise. The sieve retains its performance in Gaussian noise because the cascaded operation described in (5) means that, like wavelets, the output at large scales has a large support.

Table 2 shows the results of altering the signal to noise ratio of the images by varying the target disc intensity. In Gaussian noise, for small intensity targets, the diffusion processor is superior to the m-sieve. As the intensity of the disc increases the performance of the sieve improves, approaching that of the diffusion processor. For discs corrupted with impulsive noise the amplitude of the target has no affect on the sieve. The noise is localised to one scale and is easily removed and although the performance of the diffusion processor improves with increasing disc amplitude, the sieve remains superior.

Why is the performance of image processors in noise important? Studies of imaging noise in general purpose CCD images and photographs show that the noise model is complicated with a Gaussian pdf in which the parameters of the pdf depend on the image intensity but that, for most images, the noise level is low. We believe it is the impulsive noise model that is the most interesting since it forms the basis for a study of occlusion – an important reality that is considered now.

Synthetic test images, for example on the left of Figure 5, are generated by placing discs of differing area onto a background of amplitude 32. Each 100 by 100 pixel image contains five discs. The target is the only constant image feature.

It has an area of 901 pixels and an amplitude of 47. The four occluding discs, areas 49, 253, 481 and 709 pixels, are selected at random (with replacement). They are placed at random, have random intensity (uniform range 0 to 255) and randomly occlude one another. The right-hand side of Figure 5 shows an example real cluttered image with five cardboard discs.

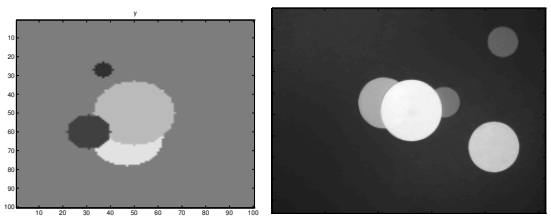

Figure 5: Occlusion image. On the left a synthetic image and on the right a real image.

Were the five objects not to have occluded one another, the sieve would produce five granules at the true positions of the objects in scale and space. With occlusion the estimates are degraded. One may judge the degradation by generating test images and plotting granule positions as a function of scale. Results for 100 real cluttered images are shown on the left-hand side of Figure 6. In Figure 6 the estimates are plotted as circles and their projections on the position and scale planes are plotted as grey dots. For the sieve results the grey dots show a clustering. There are bands corresponding to the scales of the occluding objects and there is a cluster at the centre corresponding to the target.

Repeating this experiment using a diffusion system is more complicated. Extrema must be tracked through scale-space and insignificant extrema due to finite precision effects need to be distinguished from genuine ones. Objects are denoted by maxima along the scale-space tracks.

The right-hand side of Figure 6 shows the result of performing this scale-space tracking on the 100 real images and plotting the maxima on three-dimensional plot as before.

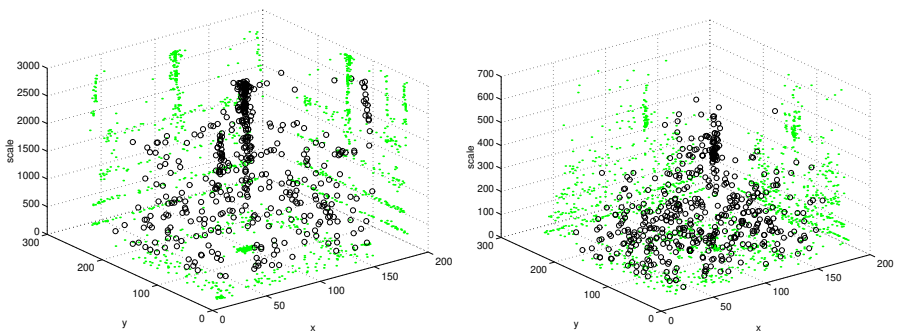

Figure 6: Scale-space estimates for real occluded images. Sieve results on the left, diffusion results on the right.

The results from the diffusion system are very confused. There is a cluster of points at large scale, corresponding to the target disc. However, at smaller scales very little structure is apparent. Similar experiments using synthetic images show the relative performance of the two systems to be the same.[3]

5 Concluding remarks

This is a stylised problem which could be solved easily using a processor that has more prior information. For example, if it is assumed that the shape of the target image is known. That is not the point. What is required is a primary vision system that obeys Marr's "principle of least commitment" [33] but preserves as much structure as possible. It appears that the sieve preserves structure and does so more robustly than the diffusion system.

As contenders for primary vision systems, sieves have some advantages. They use some of the recent results from mathematical morphology to form a processor that satisfies many of the desirable scale-space axioms and robustly separates noise and occlusion according to scale.

References

[1] A. P. Witkin. Scale-space filtering. In *8th Int. Joint Conf. Artificial Intelligence*, pages 1019–1022. IEEE, 1983.

[2] J. J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

[3] P.Perona and J.Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Patt. Anal. Mach. Intell.*, 12(7):629–639, July 1990.

[4] Lindeberg Tony. *Scale-space theory in computer vision.* Kluwer, 1994.

[5] L.M.Lifshitz and S.M.Pizer. Multi-resolution hierarchical approach to image segmentation based on intensity extrema. *IEEE Trans. Patt. Anal. Mach. Intell.*, 12(6):529–540, June 1990.

[6] J. Serra. *Image Analysis and Mathematical Morphology*, volume 1. Academic Press, 1982.

[7] J. Serra. *Image Analysis and Mathematical Morphology*, volume 2. Academic Press, 1988.

[8] J.Serra. Introduction to mathematical morphology. *Computer vision, graphics and image processing*, 35:283–305, 1986.

[9] G.Matheron. *Random sets and integral geometry.* Wiley, 1975.

[10] H. Blum. Biological shape and visual science (part 1). *J. Theoret. Biol.*, pages 205–287, 1967.

[11] S. R. Sternberg. Grayscale morphology. *Computer Vision, Graphics and Image Processing*, 35:333–335, 1986.

[12] D. Schonfeld and J. Goutsias. Optimal morphological filters for pattern restoration. In *Proc. SPIE*, volume 1199, pages 158–169, 1989.

[3] Note to reviewers: we have the synthetic results but no room to present them here.

[13] L. Vincent. Morphological area openings and closings of greyscale images. In *Workshop on Shape in Picture*. NATO, 1992.

[14] L.Vincent. Grayscale area openings and closings, their efficent implementation and applications. In J.Serra and P.Salembier, editors, *Proceedings of the international workshop on mathematical morphology and its applications to signal processing*, pages 22–27, May 1993.

[15] J.Serra and P.Salembier. Connected operators and pyramids. In *Proceedings SPIE*, volume 2030, pages 65–76, 1994.

[16] Corinne Vachier and Fernand Meyer. Extinction value: a new measurement of persistence. In Ionas Pitas, editor, *Proceedings of 1995 IEEE Workshop on nonlinear signal and image processing*, volume 1, pages 254–257, June 1995.

[17] P.Salembier and J.Serra. Flat zones filtering, connected operators and filters by reconstruction. *IEEE Transactions on Image Processing*, 8(4):1153–1160, August 1995.

[18] J.M.Sacks. Target discrimination utilizing median filters, 1984. US Patent number, 4,603,430.

[19] P. Maragos and R. W. Schafer. Morphological filters - part 1: Their set theoretic analysis and relations to linear shift-invariant filters. *IEEE Trans. Acoust, Speech, Signal Processing*, 35:1153–1169, 1987.

[20] J. A. Bangham. Median and morphological scale space filtering and zero-crossings. In *Proc. SPIE Nonlinear signal processing symp.*, volume 2180, pages 90–98, 1994.

[21] J.A.Bangham, R.Harvey, and P.D.Ling. Morphological scale-space preserving transforms in many dimensions. *J. Electronic Imaging*, 5(3):283–299, July 1996.

[22] J. A. Bangham. Data-sieving hydrophobicity plots. *Anal. Biochem.*, 174:694–700, 1988.

[23] J. A. Bangham. Properties of an improved variant on the median filter. In *in: E. Arikan, ed., Communication, Control and Signal Proc.*, volume Elsevier press, pages 1591–1597, 1990.

[24] J.A. Bangham. Properties of a series of nested median filters, namely the datasieve. *IEEE Trans. Signal Processing*, 41:31–42, 1993.

[25] J. A. Bangham, P. Ling, and R. Young. Multiscale recursive medians, scale-space and sieve transforms with an inverse. *IEEE Trans. Image Proc.*, 5:1043–1047, 1996.

[26] L. Vincent. Graphs and mathematical morphology. *Signal Processing*, 16:365–388, 1989.

[27] H.J.A.M.Heijmans, P.Nacken, A.Toet, and L.Vincent. Graph morphology. *Journal of Visual Computing and Image Representation*, 3(1):24–38, March 1992.

[28] J.A. Bangham, P.W. Ling, and R. Harvey. Nonlinear scale-space causality preserving filters. *IEEE Trans. Patt. Anal. Mach. Intelli*, 18:520–528, 1996.

[29] J. A. Bangham, S. J. Impey, and F. W. D. Woodhams. A fast 1d sieve transform for multiscale signal decomposition. In *EUSIPCO*, pages 1621–1624, 1994.

[30] R.A.Boie and I.J.Cox. An analysis of camera noise. *IEEE Trans. Patt. Anal. Mach. Intell.*, 14:671–674, 1992.

[31] G.E.Healey and R.Kondepudy. Radiometric ccd camera calibration and noise estimation. *IEEE Trans. Patt. Anal. Mach. Intell.*, 16:267–275, 1994.

[32] J. Andrew Bangham, Paul Ling, and Richard Harvey. Scale-space from nonlinear filters. In *Proc. Fifth Int. Conf. Computer Vision*, pages 163–168, 1995.

[33] D. Marr. *Vision*. W. H. Freeman and Company, New York, 1982.