

A Multiscale Technique for Automatically Detecting and Tracking CMEs in Coronagraph Data

J. P. Byrne¹ · H. Morgan¹ · S. R. Habbal¹

© Springer •••

Abstract Studying coronal mass ejections (CMEs) in coronagraph data can be challenging due to their diffuse structure and transient nature, and user-specific biases may be introduced through visual inspection of the images. The large amounts of data available from the SOHO, STEREO, and future Solar Orbiter missions, also makes manual cataloguing of CMEs tedious, and so a robust method of detection and analysis is required. This has led to the development of automated CME detection and cataloguing packages such as CACTus, SEEDS and ARTEMIS. However, the main drawbacks of these catalogues are: the CACTus method of detection fails to resolve CME acceleration profiles; the CACTus and SEEDS running-difference images suffer from spatiotemporal crosstalk; and the SEEDS and ARTEMIS detections are limited to only the LASCO/C2 field-of-view. Recently, the benefits of multiscale filtering of coronagraph data have been demonstrated in an effort to overcome current cataloguing issues. A multiscale decomposition can be applied to individual images in order to enhance the structure of CMEs whilst removing noise and small-scale features like stars. Here we present the development of a new, automated, multiscale, CME detection & tracking technique. It works by first separating the dynamic CME signal from the background corona and then characterising CME structure via a multiscale edge-detection algorithm. The detections are then chained through time to determine the CME kinematics and morphological changes as it propagates across the plane-of-sky. We demonstrate its application to a sample of LASCO data and prove its efficacy in detecting and tracking CMEs. This technique is being applied to the complete LASCO dataset, and it is planned to further develop it for implementation on the SECCHI/COR dataset in the near future.

1. Detecting and Tracking CMEs

In coronagraph images CMEs are observed as outwardly moving regions of stronger brightness intensities relative to the background corona. Different approaches to thresholding the intensity of CMEs in these images have been

¹ Institute for Astronomy, University of Hawaii, 2680
Woodlawn Drive, Honolulu, HI 96822, USA email:
jbyrne@ifa.hawaii.edu email: hmorgan@ifa.hawaii.edu email:
shadia@ifa.hawaii.edu

employed in order to detect their appearance and track their motion through the field-of-view, leading to a cataloguing of their kinematics and morphology. However, these techniques suffer several drawbacks and, as such, different catalogues can vary significantly in their description of events. We introduce a method of multiscale analysis to overcome certain drawbacks of previous detection and tracking methods. In multiscale decompositions of images, noise and small-scale features are removed to leave only larger-scale features of interest such as CMEs. This allows them to be tracked through the image sequences in order to determine their changing kinematics and morphology. Unfortunately coronal streamers tend to appear on similar size scales to CMEs, making their automatic detection difficult. Streamers do, however, tend to remain static on timescales comparable to CME propagation through the field-of-view, and contain much less angular information than the typically curved structure of CMEs, so they may be removed through spatio-temporal filtering of multiscale CME images. This chapter discusses the previous CME detection catalogues, and outlines our use of new methods of multiscale filtering to detect the CME edges in single images. We discuss our efforts to extend this to an automated CME detection algorithm. We also outline an ellipse characterisation of the CME front for study.

1.1. CME Detection Catalogues

Current methods of CME detection have their limitations, mostly since these diffuse objects have been difficult to identify using traditional image processing techniques. These difficulties arise from the varying nature of the CME morphology, the scattering effects and non-linear intensity profile of the surrounding corona, the presence of coronal streamers, and the addition of noise due to cosmic rays and solar energetic particles (SEPs) that impact the coronagraph detectors. The images are also prone to numerous instrumental effects and possible data dropouts. The following standard preprocessing methods are usually applied to optimise the images for CME studies. The coronagraph images are normalised with regard to exposure time in order to correct for temporal variations in the image statistics. A filter may be applied to remove pixel noise, for example to replace hot pixels with a median value of the surrounding pixel intensities, or to reduce the effects of background stars in the image. A correction for vignetting effects and/or lens distortion may be applied to the images. A background subtraction may also be applied, obtained from the minimum of the daily median pixels across a time span of a month. The occulting disc is normally masked, along with any data drop-outs in the images. These steps lead to a clear improvement in the image quality for CME study (Figure 1).

1.1.1. CDAW

The CME catalogue hosted at the Coordinated Data Analysis Workshop (CDAW¹) Data Center grew out of a necessity to record a simple but effective description and analysis of each event observed by SOHO/LASCO (Gopalswamy *et al.*,

¹http://cdaw.gsfc.nasa.gov/CME_list

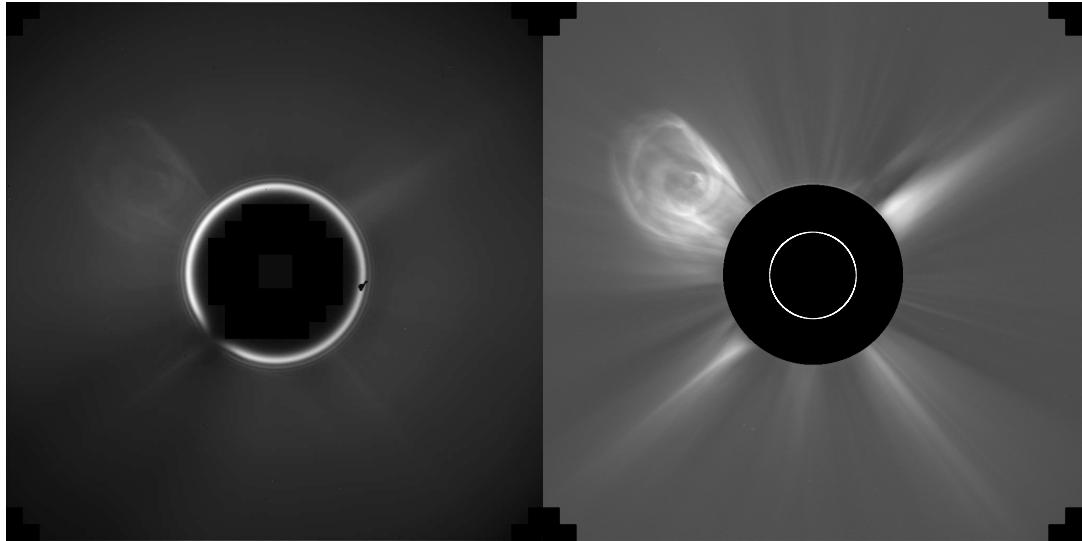


Figure 1. Raw (left) and pre-processed image (right) of a CME observed by LASCO/C2 on 1 April 2004. The pre-processing includes normalising the image statistics, subtracting the background, and masking the occulter disk. The white circle (right) indicates the relative size and position of the Sun behind the occulter.

2009). The catalogue is wholly manual in its operation, with a user tracking the CME through C2 and C3 running-difference images and producing a height-time plot of each event. A linear fit to the height-time profiles provides a 1st-order estimate for the plane-of-sky velocity, and a quadratic fit then provides a 2nd-order velocity fit and an acceleration for the event. The central position angle and angular width of the CME are also deduced from the images, and the event flagged as a halo if it spans 360° , partial halo if it spans $\geq 120^\circ$, and wide if it spans $\geq 60^\circ$. The catalogue itself lists each CME's first appearance in C2, central position angle, angular width, linear speed, 2nd-order speed at final height, 2nd-order speed at $20 R_\odot$, acceleration, mass, kinetic energy, and measurement position angle (the angle along which the heights of the CME are determined). While the human eye is supremely effective at distinguishing CMEs in coronagraph images, errors may be introduced to the manual cataloguing procedure through the biases of different operators; for example, in deciding how the images are scaled, where along the CME the heights are measured, or whether a CME is even worth including in, or discarding from, the catalogue. In an effort to overcome such biases, different automated catalogues have been developed to perform robust CME detections over large data-sets. This is also of great benefit for future missions where the data rate is expected to be too high for manual cataloguing to remain feasible.

1.1.2. CACTus

The Computer Aided CME Tracking catalogue (CACTus²; Robbrecht and Berghmans, 2004) was the first automated CME detection algorithm, in operation since 2004. It is based upon the detection of CMEs as bright ridges in time-height slices (t, r) at each angle θ around a coronagraph image. The images are preprocessed as standard, then a running-difference technique is applied and each image transformed into Sun-centred polar coordinates (r, θ), rebinned, and the C2 and C3 fields-of-view combined. These are then stacked in time, and for each angle the corresponding (t, r) slice undergoes a modified Hough transform for detecting intensity ridges across it. This works by parameterising the (t, r) slice by the variables t_0 and Δt , corresponding to the coordinate intersection point with the time axis, and the distance along the time axis respectively (together called the accumulator space; see Figure ??). So the equation of a line corresponding to an intensity ridge in the slice is given by:

$$r = \frac{r_{max} - r_{min}}{\Delta t}(t - t_0) + r_{min} \quad (1)$$

Thresholding the most significant ridges in the resultant accumulator space filters out the progression of CMEs, with the variables for each ridge characterised by onset time t_R , the velocity v_R ($\sim \frac{1}{\Delta t}$), for angle θ_R , to give a characteristic variable $I_R = (v_R, \theta_R, t_R)$. A 3D scatter plot (v, θ, t) of all detected ridges I_R is then integrated along the v -direction to identify clusters in the resulting (θ, t) map which illustrates the angular span and duration time of the detected CMEs in the coronagraph data. A median velocity across the angular span is quoted as the CME speed.

The running-difference cadence, the ridge intensity threshold, and the imposed limit on how many frames a CME may exist (and indeed the definition of a CME) all affect how successful the detection can be. However, Robbrecht and Berghmans (2004) show the algorithm to be robust in reproducing well the detections of a human user by direct comparison with the CDAW catalogue. The main drawback of the CACTus catalogue for studying CMEs is the imposed zero acceleration of the detection algorithm, since the Hough transform thresholds the ridges as straight lines whose slopes provide a constant velocity. The velocity itself may also be underestimated since it is a median across the span of the CME. The angular spans are possibly over-estimated since side outflows in the images are enhanced by the running-difference and may also include streamer deflections. It is also difficult to distinguish when one CME has fully progressed from the field-of-view and another CME has entered it, so in some cases trailing portions of a CME are detected as separate events.

1.1.3. SEEDS

The Solar Eruptive Event Detection System (SEEDS³; Olmedo *et al.*, 2008) is an automated CME detection algorithm for tracking an intensity thresholded CME

²<http://sidc.oma.be/cactus/>

³<http://spaceweather.gmu.edu/seeds/>

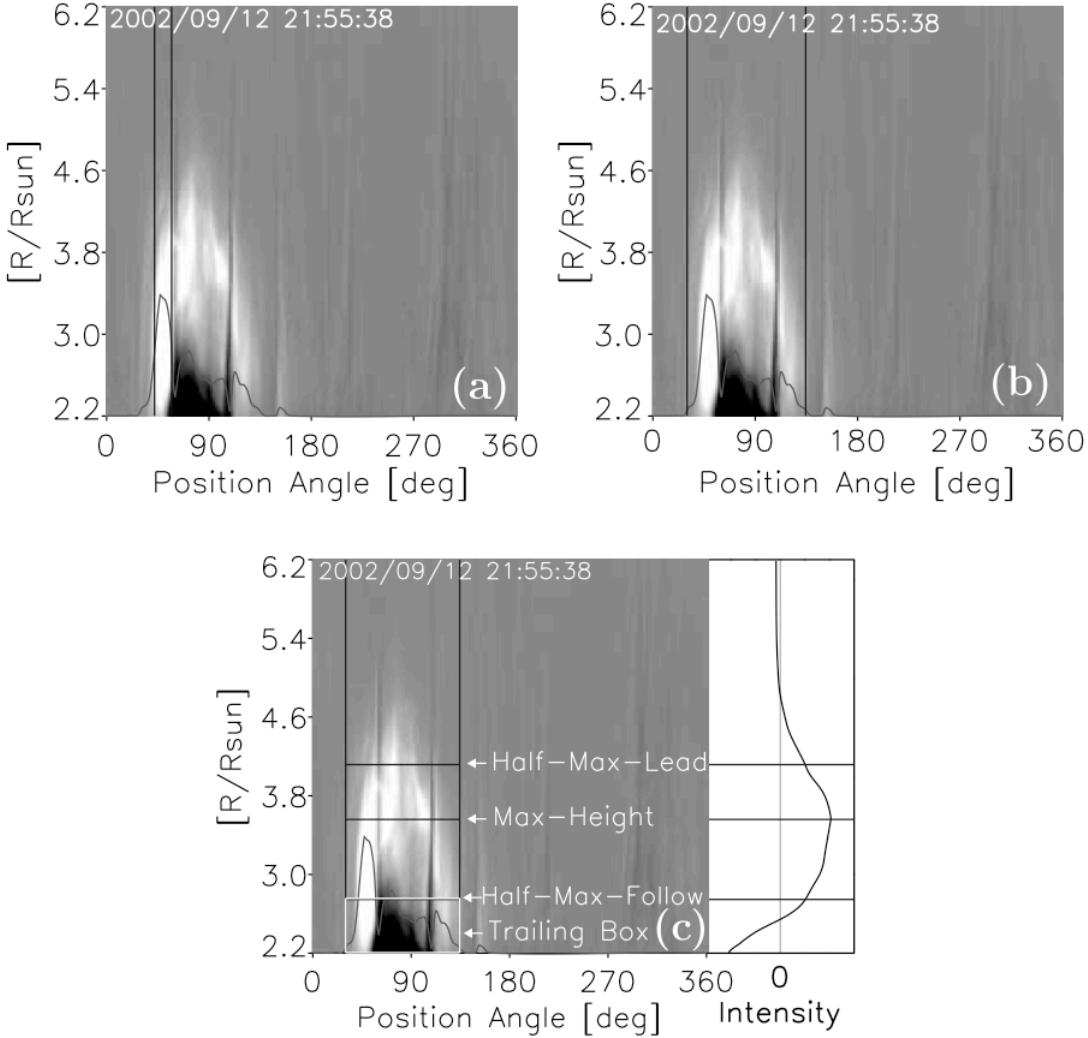


Figure 2. Example of the SEEDS CME detection and height determination, reproduced from Olmedo *et al.* (2008). (a) shows the running-difference image unwrapped into Sun-centre polar coordinates, showing a CME observed by LASCO/C2 on 12 September 2002. The black line distribution across the image represents the positive value intensity count along each angle, and the two vertical black lines mark the angular span at one standard deviation above the mean intensity. (b) shows the new angular span following the region growing technique. (c) shows the intensity within the angular span averaged across heights, and the ‘Half-Max-Lead’ is taken as the CME height in the image.

front in running-difference images from LASCO/C2. The images are preprocessed as standard, unwrapped into Sun-centred polar coordinates (r, θ), and a

normalised running-difference technique is applied using the following equation:

$$u_i = \left[n_i - n_{i-1} \left(\frac{\bar{n}_i}{\bar{n}_{i-1}} \right) \right] \frac{\alpha}{\Delta t} \quad (2)$$

where u_i is the running-difference image, \bar{n} is the mean of the pixels in the entire field-of-view of the image n , Δt is the time difference between images (in minutes), α is a constant set to approximately the smallest time difference (Δt) between any image pair, where i is the current image and $i - 1$ the preceding image. This normalised difference ensures that the mean of the new image (u_i) will effectively be zero.

The pixel intensities (positive values only) are then summed along angles and thresholded at a certain number of standard deviations above the mean intensity: $\mu + N\sigma$ (cf. Equation 12) as in Figure 2a. This determines the ‘core angles’ of the CME, and a region growing technique based on a secondary threshold of intensities in the rest of the image is applied to open the angular span to include the full CME (Figure 2b). Issues arise when streamer deflections occur that will offset the region growing technique and overestimate the CME angular width. An intensity average across the angles within the span of the CME is then determined, and where the forward portion of this intensity profile equals half its maximum value is taken as the CME height (Figure 2c). The velocity and acceleration are determined from the heights through consecutive images and these results are output with the CME position angle and angular width in the SEEDS catalogue.

Along with the issues of streamer deflections and the tracking being limited to C2 images, the choice of the ‘Half-Max-Lead’ as the CME height is dependant on the overall CME brightness, and thus any brightness changes as the CME propagates will affect this measurement. This would add to the error on the height-time profile which, along with the error in time as a result of the running-difference technique, makes it difficult to accurately determine the velocity and acceleration.

1.1.4. ARTEMIS

The Automatic Recognition of Transient Events and Marseille Inventory from Synoptic maps (ARTEMIS⁴; Boursier *et al.*, 2009) is an automated CME detection algorithm that works by identifying signatures of transients in synoptic maps. These maps are generated as (t, θ) slices for specific heights r in the corona-graph images of LASCO/C2. The images are prepared through the standard preprocessing steps. Then at a specific height (e.g., $r = 3 R_\odot$) the intensity is plotted across all angles θ for each image through time t with transient events appearing as vertical streaks through the more persistent streamer intensities (Figure 3). A method of image filtering and intensity thresholding is applied to distinguish the streaks in the synoptic map, and image segmentation then discards small features and closes off regions-of-interest (ROIs) to produce a

⁴<http://www.oamp.fr/lasco/>

binary map of the streaks. Specific parameters of these streaks are also computed, such as their total radiances, areas and centres of gravity. Merging with high-level knowledge helps to associate ROIs of the same CME if three criteria are met: the difference between the x -coordinates of the centres of two ROIs differs by less than two pixels; the difference between the y -coordinates of their centres differs by less than 60 pixels (corresponding to a 60° angular span); and the ratio of their radiances calculated at their centres (on the original synoptic map) ranges from 0.25 to 4. The result is a binary CME detection map in (t, θ) space for different heights in the corona: 3, 3.5, 4, 4.5, 5 and $5.5 R_\odot$.

With the CME detections in place, estimates of the velocity may be made. A first estimate is taken by testing a range of constant velocities $50 - 2,000 \text{ km s}^{-1}$ to determine which best matches the shifting of the CME detection in synoptic maps at subsequent heights through the corona. The binary maps are shifted by an amount corresponding to velocity steps of 10 km s^{-1} , such that the one which provides the maximum pixel value (with a minimum limit of 3) indicates the best velocity estimate of the event. A second estimate is taken by cross-correlating the detected CME ROIs on the original synoptic maps at $3 R_\odot$ and $5.5 R_\odot$ and inspecting the intensity shift in time (pixel shift in x -direction) to obtain the velocity estimate. A third estimate is taken by similar cross-correlation but specifically on each individual line of the ROIs to obtain a distribution of velocities across the angular span of the CME, the median of which is taken as the actual velocity. Boursier *et al.* (2009) compare histograms of the three different velocity estimates for the ARTEMIS CME detections over a twelve year interval and find that, globally, the three estimates are highly consistent with each other.

ARTEMIS is limited to the C2 field-of-view and it provides kinematics only in the $3 - 5.5 R_\odot$ range. The velocity determinations themselves are not specific to either the CME front nor any other identifiable feature, and carry all the inaccuracies resulting from the image rebinning, intensity averaging, filtering and segmentation techniques in generating the final detection masks.

Due to the drawbacks of each of the catalogues above, the motivation exists to study the kinematics and morphology of CMEs with as great an accuracy as possible in order to better compare with theory. To this end we outline below our application of multiscale analysis to remove small scale noise/features and enhance the larger scale CME in single coronagraph frames, allowing the CME front edges to be detected and a geometrical characterisation applied to study its propagation with increased accuracy for deriving the kinematics and morphology.

1.2. Multiscale Filtering

In this section a new multiscale method of analysing CMEs is described. The use of multiscale methods in astrophysics have proven effective at denoising spectra and images (Murtagh, Starck, and Bijaoui, 1995; Fligge and Solanki, 1997), analysing solar active region evolution (Hewett *et al.*, 2008), and enhancing solar coronal images (Stenborg and Cobelli, 2003; Stenborg, Vourlidas, and Howard, 2008). A particular application of multiscale decompositions uses high and low pass filters convolved with the image data to exploit the multiscale nature of the CME (Young and Gallagher, 2008). This highlights its intensity

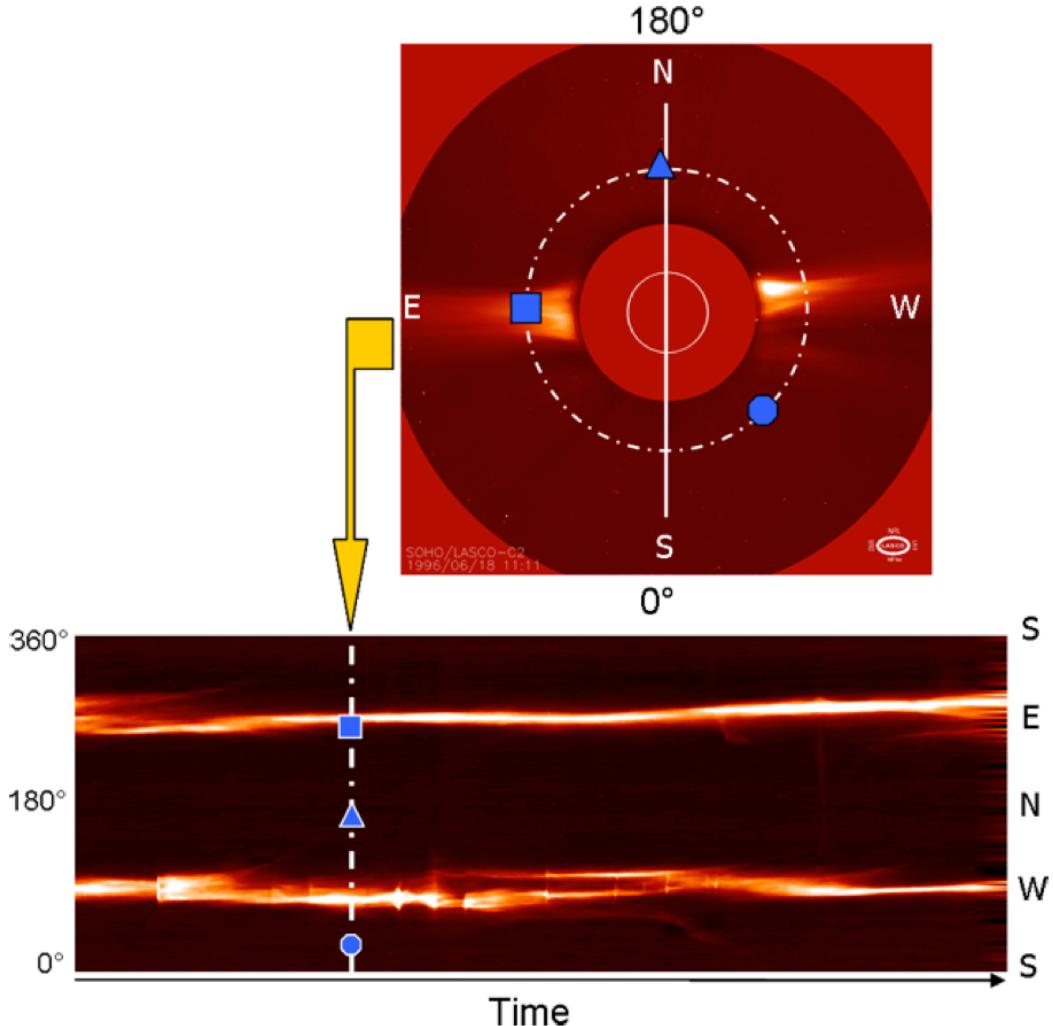


Figure 3. An example of how the synoptic maps are generated for the ARTEMIS catalogue, reproduced from Boursier *et al.* (2009). At a chosen height in the coronagraph image an annulus is unwrapped (indicated with the dashed line and blue square, circle and triangle) and these are then stacked together to illustrate how the intensity at that height changes through time. Vertical streaks represent transient events occurring on smaller time-scales than the more persistent streamers in the images.

against the background corona as it propagates through the field-of-view, while neglecting small scale features (essentially denoising the data). It also leads to the use of non-maxima suppression to trace the edges in the CME images, and Young and Gallagher (2008) show the power of multiscale methods over previous edge detectors such as Roberts (Davis, 1975) and Sobel (Duda and Hart, 1973). With these methods for defining the front of the CME we can characterise its kinematics (position, velocity, acceleration) and morphology (width, orientation)

in coronagraph images. Multiscale analysis also has the benefit of working on independent images without any need for differencing, so the temporal errors involved are on the order of the exposure time of the instrument (\sim a few seconds).

The fundamental idea behind wavelet analysis is to highlight details apparent on different scales within the data. An example of this is the suppression of noise in images, which tends to occur only on the smallest scales. Wavelets have benefits over previous methods (e.g., Fourier transforms) because they are localised in space and are easily dilated and translated in order to operate on multiple scales, the basic equation being:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{b}} \psi\left(\frac{t-a}{b}\right) \quad (3)$$

where a and b represent the shifting (translation) and scaling (dilation) of the mother wavelet ψ which can take several forms depending on the required use.

We explore a method of multiscale decomposition in 2D through the use of low and high pass filters; using a discrete approximation of a Gaussian, θ , and its derivative, ψ , respectively (Stenborg and Cobelli, 2003). Since $\theta(x, y)$ is separable, i.e. $\theta(x, y) = \theta(x)\theta(y)$, we can write the wavelets as the first derivative of the smoothing function:

$$\psi_x^s(x, y) = s^{-2} \frac{\partial \theta(s^{-1}x)}{\partial x} \theta(s^{-1}y) \quad (4)$$

$$\psi_y^s(x, y) = s^{-2} \theta(s^{-1}x) \frac{\partial \theta(s^{-1}y)}{\partial y} \quad (5)$$

where s is the dyadic scale factor such that $s = 2^j$ where $j = 1, 2, 3, \dots, J$. Successive convolutions of an image with the filters produces the scales of decomposition, with the high-pass filtering providing the wavelet transform of image $I(x, y)$ in each direction:

$$W_x^s I \equiv W_x^s I(x, y) = \psi_x^s(x, y) * I(x, y) \quad (6)$$

$$W_y^s I \equiv W_y^s I(x, y) = \psi_y^s(x, y) * I(x, y) \quad (7)$$

Akin to a Canny edge detector (Young and Gallagher, 2008), these horizontal and vertical wavelet coefficients are combined to form the gradient space, $\Gamma^s(x, y)$, for each scale:

$$\Gamma^s(x, y) = [W_x^s I, W_y^s I] \quad (8)$$

The gradient information has an angular component α and a magnitude (edge strength) M :

$$\alpha^s(x, y) = \tan^{-1} (W_y^s I / W_x^s I) \quad (9)$$

$$M^s(x, y) = \sqrt{(W_x^s I)^2 + (W_y^s I)^2} \quad (10)$$

The resultant horizontal and vertical detail coefficients, and the magnitude and angular information are illustrated in Figure 4.

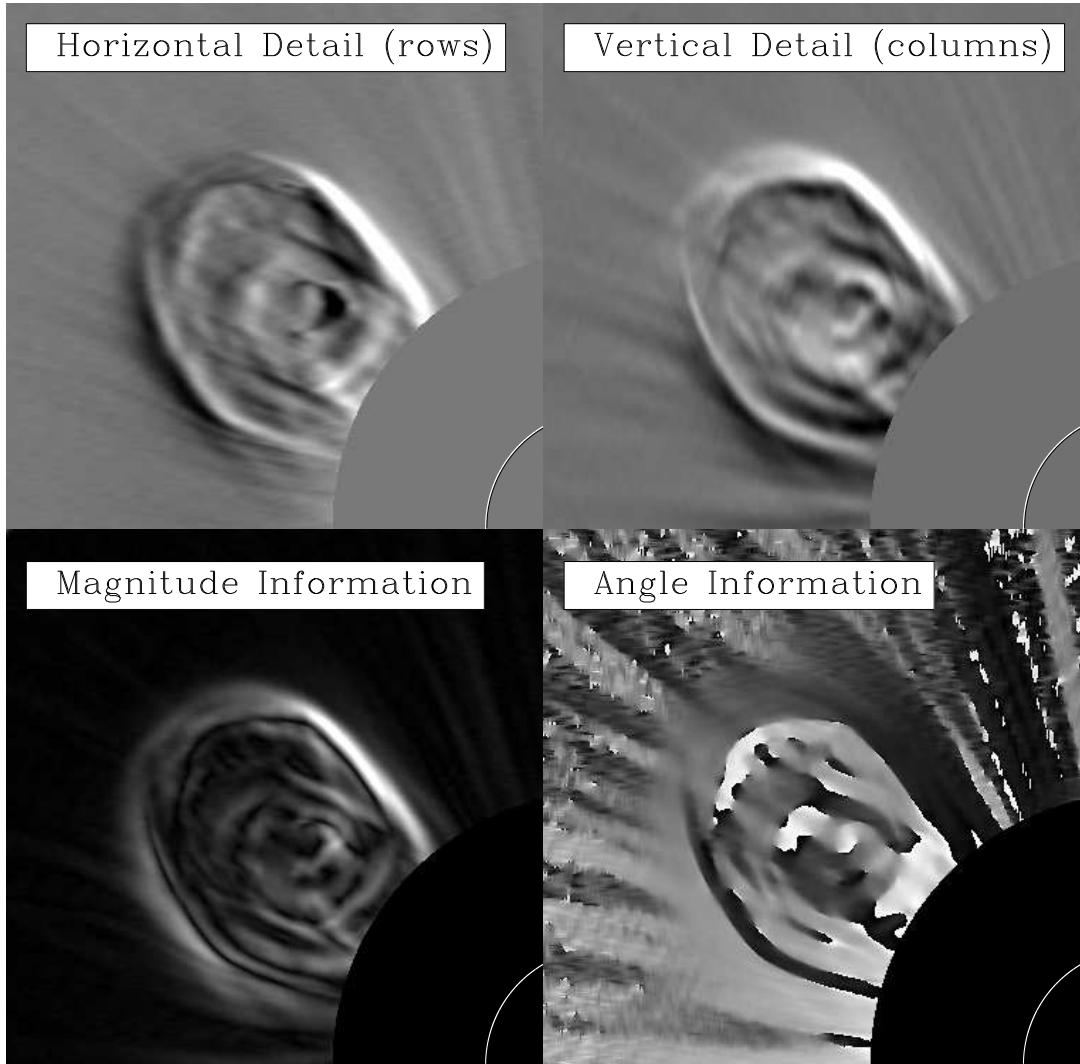


Figure 4. Top left, the horizontal detail, and top right, the vertical detail from the high-pass filtering at one scale of the multiscale decomposition (called the rows and columns respectively). Bottom left, the corresponding magnitude (edge strength) and bottom right, the angle information ($0 - 360^\circ$) taken from the gradient space, for a CME observed in LASCO/C2 on 1 April 2004 (Byrne *et al.*, 2009).

At a particular scale the signal-to-noise ratio of the CME is highest and this is the optimum scale for determining the edges in the image. The angular component α of the gradient specifies a direction which points across the greatest intensity change in the data (an edge). A threshold is specified with regard to this gradient direction in order to chain pixels along maxima to highlight the edges. The changes in magnitude and angular information may then be implemented in

a spatio-temporal filter for distinguishing those edges corresponding to the CME only. Overlaying a mesh of vector arrows on the data shows how the combined magnitude and angular information illustrate the progression of the CME. Each vector is rooted on a pixel in the gradient space, and has a length corresponding to the magnitude M with an angle from the normal α (Figure ??). Using this information, it becomes possible to create a specific detection mask which is used to identify the edges along the CME front to study its propagation (as done for a sample of events in Chapter ??). However, for the cases of faint CMEs or strong streamer deflections, the filter is presently limited by exploiting the information from only one scale and ignoring all other scales, meaning it currently often requires the user to remove/include certain edges that the algorithm has mistakenly retained/discard. Extending the algorithm to work on more than one scale may help alleviate this issue in order to develop a fully automated CME detection and characterisation routine, as outlined in the following section.

1.3. Automated Multiscale CME Detection

For the most part, CMEs exist on size scales larger than noise and any small scale features in coronagraph images, such as stars or planets, that are redundant for studying CME propagation. This fact has led to the development and implementation of multiscale decompositions that highlight the CME in images from SOHO/LASCO and STEREO/SECCHI (Young and Gallagher, 2008; Byrne *et al.*, 2009; Stenborg and Cobelli, 2003). However, coronal streamers (plasma outflows from open magnetic field regions on the Sun) can persist through coronagraph images with significant brightness intensities and tend to appear on similar scales as the CME in multiscale image analysis. If a CME propagates through an image with a strong streamer present, it becomes difficult to distinguish the two features by intensity thresholding alone, and this is one reason why differencing techniques have been widely used in CME analysis. In an effort to move away from differencing and the large errors involved in the subtraction of images from each other, since the goal is to obtain kinematics with the greatest precision, we endeavour to separate CME and streamer features from one another using multiscale methods alone. These efforts involve exploiting the angular distribution that exists across a curved CME front compared to the more linear streamers in single independent images. To do this, the coronagraph images must first be normalised for their radial gradient in intensity, since the drop-off across the field-of-view is too steep to effectively segment a single entire streamer from the inner to outer edge of an image. This can serve to enhance the noise at the edge of the images, but this is again suppressed by the multiscale analysis. Occasions when the CME propagates directly in front of, or behind, a streamer remain problematic, as do strong streamer deflections that can occur when a CME propagates into or expands alongside a streamer.

1.4. Normalising Radial Graded Filter (NRGF)

Since the brightness drop-off of the corona is large, falling from approximately $10^{-6} - 10^{-9}$ B_\odot across heights of $1 - 6 R_\odot$ (Kimura and Mann, 1998), a method

for radially normalising coronagraph images to enhance features across this steep intensity gradient was developed by Morgan, Habbal, and Woo (2006). It works by normalising the intensity in radial coordinates of the image according to the equation

$$I'(r, \phi) = \frac{I(r, \phi) - I(r)_{<\phi>}}{\sigma(r)_{<\phi>}} \quad (11)$$

where $I'(r, \phi)$ is the processed and $I(r, \phi)$ is the original intensity at height r and position angle ϕ , and $I(r)_{<\phi>}$ and $\sigma(r)_{<\phi>}$ are the mean and standard deviation of intensities calculated over all position angles at height r . Figure 5 shows the result when the NRGF is applied to a LASCO/C2 image of the 1 April 2004 CME.

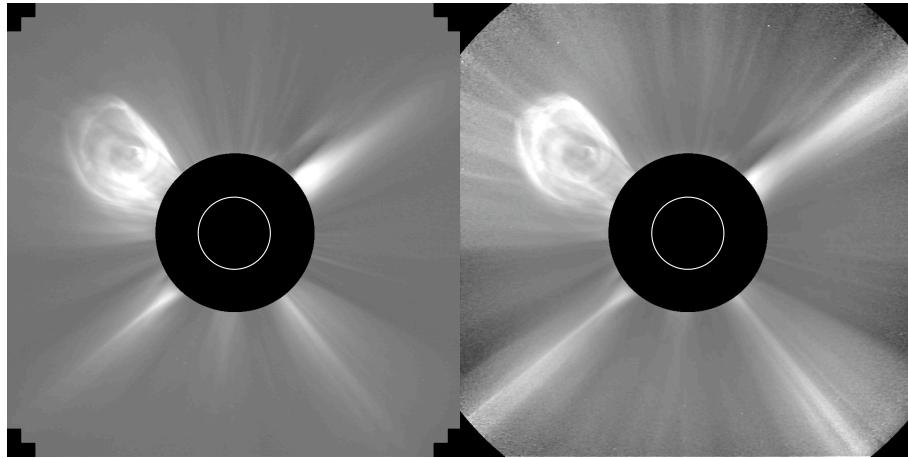


Figure 5. A normalised, background subtracted, LASCO/C2 image (left) of a CME on 1 April 2004, and the resulting NRGF image (right). The image radial intensity is scaled such that structure along streamers and the CME becomes visible across the field-of-view

The multiscale decomposition introduced in Young and Gallagher (2008) provides magnitude and angular information of the edges in the image. This information is combined to chain the strongest edges within the image on a scale that provides an optimal signal-to-noise ratio for studying the CME. Byrne *et al.* (2009) obtain the CME front edges in this manner, which are then used to fit an ellipse to characterise the CME propagation in the image sequence. In order to automate the algorithm, thresholds on the magnitude information (e.g., CME edges appear on larger scales than noisy features) and angular information (e.g., CME edges appear more curved than streamer edges) were investigated. The thresholding is strengthened by the inclusion of more than one scale in localising the CME and distinguishing it from the streamers, detailed below.

1.4.1. Thresholding

The magnitude information corresponds to the strength of the edges in the image, and so can be thresholded to discard the small scale noise. For the NRGF

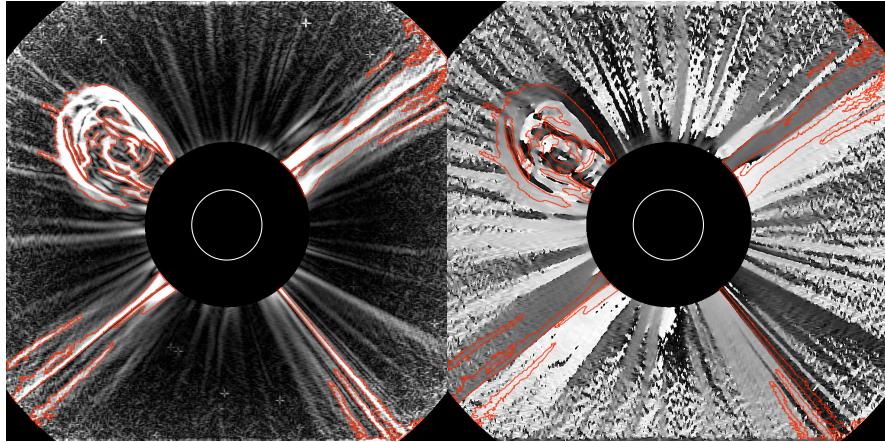


Figure 6. A chosen scale of the decomposed NRGF image provides a magnitude image of the edge strengths displayed on the left, which is thresholded at one standard deviation from the mean intensity to obtain contoured regions of interest that could contain a CME (sample contours indicated in red). As shown, the streamers have edges which appear on the same scale as the CME edges in this image. The angular information from the decomposition is displayed on the right, and the contoured regions of interest overlaid for comparison. The grey scale indicates angles from $0\text{--}360^\circ$ and it is clear that streamers tend to have a linear grey scale while the CME has a gradient of greys across the scale.

image (Figure 5), a hard threshold T is set at one standard deviation σ of the mean μ of the image intensity to contour regions of interest that may be a CME according to the equations:

$$T = \mu + N\sigma = \underbrace{\frac{1}{n} \sum_{i=1}^n x_i}_{\text{mean}} + N \underbrace{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}}_{\text{standard deviation}} \quad (12)$$

where x_i are the pixel intensity values of the image, and we choose the number of standard deviations $N = 1$. The left image of Figure 6 illustrates this thresholding with a sample of contoured regions (outlined in red) on the multiscale decomposition of a CME observed in LASCO/C2 on 1 April 2004.

It is apparent from the right of Figure 6 that the CME will contain edges whose normals are widely distributed across $0\text{--}360^\circ$ compared to the more linear, radially directed, streamer edges. The angular distribution of each region is determined and then normalised and folded into $0\text{--}180^\circ$ range, centred on 90° , due to the symmetry of the edge normals. This is illustrated for four selected contour regions in Figure 7. The resulting angular distributions are then thresholded with regard to their median value, since the distribution of angles across the CME will be wider and have a higher median value than for a distribution of angles along a streamer.

This thresholding is repeated across four scales of the multiscale decomposition, neglecting smaller scales dominated by noise and larger scales that are

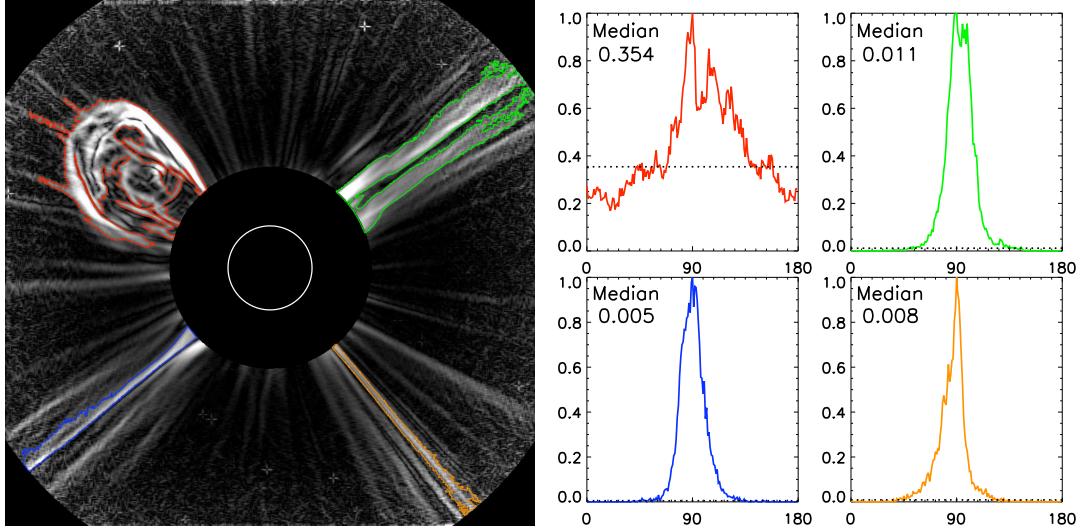


Figure 7. Left: four contoured regions (at one standard deviation of the mean image intensity) highlighted on the magnitude information from the multiscale decomposition of the 1 April 2004 CME. Right: the corresponding angular distribution of each region, normalised and folded into the $0 - 180^\circ$ range (centred on 90°). The angular distribution may be thresholded with respect to its median value to distinguish regions corresponding to CMEs from those along streamers.

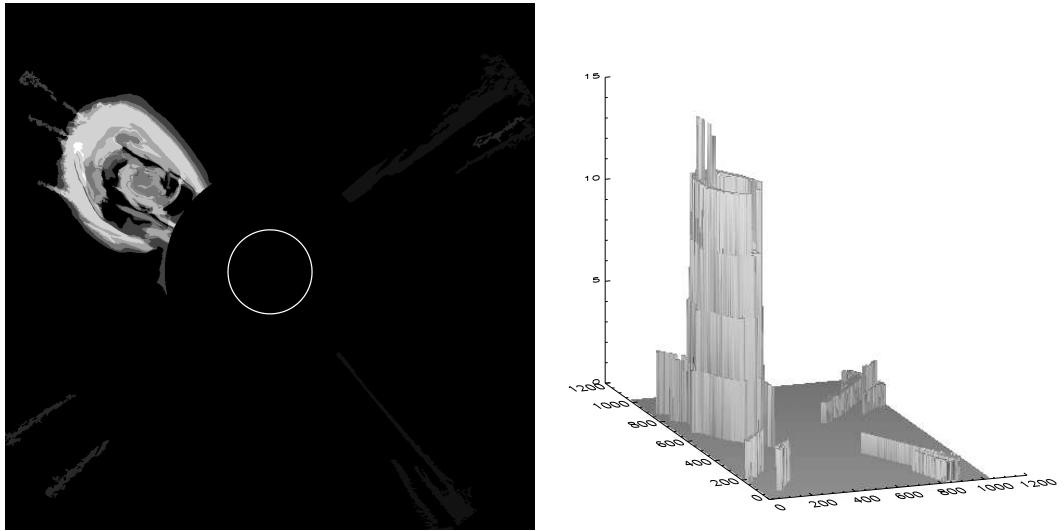


Figure 8. The resulting CME detection mask from combining the thresholded regions of strongest magnitude and angular distribution at four scales of the decomposition. The 3D representation on the right illustrates the pixel values of the mask.

overly smoothed. Assigning a score to the regions that may contain a CME at each scale, it becomes possible to build a detection mask as in Figure 8.

The scoring system is chosen arbitrarily to work best with the chosen thresholds, and these may be changed and refined as an analysis of more CMEs is done. For example, the current thresholds from working on a sample of ~ 10 CMEs are as follows:

1. The magnitude information is thresholded at one standard deviation (1σ) of the mean intensity across the image.
2. The 15 largest contoured regions across the image are investigated (there are rarely more than ~ 5 streamers of similar intensity to a CME, and we allow for disjointed contours along structures).
3. If the median angular value is $> 20\%$ of the angular distribution peak then the region is deemed a CME and assigned a score of 3 (the pixels in that region of the mask are given the value 3). If it is $> 10\%$ the score is 2 (potential CME structure), or $> 5\%$ the score is 1 (weak CME structure or portion thereof).
4. The final CME mask through the combination of scores at each scale results in a dominant region that localises the CME front in the image and can be used to characterise the front, or input into a spatio-temporal filter if subsequent CME images are available in order to refine the masked region whenever streamers are still present.

The resultant set of detection masks for three of the frames of the CME observed on 1 April 2004 are shown in Figure 9.

1.4.2. Faint CMEs and Streamer Interactions/Deflections

Due to the nature of the hard thresholds in place on the magnitude and angular information, there are problems which arise when the algorithm mistakenly disregards a CME or includes a streamer, or portions thereof. Firstly, if a CME is faint enough that the intensity falls below the 1σ magnitude threshold, it will not be detected as a region of interest in the image. Secondly, if the CME interacts with a streamer, the two features may be contoured together and this will skew the angular distribution and affect the detection mask. And thirdly, if the CME causes a significant streamer deflection, it will lead to a wider distribution of angles along the streamer and the algorithm may thus detect it as part of the CME. This is why the above scoring system was introduced in an effort to minimise these effects, which are highlighted in Figure 10 for a CME observed on 23 April 2001. The event is too faint compared with the streamers across it for it to be easily distinguished in the image, and parts of the streamers are then mistakenly included in the final detection mask. This is where the current algorithm requires a user to specify which parts of the edges correspond to the CME for characterisation. Such limitations in current wavelet analysis of CMEs may be overcome by extending these algorithms to work with ridgelets or curvelets that better suit the curved form of a typical CME front as discussed in ?. Furthermore, it should be noted that a relatively small number of CMEs exhibit a narrow, spiky or strongly kinked structure which may not be readily distinguished from streamers in an automated fashion, nor have an adequately high angular distribution for automatic detection in the images. These CMEs may also not be satisfactorily characterised with an ellipse, discussed in the

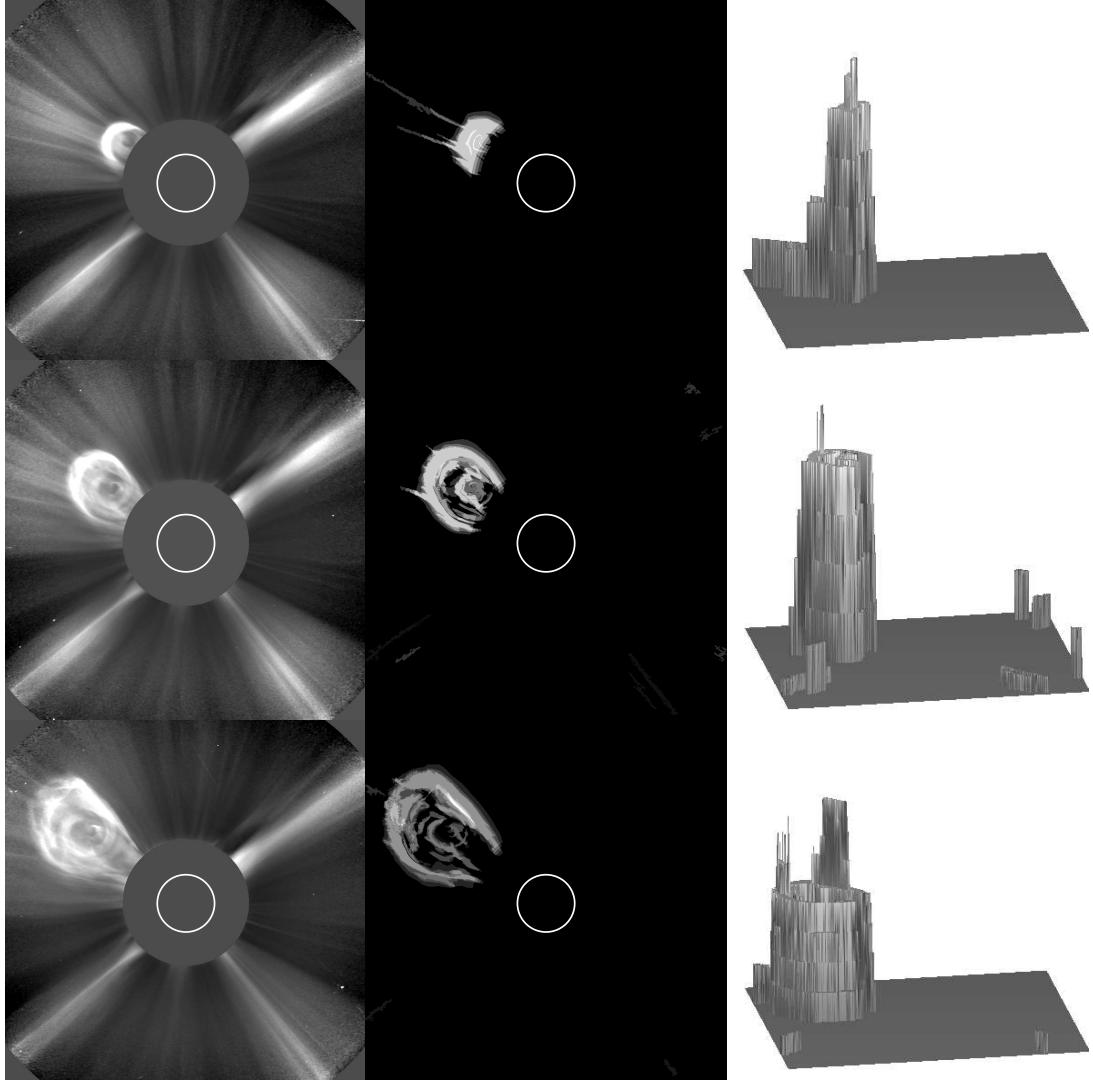


Figure 9. The NRGF (left) and resulting detection masks (middle and right) for different frames of a CME observed by LASCO/C2 at times of 00:00 UT (top), 00:40 UT (middle) and 01:20 UT (bottom) on 2 April 2004. The location of the CME front is highlighted very efficiently by this method, although the detection masks may contain artefacts of the chosen thresholds which must be discarded when characterising the CME front.

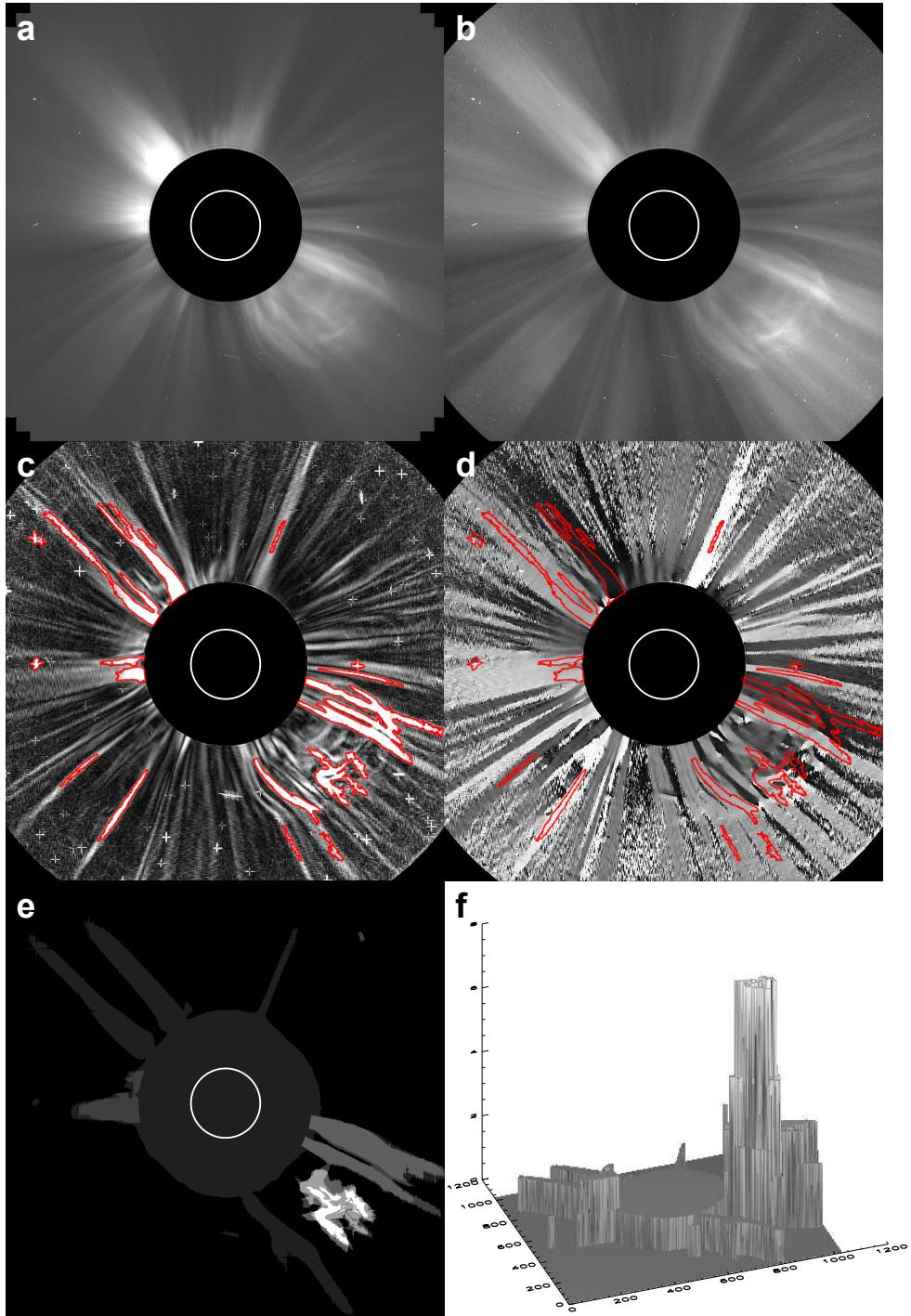


Figure 10. Example of the difficulty in detecting the faint CME observed by LASCO on 23 April 2001. If the CME is far-sided and/or of low intensity it becomes difficult to threshold its edges in the image compared to the coronal streamers. (a) is the pre-processed CME image. (b) is the NRGF image. (c) and (d) show the magnitude and angular information from the multiscale decomposition with intensity contours overlaid in red. (e) and (f) show the resulting CME detection mask in 2D and 3D respectively.

following section, however they are not typical of the commonly observed curved nature of CMEs and thus represent a small class of events outside the limits of these methods as they currently stand.

References

- Boursier, Y., Lamy, P., Llebaria, A., Goudail, F., Robelus, S.: 2009, The ARTEMIS Catalog of LASCO Coronal Mass Ejections. Automatic Recognition of Transient Events and Marseille Inventory from Synoptic maps. *Solar Physics* **257**, 125–147. doi:10.1007/s11207-009-9370-5.
- Byrne, J.P., Gallagher, P.T., McAteer, R.T.J., Young, C.A.: 2009, The kinematics of coronal mass ejections using multiscale methods. *Astronomy & Astrophysics* **495**, 325–334. doi:10.1051/0004-6361:200809811.
- Davis, L.S.: 1975, A Survey of Edge Detection Techniques. *Computer Graphics Image Processing* **4**, 248–270.
- Duda, R.O., Hart, P.E.: 1973, *Pattern classification and scene analysis*. New York, NY: Wiley.
- Fligge, M., Solanki, S.K.: 1997, Noise reduction in astronomical spectra using wavelet packets. *Astronomy & Astrophysics Supplemental* **124**, 579–587. doi:10.1051/aas:1997208.
- Gopalswamy, N., Yashiro, S., Michalek, G., Stenborg, G., Vourlidas, A., Freeland, S., Howard, R.: 2009, The SOHO/LASCO CME Catalog. *Earth Moon and Planets* **104**, 295–313. doi:10.1007/s11038-008-9282-7.
- Hewett, R.J., Gallagher, P.T., McAteer, R.T.J., Young, C.A., Ireland, J., Conlon, P.A., Maguire, K.: 2008, Multiscale Analysis of Active Region Evolution. *Solar Physics* **248**, 311–322. doi:10.1007/s11207-007-9028-0.
- Kimura, H., Mann, I.: 1998, Brightness of the solar F-corona. *Earth, Planets, and Space* **50**, 493–499.
- Morgan, H., Habbal, S.R., Woo, R.: 2006, The Depiction of Coronal Structure in White-Light Images. *Solar Physics* **236**, 263–272. doi:10.1007/s11207-006-0113-6.
- Murtagh, F., Starck, J., Bijaoui, A.: 1995, Image restoration with noise suppression using a multiresolution support. *Astronomy & Astrophysics Supplemental* **112**, 179.
- Olmedo, O., Zhang, J., Wechsler, H., Poland, A., Borne, K.: 2008, Automatic Detection and Tracking of Coronal Mass Ejections in Coronagraph Time Series. *Solar Physics* **248**, 485–499. doi:10.1007/s11207-007-9104-5.

- Robbrecht, E., Berghmans, D.: 2004, Automated recognition of coronal mass ejections (CMEs) in near-real-time data. *Astronomy & Astrophysics* **425**, 1097–1106. doi:10.1051/0004-6361:20041302.
- Stenborg, G., Cobelli, P.J.: 2003, A wavelet packets equalization technique to reveal the multiple spatial-scale nature of coronal structures. *Astronomy & Astrophysics* **398**, 1185–1193. doi:10.1051/0004-6361:20021687.
- Stenborg, G., Vourlidas, A., Howard, R.A.: 2008, A Fresh View of the Extreme-Ultraviolet Corona from the Application of a New Image-Processing Technique. *Astrophysical Journal* **674**, 1201–1206. doi:10.1086/525556.
- Young, C.A., Gallagher, P.T.: 2008, Multiscale Edge Detection in the Corona. *Solar Physics* **248**, 457–469. doi:10.1007/s11207-008-9177-9.

