Colour Morphological Scale-Spaces for Image Segmentation

David Gimenez and Adrian N. Evans
Department of Electronic and Electrical Engineering
University of Bath
Bath, BA2 7AY, UK
eepdg@bath.ac.uk , A.N.Evans@bath.ac.uk

Abstract

Morphological scale-spaces have become an important tool for analysing greyscale images. However, their extension to colour images has proven elusive until recently. In this paper an original evaluation of two recently proposed colour sieves is presented, both algorithmically and in terms of their computational and segmentation performance. A new colour sieve structure is also proposed, motivated by the relative advantages of the two sieves previously studied. A quantitative evaluation of the segmentation performance using a set of images with human ground truth from the Berkeley dataset shows the new method to produce the best segmentation performance.

1 Introduction

Morphological scale-space filters provide an attractive alternative to diffusion methods for the hierarchical analysis, segmentation [16] and classification [1] of greyscale images. Current greyscale morphology scale-spaces can be considered to belong to two categories: those based on the use of fixed structuring elements [13, 4] and those that employ area operators [2, 1]. Both classes are implemented by the application of successive openings and closings of increasing scale to produce a tree-based image representation. Of the two approaches the latter is the most attractive as it obeys the property of strong causality and, unlike the former, does not require any a priori knowledge of the shape of objects present.

In common with other morphological methods, major difficulties are encountered when trying to extend morphological scale-spaces to colour or other multichannel images as vector values cannot be placed in an unambiguous order. There have been some attempts to propose definitions for colour openings and closings for fixed structuring elements [3] that could be employed for scale-spaces but it is the advantages of the area morphology approach that have lead to its extension to colour receiving recent attention. In particular, in 2003 two approaches to colour morphological scale-spaces were presented, here referred to as the convex colour sieve (CCS) [8, 9] and the vector area morphology sieve (VAMS) [5, 6]. These sieves employ connected operators which operate by altering the colour of connected regions of constant colour, called flat zones, to produce regions that show some correspondence with image objects and, in the greyscale case, have been formally related to segmentation algorithms based on region merging/classification [7].

This paper presents a comparison of the two colour sieve methods, both algorithmically and in terms of evaluating their extrema definition, processing speed and segmentation performance. In addition, a new colour sieve structure is proposed that seeks to combine the relative merits of the two methods. Section 2 briefly reviews greyscale sieves and describes the CCS and VAMS algorithms. Section 3 presents the new colour sieve structure and a comprehensive performance evaluation is undertaken in section 4, including a quantitative evaluation of the sieves' segmentation performance using the Berkeley segmentation dataset. Finally, discussion and conclusions are given in section 5.

2   Colour Morphological Scale-Spaces

A greyscale area open-close (AOC) sieve can be formed by successive area openings and closings of increasing scale [1] and, for the image X, is defined by

$$AOC_i(X) = \varphi_i^a \gamma_i^a (\varphi_{i-1}^a \gamma_{i-1}^a (\ldots (\varphi_2^a \gamma_2^a (\varphi_1^a \gamma_1^a (X))))) \qquad (1)$$

where γ_i^a and φ_i^a are respectively area openings and closings to an area limit i. This sieve structure is also known as an $\mathscr{M}$ sieve [2] and forms a tree as the maxima and minima are merged with their closest greyscale neighbour. In a similar manner, the closings can be performed before the openings, giving rise to an area close-open (ACO) or $\mathscr{N}$ sieve. Combining the maxima and minima in this manner ensures the property of strong causality, with no new extrema being generated as the scale increases. Alternatively, maxima and minima can be treated separately to give the max- and min-trees of [14].

Algorithmically, area morphological scale-space algorithms follow these steps:

1. Identification of extrema regions;
2. Merge all scale 1 extrema regions with their nearest neighbour;
3. Repeat step 2 with increasing scale until only 1 region remains.

The CCS and the VAMS were developed independently and first presented within a few days of each other at Scale-Space 2003 [8] and the 2003 IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing [5] respectively. Although they differ in some aspects, the two techniques follow the same general algorithm:

1. Identification of extrema regions;
2. Merge all scale 1 extrema regions with their nearest neighbour;
3. Repeat steps 1 and 2 until no extrema are found at current scale;
4. Repeat steps 1 to 3 with increasing scale until only 1 region remains.

Comparing the greyscale and colour scale-space algorithms, it can be seen that the latter requires an additional stage (step 3) to ensure itempotence. This is because the process of merging in vector spaces can result in the creation of new extrema in the area of influence of the merged regions. Both the CCS and the VAMS use the Euclidean distance to select the closest region for merging with in Step 2, although they differ in how ties are resolved: the CCS uses luminance, then individual colour channels while the VAMS uses scan order. They also both only process extrema as it is not possible to differentiate between maxima and minima for vector values. Therefore, their main difference lies in the mechanism for determining extreme regions in Step 1 of the algorithm and to provide a comparison between the two techniques their approaches are discussed in detail below.
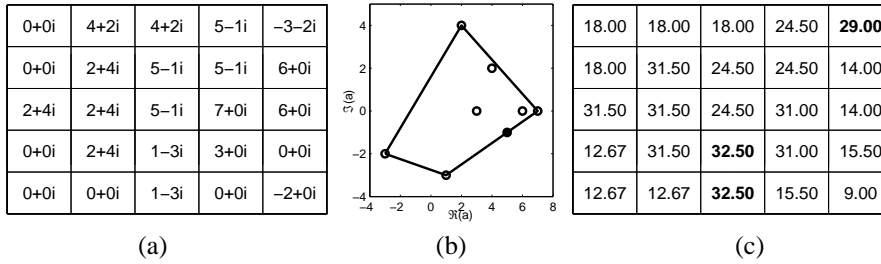
0+0i	4+2i	4+2i	5−1i	−3−2i
0+0i	2+4i	5−1i	5−1i	6+0i
2+4i	2+4i	5−1i	7+0i	6+0i
0+0i	2+4i	1−3i	3+0i	0+0i
0+0i	0+0i	1−3i	0+0i	−2+0i

18.00	18.00	18.00	24.50	**29.00**
18.00	31.50	24.50	24.50	14.00
31.50	31.50	24.50	31.00	14.00
12.67	31.50	**32.50**	31.00	15.50
12.67	12.67	**32.50**	15.50	9.00

(a) (b) (c)

Figure 1: Colour sieves example using 8nn connectivity. (a) Complex image, (b) convex hull for $5 - 1i$ region and (c) VAMS scalar image with extrema marked in bold.

2.1 Determination of Extrema

Unlike greyscale sieves, the multivariate data associated with colour sieves cannot be unambiguously ordered and as the CCS and VAMS use different extrema definitions they produce different extrema for a given image. As colour sieves generate additional extrema at each scale as a result of the region merging process the behaviour of a colour sieve is in a large part determined by the proportion of image regions marked as extreme, which in turn depends on the extrema definition. If a high proportion of extrema regions is found then the repetition of steps 1 and 2 of the algorithm will result in few, if any, regions of area less than the current scale surviving while a low proportion of extrema regions will leave a significant proportion of the image untouched until relatively large scales. In addition, the proportion of extrema regions has a significant effect on the processing time.

The approach to extrema definition adopted by the CCS is one based on a local convex hull. For a set S of points in d-dimensional space, the convex hull is the smallest convex polygon containing all the points of S. The CCS forms a local convex hull for each pixel and its connected neighbours and then defines the pixel as extreme if it lies on the edge of the hull. This approach has the advantage that the topology of the local hull is unaffected by linear axes transformations and monotonic scaling but can result in large proportion of extrema. For example, consider the complex image shown in figure 1(a). The local convex hull for the $5 - 1i$ region in figure 1(b) shows the region to be extreme as it lies on the edge of the hull. However, examination of the local convex hull for the other regions in the image shows that they are all extreme. Part of this problem results from degenerate cases: a d-dimensional hull requires at least $(d + 1)$ different points to be non-degenerate and although an approach to reduce the number of extrema for degenerate cases was presented in [8], they can still correspond to a significant proportion of the total extrema. This result is confirmed by figure 2 which shows the initial extrema for the 186×230 colour test image Lily. The CCS result in figure 2(b) classifies the majority of the image as extrema, a finding confirmed by other images. With 4nn connectivity the number of dimensions is reduced and the CCS, if anything, finds even more extrema. Therefore the action of the CCS is very aggressive and the total number of regions will rapidly decrease with increasing area size. However, classifying the majority of pixels as extreme is counter-intuitive as extrema are associated with outlying values. In addition, there are many connected extrema which cannot simply be explained as alternating maxima and minima, which cannot be differentiated in colour sieves.
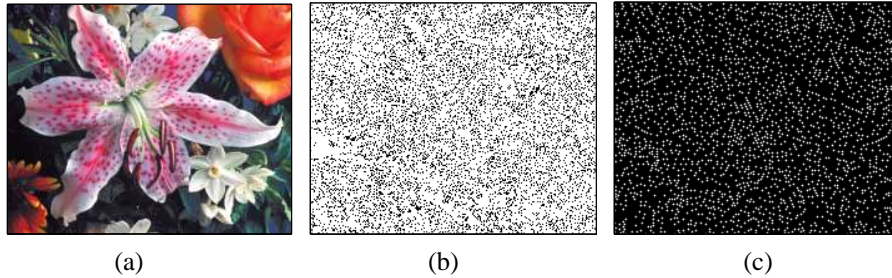
Figure 2: CCS and VAMS extrema (white pixels) using 8nn connectivity. (a) Original image Lily, (b) CCS extrema and (c) VAMS extrema.

The VAMS uses reduced ordering to form a scalar image in which colour extrema correspond to regional maxima. In the scalar image, the value of each pixel is initially given as the sum of the vector differences between the pixel and its connected neighbours. The scalar values for all the pixels in each flat zone are then summed and normalised by the region's area and this value is assigned to all pixels within the region, see figure 1(c). Finally, extrema are identified as the maxima of the scalar image. In figure 1(c) there are only two extreme regions corresponding to the flat zones with values $-3-2i$ and $1-3i$ in figure 1(a) and although it is theoretically possible for neighbouring regions to be marked as extreme if they have exactly the same scalar value in practice this rarely happens. The extrema found by the VAMS for the Lily image are presented in figure 2(c) and confirm the findings of figure 1 that the VAMS produces far fewer extrema than the CCS. As the VAMS extrema correspond to pixels whose colour is very different from their neighbours (either isolated regions or at positions of high gradient), they can therefore be considered as "true" extrema. However, if extrema are viewed as seeds from which the image is altered through merging, then a low extrema proportion may result in much of the image being unaffected by the sieving process until larger scales.

3   Vector area morphology open-close sieve

Summarising the previous section, the CCS has many extrema and an aggressive sieving action while the VAMS finds fewer, more meaningful extrema and is less aggressive. In addition, while the convex hull used by the CCS is quite inflexible with regards to its extrema definition, changes can be made to the VAMS structure to try and combine the advantages of the two methods. To this end, the vector area morphology open-close sieve (VAMOCS) is proposed that applies both area openings and closings to the scalar surface produced by the VAMS. While the maxima in the scalar surface correspond to image extrema, local minima mark regions that are closer in value to their neighbouring regions than other regions in their connected neighbourhood. Processing the minima essentially merges regions that are in relatively flat parts of the image in a manner reminiscent to that of [15], which introduces a bound Δ on the allowable greyscale fluctuations within a flat zone. Providing the merging is handled sensibly closings can increase the numbers of seeds without adversely affecting the segmentation performance. Here, the merging rules

Figure 3: Colour sieve results for Lily test image using 8nn connectivity. Top to bottom: CCS, VAMOCS (openings only), VAMOCS (closings only) and VAMOCS (combined openings and closings). Left to right: area = 10,100,1000 and 10000.

proposed by Salembier and Garrido [14] are used for closings. The VAMOCS algorithm also has an additional minor modification in which the sum of scalars for the pixels within a flat zone is normalised by its perimeter rather than its area. As the distance between all pixels within a flat region is zero, this approach effectively calculates the average vector difference per unit perimeter for all pixels on the perimeter of a region and makes the sieving action less dependent on the complexity of the regions' geometry.

Figure 3 shows the results of sieving the Lily image at selected scales using the CCS and the VAMOCS. Also shown are the VAMOCS results using openings and closings only (rows 2 and 3 of figure 3 respectively). Separate openings and closings both preserve colour edges while creating large flat regions although their filtering actions differ, with the former removing outliers and the latter leaving islands of extrema as the relatively flat regions are extended. The combined VAMOCS results shows the openings and closings to compliment each other, producing a colour sieve with an action similar in aggressiveness to the CCS. However, comparing the VAMOCS extrema in figure 4(a) with those of figure 2 shows the VAMOCS to produce many fewer extrema than the CCS and roughly equal in number to the greyscale extrema shown figure 4(b). The VAMOCS therefore appears to have achieved its aims and a comprehensive performance analysis of all the colour sieves is undertaken in the next section.
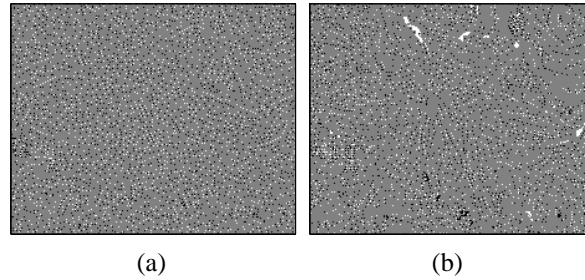
Figure 4: Extrema for Lily test image produced by (a) VAMOCS and (b) greyscale AOC sieve. Maxima shown in white and minima in black.

4 Performance Evaluation

To evaluate the performance of the colour sieves, a number of criteria were used including processing time and the proportion of the image regions defined as extreme. One of the most useful applications of these sieves is colour image segmentation and an initial study of the segmentation performance of the CCS was undertaken in [9]. Here, the Berkeley segmentation dataset is used to provide the basis for a quantitative evaluation. The colour sieves were implemented in C++ using an approach based on the pixel-queue algorithm [12] and then run as mex files under Matlab. Eight nearest neighbours connectivity and the Euclidean distance metric were used for all sieves.

The first aspect to be considered is the processing time, given in figure 5, which also includes the greyscale AOC (GS-AOC) sieve for comparative purposes. As the processing times are related to the number of extreme regions, each of which has to be merged and updated, the percentage of extrema for each sieve is also plotted, with exact values for selected scales given in table 1. As expected, the greyscale sieve has the lowest processing time, although the VAMS is only fractionally slower for areas < 100. The CCS has the highest processing time reflecting its high proportion of extrema, although its rapid reduction in the total number of regions produces a relatively constant processing time for areas > 100. In contrast, the processing time for the VAMS increases with scale and at

scale	1	2	10	50	100	500	1000
VAMS	2209	1054	151	21	8	1	1
	41207	37199	27547	19660	16102	8024	6871
CCS	33616	12852	2211	456	223	44	23
	41207	16402	2878	572	279	53	25
GS-AOC	4062	1973	424	110	63	11	7
	41207	36553	28457	20697	17040	9114	6150
VAMOCS	4388	2353	476	112	16	8	9
	41207	33882	14326	3856	1845	298	132

Table 1: Variation of proportion of image extrema with scale. Fractions give the number of extreme regions the total number of regions.
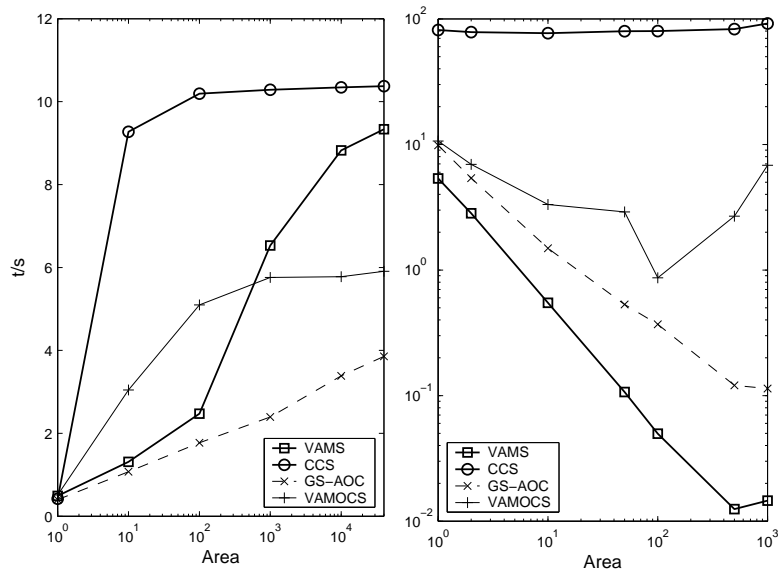
Figure 5: Processing times and proportion of extrema regions for Lily test image.

high scales approaches that of the CCS. The VAMOCS is more expensive than the VAMS for scales < 100 as it has more extrema to process. However, as it reduces the total number of regions more rapidly (see table 1), larger scales require little extra processing.

An objective measure of segmentation performance is obtained by using the Berkeley Segmentation Dataset and Benchmark available at http://www.cs.berkeley.edu/projects-/vision/grouping/segbench [10]. The ground truth for each image in the dataset is given as the collection of boundaries produced by all human subjects. A quantitative performance measure is provided by precision-recall (P-R) curves, where precision is the probability of boundary pixels being correctly identified and recall is the amount of boundary pixels detected. The F-measure is given by the harmonic mean of precision and recall along the curve, with its maximum providing a single measure segmentation performance [11]. In addition, it allows both edge- and region-based methods to be compared and avoids the trivial cases where the GCE of [10] gives zero error. Unlike [9] where the evaluation takes the most representative set of regions near the root of the tree, trimming branches to the level (attribute) required, here the entire tree is sieved to the same level. Although suboptimal in terms of segmentation performance, this allows a focus on the global threshold while maintaining itempotency.

P-R curves for 100 images from the dataset were generated by sieving each image with increasing scale until the total number of regions was $\leq n$ and plotting precision and recall for a range of values of n. A fixed number of regions was used in preference to a constant area (or other attribute) since it is less dependent on image content and also allows comparison with other attributes. This approach is also compatible with the dataset definition where a small $(2 - 30)$ number of equally important regions is suggested.

The P-R curves for one image from the dataset produced by the GS-AOC sieve, CCS and VAMOCS are shown in figure 6, with the point producing the maximum F-measure marked in bold. To help visualise these results, figure 7 presents the corresponding human
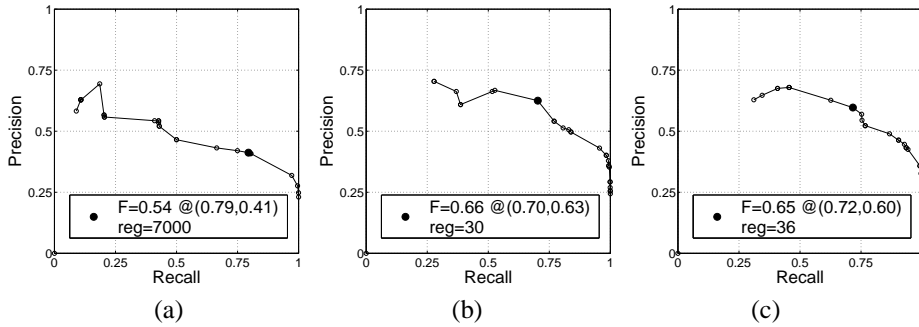
Figure 6: P-R curves for (a) GS-AOC Sieve, (b) CCS and (c) VAMOCS. The maximum F-measure and number of regions at which it occurred are also shown.

and colour sieve segmentations. These results are for just one image and a more comprehensive comparison is given by averaging the P-R curves and peak F-measure for all 100 images, see table 2. The target number of regions at which the peak F-measure occurred is also shown as a technique that can achieve a high F-measure with the minimum number of regions is preferable to one with more regions. The results in table 2 show that both the CCS and VAMOCS out-perform the greyscale sieve, although the VAMS does not. The VAMOCS has the highest average F-measure of 0.51, achieved with a target number of 50 regions. Table 2 also presents a set of results for the contrast attribute. These follow the same trend as the area attribute with the VAMOCS result being 0.02 higher than the CCS, albeit with an increased number of regions.

Method	Attribute	F-measure (R,P)	Regions
GS-AOC		0.46 (0.50,0.44)	300
CCS	Area	0.49 (0.61,0.41)	30
VAMS		0.40 (0.51,0.32)	100
VAMOCS		0.51 (0.61,0.43)	50
GS-AOC		0.47 (0.55,0.41)	1000
CCS	Contrast	0.50 (0.65,0.41)	30
VAMS		0.35 (0.42,0.30)	500
VAMOCS		0.52 (0.65,0.43)	90

Table 2: Figure of merit comparison using the Berkeley dataset

5   Discussion and conclusions

Two recently proposed colour morphological scale-space sieves have been evaluated algorithmically and in terms of their definition of extrema. The link between proportion of image regions defined as extreme and the processing speed was also investigated. The CCS was found to produce large numbers of extrema and show an aggressive sieving action whereas the VAMS has fewer extrema and therefore a lower computational cost. To
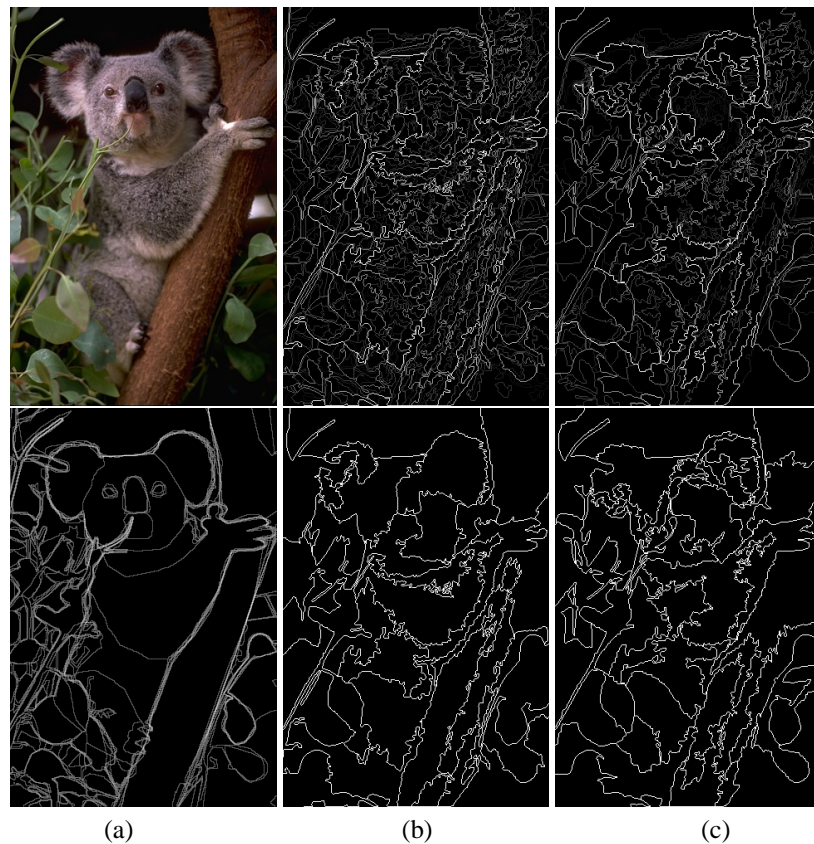
Figure 7: Colour sieve segmentation of Koala image. (a) Original image and collection of human segmentations, (b) and (c) CCS and VAMOCS results for different target number of regions (top) and best result (bottom).

combine their relative advantages, the VAMOCS was proposed. The VAMOCS works by also processing image minima which correspond to regions in "nearly flat" parts of the image. For scales >100, the VAMOCS has the lowest processing time of all colour sieves despite its aggressive action.

A major application area of greyscale connected sieves has been image segmentation and to assess any advantage gained by applying colour sieves to this task a quantitative evaluation of its segmentation performance was undertaken using the methodology of [11]. The new VAMOCS produced the best average segmentation performance over 100 images, showing the benefits conferred by using colour. Although its overall performance falls short of that of state-of-the-art colour segmentation techniques such as the combined brightness/colour/texture gradients of [11], the potential of the VAMOCS for segmentation has been demonstrated. An improved segmentation performance can be achieved by more advanced post-processing of the colour tree and algorithms for this, and the development of improved attributes, are areas of current research.

References

[1] S.T. Acton and D.P. Mukherjee. Scale-space classification using area morphology. *IEEE Trans. Image Processing*, 9(4):623–635, April 2000.

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

[3] M.L. Comer and E.J. Delp. Morphological operations for color image processing. *Journal of Electronic Imaging*, 8(3):279–289, 1999.

[4] E.R. Dougherty, J.T. Newell, and J.B. Pelz. Morphological texture-based maximum likelihood pixel classification based on local granulometric moments. *Pattern Recognition*, 25(10):1181–1198, 1992.

[5] A.N. Evans. Extending area morphology to multivariate images. In *Proc. 6th IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, June 2003.

[6] A.N. Evans. Vector area morphology for motion field smoothing and interpretation. *IEE Proc. Vision, Image and Signal Processing*, 150(4):219–226, August 2003.

[7] D. Gatica-Perez, C. Gu, M.T. Sun, and S. Ruiz-Correa. Extensive partition operators, gray-level connected operators, and region merging/classification segmentation algorithms: theoretical links. *IEEE Trans. Image Processing*, 10(9):1332–1345, September 2001.

[8] S. Gibson, R. Harvey, and G.D. Finlayson. Convex colour sieves. In *Proc. 4th International Conference on Scale Space Methods in Computer Vision*, volume LNCS 2695, pages 550–563, 2003.

[9] S. Gibson, R. Harvey, and G.D. Finlayson. Evaluating a colour scale-space. In *Proc. British Machine Vision Conference*, June 2003.

[10] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. 8th Int'l Conf. Computer Vision*, volume 2, pages 416–423, July 2001.

[11] D.R Martin, C.C. Fowlkes, and J. Malik. Learning to detect natural image boundaries using local brightness, colour and texture cues. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26(5):530–549, May 2004.

[12] A. Meijster and M.H.F. Wilkinson. A comparison of algorithms for connected set openings and closings. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(4):484–494, April 2002.

[13] K.R. Park and C.N. Lee. Scale-space using mathematical morphology. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(11):1121–1126, November 1996.

[14] P. Salembier and L. Garrido. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Trans. Image Processing*, 9(4):561–576, April 2000.

[15] P. Salembier, A. Oliveras, and L. Garrido. Antiextensive connected operators for image and sequence processing. *IEEE Trans. Image Processing*, 7(4):555–570, April 1998.

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