Vector area morphology for motion field smoothing and interpretation

A.N. Evans

Abstract: A new nonlinear technique for filtering motion fields and other multivariate data is introduced. The method is developed from mathematical morphological area openings and uses a vector to scalar transform, in which each vector is replaced by the sum of the distances to its connected neighbours, to control the growth of extrema regions. As the filter either perfectly preserves or completely removes image components, it is able to remove noise without altering significant features. In addition, at larger area sizes a meaningful interpretation of the underlying structure is achieved. Results show that the vector area morphology sieve performs well in comparison to the widely used vector median filter.

1 Introduction

Nonlinear filters have long been established as a popular class of methods for many digital imaging applications. Arguably the most widely used nonlinear technique is the median filter, mainly due to its ability to remove noise without excessive edge smoothing or generating any values not present in the original image. Many of the techniques that have been developed for scalar-valued images can also be applied to multivariate or vector data such as that found in motion fields, colour images and multi- and hyper-spectral images. However, the extension of nonlinear methods to multivariate data is not straightforward. For example, applying a median filter to each component of an RGB image can produce vectors that were not present in the original image and edge jitter. These problems are overcome with the development of the vector median (VM) filter for multivariate signals [1] that has been used in application areas such as colour image processing and motion field regularisation [2, 3]. Unfortunately the VM filter also has the same lack of controllability as the scalar median filter, the only tuning parameter being the mask size. A finer degree of control is provided by the weighted vector median (WVM) filter developed by Viero *et al*. [4] which operates in a similar manner to the weighted median (WM) filter. This adds an additional degree of control but requires a suitable mechanism for determining the weights.

Another disadvantage of the median filter is that it is not idempotent unless the input is a root signal which results after $n/2$ repeated applications, where n is the filter length. Alternatively, the property of idempotency is exhibited by another class of nonlinear operators, namely mathematical morphology (MM) openings and closings. This is one of the many reasons underpinning the widespread popularity of MM for noise reduction, shape and multi-scale analysis, edge detection and many other applications for processing and analysing binary and greyscale images.

The extension of MM operations to colour image processing is proposed in [5]. To perform the morphological operations each multivariate value is reduced to a scalar using a distance metric, resulting in a reduced ordering. This approach relies on the use of a structuring element, the shape and size of which influences the output image. Further, to successfully remove noise it is necessary to apply multiple operators consisting of different structuring elements for both the opening and closing stages, selecting the maximum and minimum outputs, respectively.

For single-valued images the need for multiple structuring elements is overcome with the development of connected set granulometries [6–9]. Of these, area openings and closings [7, 8], later generalised as attribute openings [10], have proved a useful tool for image analysis and classification [11], and provide a fine degree of control over the filtering action [12]. Compared with traditional morphological openings, area openings have the advantage of not being based on a fixed structuring element. Instead, they remove all maxima connected components of an image with an area less than the area limit λ and have been shown to be equivalent to a maximum of openings with all possible connected structuring elements with λ elements. Greyscale area openings remove all light structures up to the area limit, and area closings operate in a similar manner on dark structures. Successive applications of increasing scale area openings and closings gives rise to morphological area sieves that have the advantage of being able to use area or any other attribute to control the sieving action [10]. At each sieve scale the image has extrema regions that are flat zones at least equal to the current area size, with other image components completely preserved. The extension of area morphology to vector images therefore has the potential for providing a finer degree of control over filter action than the VM filter without the requirement for multiple structuring elements of existing morphological methods for vector images.

This paper describes a new approach for applying an area sieve to vector images. The filtering action is controlled by a transform in which each vector is replaced by the sum of the distances to its connected neighbours. This is more powerful

IEE Proceedings online no. 20030521

doi: 10.1049/ip-vis:20030521

Paper first received 17th July 2002 and in revised form 23rd January 2003

The author is with the Department of Electronic and Electrical Engineering, University of Bath, Bath BA2 7AY, UK

than the distance transform of [5] where the Euclidean norm of each pixel is the metric used for the reduced ordering as it takes into account both the magnitude and direction of the multivariate values. In the transformed image the vector extrema can be identified as those positions with a sum of distances greater than any in a local neighbour, defined using local 4 of 8 nearest neighbour connectivity. This contrasts with the definition of the vector median, which is the vector whose sum of distances is the minimum in the local neighbourhood. Once identified, those extrema regions that are smaller than the current sieve size λ are replaced by the closest vector from the connected neighbourhood, creating flat zones in the multivariate image.

2 Vector median filtering

The VM filter exhibits the same desirable properties of the scalar median filter but also suffers from some of its drawbacks, in particular the relatively coarse degree of control provided by the selection of mask size and shape.

Given a set of N vectors $\mathcal{V}$ the vector median $\vec{\boldsymbol{x}}_{VM} \in \mathcal{V}$ is defined by

$$\sum_{\vec{\boldsymbol{x}}_i \in \mathcal{V}} \|\vec{\boldsymbol{x}}_{VM} - \vec{\boldsymbol{x}}_i\|_p \leq \sum_{\vec{\boldsymbol{x}}_i \in \mathcal{V}} \|\vec{\boldsymbol{x}}_j - \vec{\boldsymbol{x}}_i\|_p \quad \forall j \in \mathcal{V} \tag{1}$$

where the norm $\|\cdot\|_p$ defines the metric used to convert from a vector to a scalar value. In practice the city-block distance or the Euclidian distance is used, for $p = 1$ or 2, respectively.

Direct calculation of (1) requires the distance between all possible vector pairs to be computed, which is too computationally expensive for most practical applications. To reduce the complexity, a faster running algorithm based on the samples entering and leaving the window at each location can be used. An alternative strategy is the development of fast algorithms for minimising (1). The latter approach has been used for both the 1-norm [13] and the squared Euclidean norm. A thorough review of the computational performance of fast algorithms is given in [14]. To date, no fast calculation has been found for the Euclidean norm, which is the best performing metric. Instead, effort has been put into the development of an efficient pseudo-norm with only slightly reduced performance [15].

The only free parameter available for tuning the VM is the mask size. A finer degree of control is achieved by the WVM filter, generalised from the scalar case in [4] and extended to real-valued weights by Alparone *et al.* [3]. Given a set of weights w_i, $i = 1, 2, \ldots, N$, the WVM $\vec{\boldsymbol{x}}_{WVM}$ is defined by

$$\sum_{\vec{\boldsymbol{x}}_i \in \mathcal{V}} w_i \|\vec{\boldsymbol{x}}_{WVM} - \vec{\boldsymbol{x}}_i\|_p \leq \sum_{\vec{\boldsymbol{x}}_i \in \mathcal{V}} w_i \|\vec{\boldsymbol{x}}_j - \vec{\boldsymbol{x}}_i\|_p \quad \forall j \in \mathcal{V} \tag{2}$$

An alternative derivation of the weights for use as a motion field post-filter within a video codec is given by a ratio of displaced frame difference values in [16]. The application area for this latter scheme is the integration of motion field post-processing within an H.263 video codec. However, to achieve any overall coding gains a conditional decision on whether to retain the original vector or replace with a filtered vector is required.

3 Vector area morphology

Area morphology belongs to the attribute-based class of operators. Unlike basic morphological operators, area morphology does not require a fixed structuring element. Instead, it adapts to the contents of the image using an attribute, for example area, contrast or complexity, to control which image components are removed. A greyscale area open (AO) is given by

$$\gamma_\lambda(X) = \bigvee_{B \in A_\lambda} (X \circ B) \tag{3}$$

where A_λ is the set of connected subsets whose area is $\geq \lambda$. For this function to be an area opening, it must also satisfy the increasing criteria:

$$X \subseteq Y \Rightarrow \Psi(X) \subseteq \Psi(Y) \tag{4}$$

An area open sieve given by

$$AO_\lambda(X) = \gamma_\lambda(\gamma_{\lambda-1}(\ldots(\gamma_2(\gamma_1(X))))) \tag{5}$$

successively increases the sieve scale to remove connected maxima of increasing area, producing a size or granulometry distribution, and creating flat zones in the filtered image.

As only regional maxima are affected by an area opening, pixels that do belong to maxima need not be processed. This fact is used by Vincent to develop an algorithm for the efficient computation of area openings [7], which can easily be modified to the area sieve form, as shown in Fig. 1. This implementation provides the starting point for the development of the VAM algorithm. For a recent review of this and other attribute openings and closings implementations see [17].

The aim of VAM is identical to that of the scalar case: to remove extrema regions less than a given area size, the difference being that the components removed and flat zones created consist of vectors. In area morphology the greyscale intensity value is used as the criteria to determine regional minima and maxima. As there is no natural mechanism for ordering multivariate data the local extrema are not well defined. The standard approach is marginal ordering, also

```
1  Extract the regional maxima of I.
2  For area λ = 1 to max:
   2.1  For each regional maximum m do:
      2.1.1  If area of m > λ, go to next maximum
      2.1.2  Scan the neighbours of m and select the one
             with the highest greyscale value.
      2.1.3  If the greyscale value is greater than that
             of m, remove form list.
      2.1.4  Else, set all pixels of m to the greyscale
             value.
```

Fig. 1 *Area open sieve algorithm*

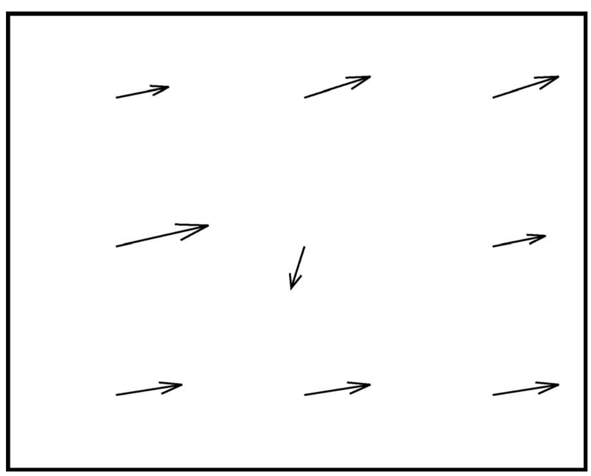

Fig. 2 *Example motion field*

0 0	1 0	5 -2	3 -1	2 0	1 0	3 0	28 24	2 0	2 0	2 0
0 0	0 0	-22 10	2 1	2 0	8 -4	2 -1	1 -1	2 0	2 0	2 0
0 0	5 -2	3 -1	1 0	1 0	3 0	1 -1	2 0	2 0	2 0	2 0
0 -1	0 0	2 -1	3 -1	1 -1	1 -1	2 -1	2 0	2 0	-12-12	2 0
0 -3	3 0	3 -1	2 -1	1 -1	1 -1	2 -1	2 0	-18-14	-22-20	1 -1
0 -9	4 2	1 -1	2 -1	2 -1	1 -1	2 0	2 0	3 0	2 0	2 0
0 3	2 -2	2 -1	2 -1	1 -2	2 -1	2 0	2 0	1 -1	1 0	-10 -8
0 -1	1 -1	2 -1	2 -1	2 -1	2 -2	2 -1	2 0	1 0	1 0	1 0
0 0	3 -2	2 -1	2 -2	2 -2	2 -2	2 -1	0 -2	1 -1	1 0	1 0

a

1	42	61	46	14	17	65	252	52	0	0
8	51	281	51	17	79	69	64	52	0	0
8	80	54	45	21	25	19	6	28	26	26
14	25	14	12	8	6	8	38	104	180	72
26	34	14	5	5	5	7	37	223	291	72
54	53	17	5	5	8	6	38	89	104	67
33	31	8	3	14	6	6	5	14	24	97
13	16	5	4	6	7	8	11	7	21	19
8	16	4	4	2	3	6	15	7	1	0

b

1				2				3	3	3
1	1			2			4	3	3	3
1		5	6	6		4	3	3	3	3
	1	7	5	4	4	8	3	3		3
		5	7	4	4	8	3			
			7	7	4	3	3		9	9
		7	7		7	3	3		10	
		7	7	7	11	7	3	10	10	10
		7	11	11	11	7			10	10

c

19	42	61	46	16	17	65	252	27	27	27
19	19	281	51	16	79	69	16	27	27	27
19	80	27	33	33	25	16	27	27	27	27
14	19	6	27	16	16	8	27	27	180	27
26	34	27	6	16	16	8	27	223	291	72
54	53	17	6	6	16	27	27	89	86	86
33	31	6	6	14	6	27	27	14	12	97
13	16	6	6	6	4	6	27	12	12	12
8	16	6	4	4	4	6	15	7	12	12

d

Fig. 3 *Vector to scalar transform process*

a Original motion field (see Fig. 5*a* for vector representation)
b Vector to scalar transform of *a* using (7) with 8 nearest neighbour connectivity and $p = 1$
c Flat zones
d Modified transform values

1 Extract the regional maxima of $d[\dot{x}_i]$.

2 For area λ = 1 to max:

  2.1 While min area of regional maxima < λ:

    2.1.1 For each regional maximum m do:

      2.1.1.1 If area of $m > \lambda$, go to next maximum.

      2.1.1.2 Scan the neighbours of m and select the one that minimises $||v_c - v_n||_p$.

      2.1.1.3 Set all vectors in m equal to closest neighbour and update $d[\vec{x}_i]$.

      2.1.1.4 If m is no longer a regional maximum, remove from list.

      2.1.1.5 If any new maximum regions are created, add to list.

Fig. 4 *VAM sieve algorithm*

known as component-wise filtering, in which the morphological filter is applied to each individual channel. When the channels are correlated, for example in the case of colour images, the decision to remove or enhance components in each channel is taken on an individual basis, and this can result in edge jitter [1].

The approach adopted here is to use a scalar image derived from the original vector data to control the filtering of the original vector data. A similar approach is proposed by Comer and Delp for noise reduction in colour images using the Euclidean norm as the metric to produce the scalar image [5]. The shortcomings of using the Euclidean norm as the metric for each pixel can be illustrated by a simple example – see Fig. 2. The vector with the greatest magnitude and therefore the largest Euclidean norm can clearly be identified (centre left). However, on inspection of the vector set the vector in the centre clearly stands out as its direction is

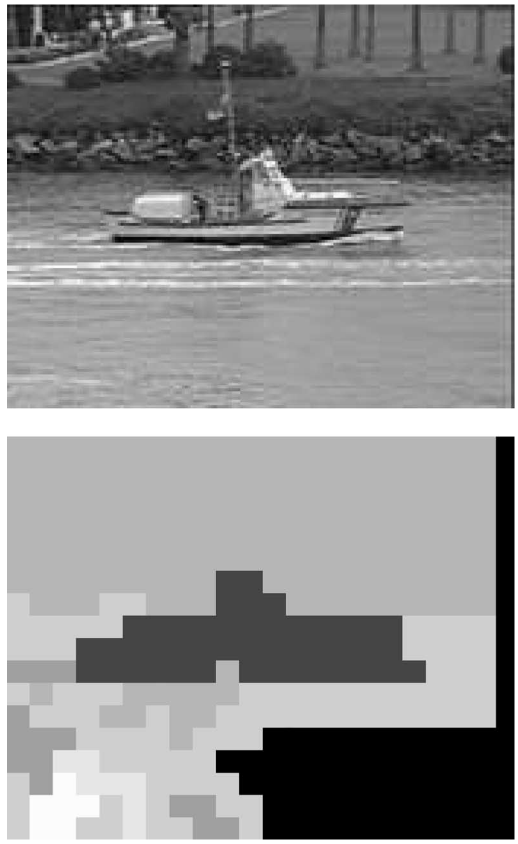

Fig. 6 *Original image and flat zones from VAM sieve for $\lambda = 17$*

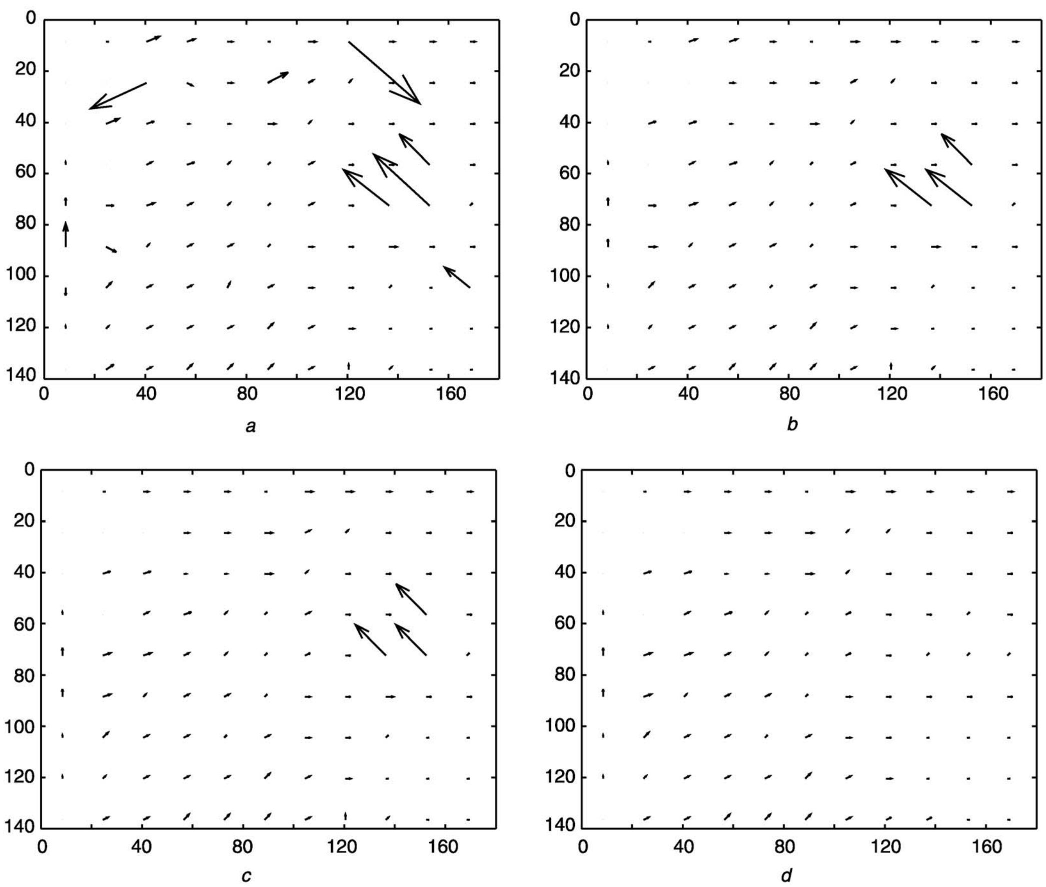

Fig. 5 *Vector area sieve results*

a Original motion field
b Sieve result for $\lambda = 2$
c $\lambda = 3$
d $\lambda = 4$

Fig. 7 *Vector area sieve results for (in raster order) area = 1, 2, 3, 4, 5 and 6*

markedly different from the others; this vector is a natural outlier and can therefore be considered the local extremum.

In (1) the vector median is defined as the vector that has the minimum distance to all other vectors in the set. In this spirit, the vector extremum $\vec{\boldsymbol{v}}_{VE}$ can be defined as the vector that is furthest from its neighbours, given by

$$\sum_{\vec{\boldsymbol{x}}_i \in \mathcal{V}} \|\vec{\boldsymbol{x}}_{VE} - \vec{\boldsymbol{x}}_i\|_p \geq \sum_{\vec{\boldsymbol{x}}_i \in \mathcal{V}} \|\vec{\boldsymbol{x}}_j - \vec{\boldsymbol{x}}_i\|_p \quad \forall j \in \mathcal{V} \tag{6}$$

For VAM, the right hand side of (6) can be used as a distance transform $d[\vec{\boldsymbol{x}}_i]$ that maps each $\vec{\boldsymbol{x}}_i$ to a scalar by replacing each vector by the sum of the distances to its connected neighbours. Thus,

$$d[\vec{\boldsymbol{x}}_i] = \sum_{\vec{\boldsymbol{x}}_j \in \mathcal{N}_i} \|\vec{\boldsymbol{x}}_j - \vec{\boldsymbol{x}}_i\|_p \tag{7}$$

where $\mathcal{N}_i$ is the set of connected neighbours of $\vec{\boldsymbol{x}}_i$.

Although (7) provides an attractive mechanism for identifying natural outliers it does have some disadvantages. One of these disadvantages results from the transform being a many-to-many mapping: different vectors can have the same transform values and identical vectors can have different transform values. This contrasts with the transform of [5], which is a many-to-one-mapping. Scalar area openings work by processing (enlarging) those flat zones that are regional maxima. To ensure that each flat zone in $d[\vec{\boldsymbol{x}}_i]$ is treated as a single entity, all vectors belonging to a flat zone are assigned the mean value of the zone. This process is illustrated in Fig. 3. Fig. 3*a* is a motion field produced by a rate constrained video codec from two frames of the well known foreman sequence. The vector to scalar transform of the field produced by (7) is shown in Fig. 3*b* and its flat zones in Fig. 3*c*. Finally, the modified transform values are given in Fig. 3*d*.

Once this transform has been performed, regional extrema can be identified in the same manner as for scalar area morphology. The extrema regions consist of local maxima or minima vectors as both of these will give rise to peaks in the distance transform surface. Therefore, unlike

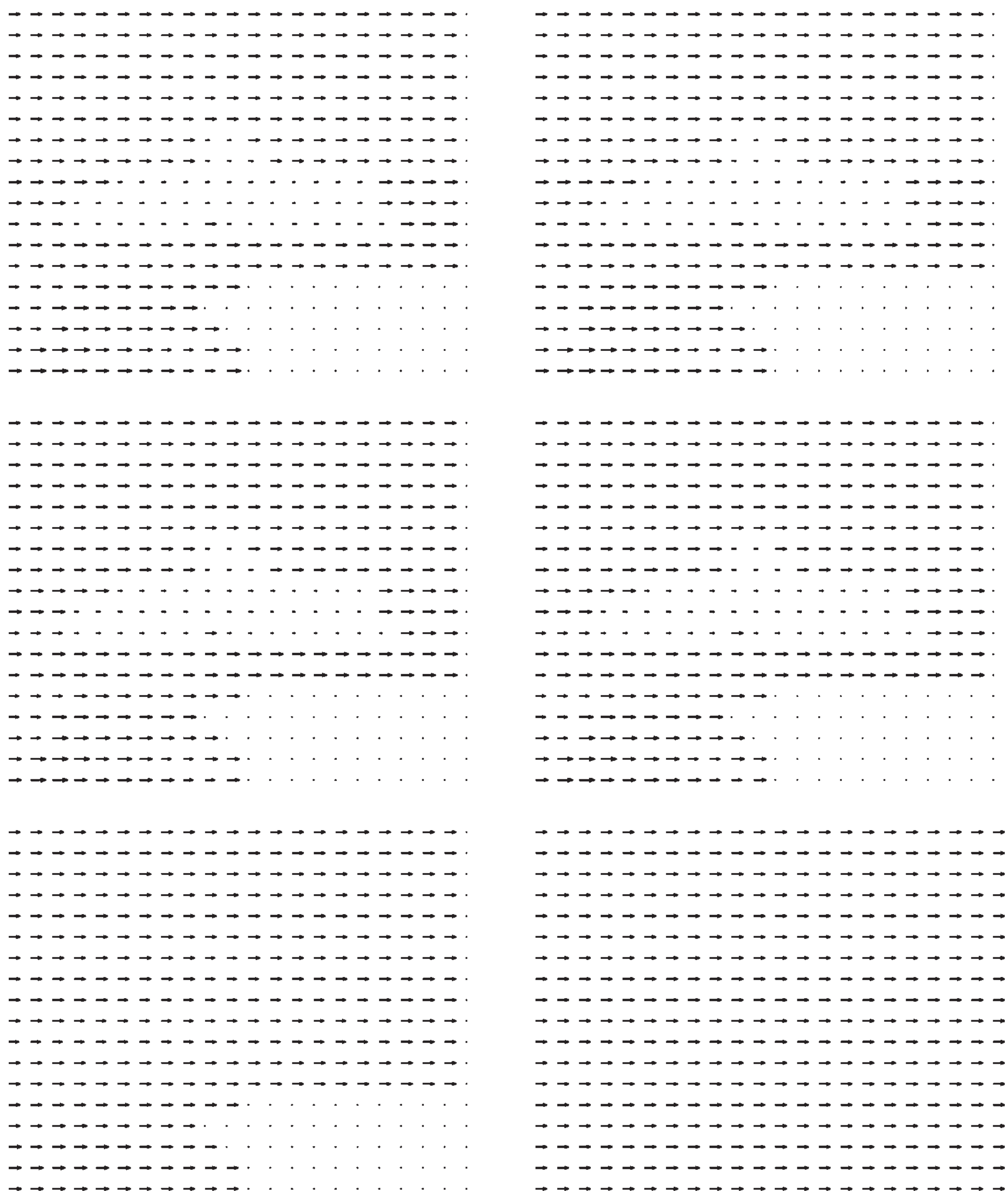

Fig. 8 *Vector area sieve results for (in raster order) area = 7, 8, 12, 17, 64 and 72*

scalar area morphology, which requires both opening and closings, VAM only requires area openings.

The next step of the area sieve algorithm in Fig. 1 is to assign the pixels in each maxima to the greyscale values of the neighbour that is closest to it. The analogue in VAM is to replace those vectors that constitute regional maxima in $d[\vec{x}_i]$ by the closest vectors from the connected neighbourhood of the maxima, determined by a norm such as the Euclidean distance. This is easily achieved in practice as the neighbours of each regional maximum are held in a list for computational purposes [7, 17]. However, vector replacement will also alter the $d[\vec{x}_i]$ value of the regional maxima's neighbours. As a consequence, after updating $d[\vec{x}_i]$, two checks need to be performed: firstly, to determine if each region is still a maximum and, secondly, whether any new maxima have been created. For the second check the only positions that need to be tested are those whose $d[\vec{x}_i]$ value has changed and their neighbours. These are the neighbours of the original regional maxima and the neighbours of the neighbours. The possibility of creating new maxima with increasing scale is a property that has been found elsewhere [18]. At each scale the area of any newly created maxima is increased until it either meets the current area limit or the region is no longer a maximum.

The algorithm for a VAM sieve, incorporating the above considerations, is given in Fig. 4. To illustrate its operation it has been applied to the motion field of Fig. 3*a*, and the results are presented in Fig. 5. As can be seen, as the scale increases maxima regions are removed and replaced with the most similar vector from the regions' neighbours, creating flat zones equal to the current area size λ.

4 Experimental results and discussion

To illustrate the effectiveness of the VAM sieve for motion field smoothing and interpretation it is applied to a motion field produced by two frames from towards the end of the 'Coastguard' sequence. The first of the frames is shown in Fig. 6. Block matching was used to generate a motion field with a block size of 8×8 and a search range was

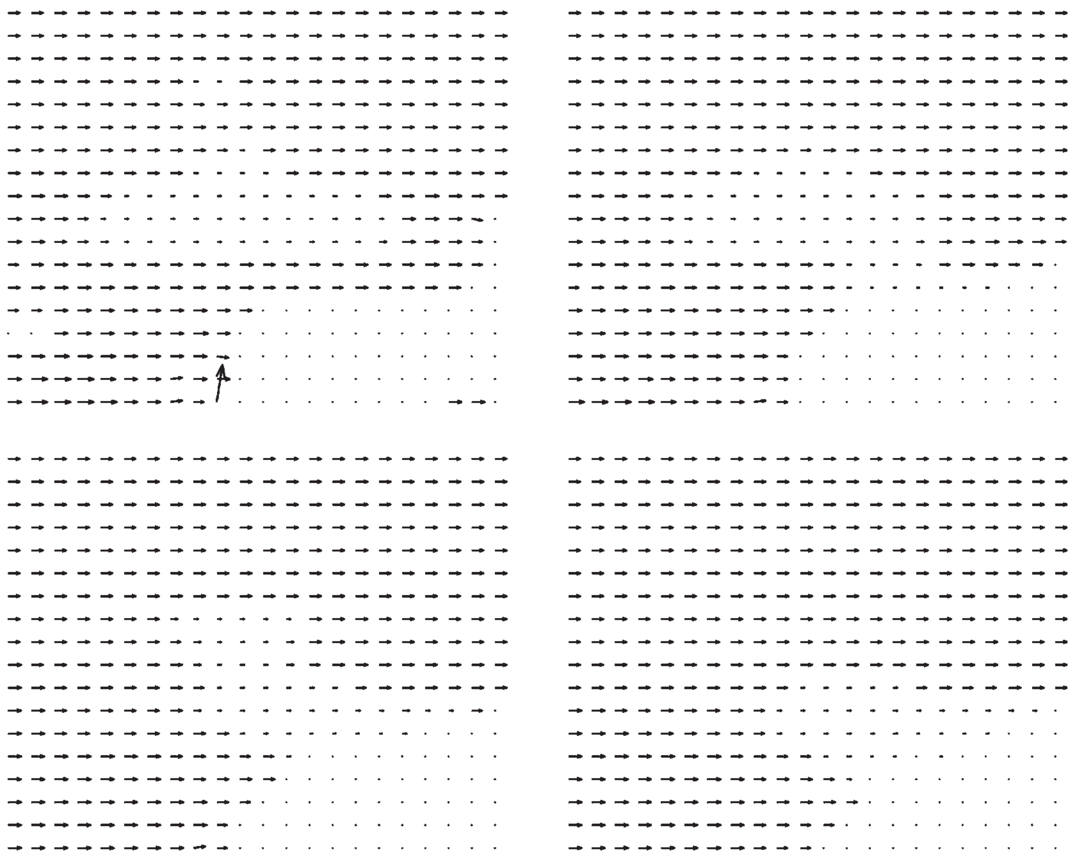

Fig. 9 *Median filter results for 'Coastguard' motion field of Fig. 7 top left*
Mask size (in raster order) 3, 5, 7 and 9

± 16 pixels, to a $\frac{1}{2}$ pixel resolution. The original motion field and the VAM sieve results for each area size that removed a component from the motion field are shown in Figs. 7 and 8. Some structure is discernible in the original motion field but it is corrupted by many noisy vectors. As the sieve scale increases, successively larger flat zones that constitute regional maxima are removed. At a scale of 4 the motion field is 'cleaned' and provides a de-noised interpretation of the original field that is both smooth and consistent. Note that as a VAM sieve either completely removes or completely preserves structures the larger scale boundaries are not adversely affected by the filtering. This contrasts with vector median filtering (see Fig. 9), where the removal of noise also modifies the boundaries present in the motion field, thus changing the shape of objects undergoing a common motion.

As the scale increases, the VAM sieve provides a more meaningful motion field interpretation, with each flat zone corresponding to a region of uniform motion in the original field. Successively larger components are removed up until scale $s = 72$, when the whole image is assigned the vector (4, 0), which gives the global motion. At an area size of $\lambda = 17$, seven vectors are present and the flat zones for each are shown in comparison with the original image in Fig. 6. The boat in the centre of the image is clearly defined and has a motion vector of (1.5, 0). The vector median filter with increasing mask size (see Fig. 9) completely fails to achieve this interpretation.

These results show that the VAM sieve performs well at motion field smoothing and interpretation. There are two main factors behind its success. The first of these is the use of the connected neighbours for each vector to generate the scalar surface from the vector values. This contrasts with the use of a vector norm which, as a magnitude operation, essentially discards any information on directional information. Secondly, the area opening filter structure provides a much finer degree of control over the filter action. The difficulties in extending the area opening approach to multivariate data are successfully overcome, as detailed in the algorithm description, to give a consistent vector image interpretation. The results presented compare very favourably with those of VM filtering.

Although the work presented here is for motion fields, the techniques are also applicable to multivariate data of any size and/or any number of dimensions, for example colour images and multi-spectral imagery. In addition, just as with scalar area and closings, the use of other attributes to control the filter action is possible.

5 References

1 Astola, J., Haavisto, P., and Neuvo, Y.: 'Vector median filters', *Proc. IEEE*, 1990, **78**, pp. 678–689
2 Vardavoulia, M.I., Andreadis, I., and Tsalides, P.: 'A new vector median filter for colour image processing', *Pattern Recognit. Lett.*, 2001, **22**, (6, 7), pp. 675–689
3 Alparone, L., Barni, M., Bartolini, F., and Caldelli, R.: 'Regularization of optic flow estimates by means of weighted vector median filtering', *IEEE Trans. Image Process.*, 1999, **8**, (10), pp. 1462–1467
4 Viero, T., Oistamo, K., and Neuvo, Y.: 'Three-dimensional median-related filters for colour image sequence filtering', *IEEE Trans. Circuits Syst. Video Technol.*, 1994, **4**, (2), pp. 129–142
5 Comer, M.L., and Delp, E.J.: 'Morphological operations for color image processing', *J. Electron. Imaging*, 1999, **8**, (3), pp. 279–289
6 Salembier, P., and Serra, J.: 'Flat zones filtering, connected operators and filters by reconstruction', *IEEE Trans. Image Process.*, 1995, **4**, (8), pp. 1153–1160
7 Vincent, L.: 'Morphological area openings and closings, their efficient implementation and applications'. Proc. EURASIP Workshop on Mathematical morphology and its application to signal processing, Barcelona, Spain, May 1993, pp. 22–27

8 Vincent, L.: 'Morphological area openings and closings for grey-scale images', Proc. NATO Shape in Picture Workshop: Mathematical description of shape in grey-level images, 1993, pp. 196–208
9 Bangham, J.A., Harvey, R., Ling, P.D., and Aldridge, R.V.: 'Morphological scale-space preserving transforms in many dimensions', *J. Electron. Imaging*, 1996, **5**, (3), pp. 283–299
10 Breen, E.J., and Jones, R.: 'Attribute openings, thinnings and granulometries', *Comput. Vis. Image Underst.*, 1996, **64**, pp. 377–389
11 Acton, S.T., and Mukherjee, D.P.: 'Scale-space classification using area morphology', *IEEE Trans. Image Process.*, 2000, **9**, (4), pp. 623–635
12 Young, N., and Evans, A.N.: 'Psycho-visually tuned area-morphology tools for improved image compression'. Proc. Int. Symp. on Mathematical morphology VI, Sydney, Australia, April 2002, pp. 185–195
13 Barni, M.: 'A fast algorithm for 1-norm vector median filtering', *IEEE Trans. Image Process.*, 1997, **6**, (10), pp. 1452–1455
14 Barni, M., and Cappellini, V.: 'On the computational complexity of multivariate median filters', *Signal Process.*, 1998, **71**, (1), pp. 45–54
15 Barni, M., Buti, F., Bartolini, F., and Cappellini, V.: 'A quasi-Euclidean norm to speed up vector median filtering', *IEEE Trans. Image Process.*, 2000, **9**, (10), pp. 1704–1709
16 Alparone, L., Barni, M., Bartolini, F., and Santurri, L.: 'An improved H.263 video coder relying on weighted median filtering of motion vectors', *IEEE Trans. Circuits Syst. Video Technol.*, 2001, **10**, (2), pp. 235–240
17 Meijster, A., and Wilkinson, M.H.F.: 'A comparison of algorithms for connected set openings and closings', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2002, **24**, (4), pp. 484–494
18 Florack, L.M.J.: 'Scale-space theories for scalar and vector images', in 'Scale space and morphology in computer vision' LNCS 2106 (Springer-Verlag, 2001), pp. 193–204