Solar Phys (2010) 262: 481–494
DOI 10.1007/s11207-009-9495-6

Tracking of Coronal White-Light Events by Texture

N. Goussies · G. Stenborg · A. Vourlidas · R. Howard

Received: 14 May 2009 / Accepted: 16 December 2009 / Published online: 20 January 2010

Abstract The extraction of the kinematic properties of coronal mass ejections (CMEs) from white-light coronagraph images involves a significant degree of user interaction: defining the edge of the event, separating the core from the front or from nearby unrelated structures, *etc*. To contribute towards a less subjective and more quantitative definition, and therefore better kinematic characterization of such events, we have developed a novel image-processing technique based on the concept of "texture of the event". The texture is defined by the so-called gray-level co-occurrence matrix, and the technique consists of a supervised segmentation algorithm to isolate a particular region of interest based upon its similarity with a pre-specified model. Once the event is visually defined early in its evolution, it is possible to automatically track the event by applying the segmentation algorithm to the corresponding time series of coronagraph images. In this paper we describe the technique, present some examples, and show how the coronal background, the core of the event, and even the as-

Solar Image Processing and Analysis
Guest Editors: J. Ireland and C.A. Young

N. Goussies
Facultad de Cs. Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina
e-mail: ngoussie@dc.uba.ar

N. Goussies
George Mason University, Fairfax, VA 22030, USA

G. Stenborg (✉)
Interferometrics, Inc., Herndon, VA 20171, USA
e-mail: guillermo.stenborg.ctr.ar@nrl.navy.mil

A. Vourlidas · R. Howard
U.S. Naval Research Laboratory, Washington, DC 20375, USA

A. Vourlidas
e-mail: angelos.vourlidas@nrl.navy.mil

R. Howard
e-mail: russ.howard@nrl.navy.mil

sociated shock (if one exists) can be identified for different kind of CMEs detected by the LASCO and SECCHI coronagraphs.

Keywords CME · Coronagraph · Automatic tracking

1. Introduction

The white-light coronagraphs onboard SOHO (LASCO: Brueckner *et al.*, 1995) and most recently STEREO (SECCHI: Howard *et al.*, 2008) have detected thousands of dynamic events, in particular Coronal Mass Ejections (CME). The CME morphology can vary depending on projection effects, kinematic properties, and source region topology. The largest of these CMEs are a key contributor of space weather as they can become geoeffective under certain conditions. Hence it is important to reliably characterize their morphological and kinematic properties as a first step to predict their effects. However, there are no universally accepted rules on how to define the event boundaries, which makes it difficult to agree on the kinematics of a given event (*e.g.*, Byrne *et al.*, 2009). The lack of agreement is reflected in both manual (*e.g.*, CDAW CME Catalog: Yashiro *et al.*, 2004; http://cdaw.gsfc.nasa.gov/CME_list/) and automated CME catalogs (*e.g.*, CACTUS: Robbrecht, Berghmans, and Van der Linden, 2009; SEEDS: Olmedo *et al.*, 2008).

One of the fundamental methods used by human vision to discriminate between surfaces (background) and objects (foreground) is that of texture analysis. This approach, alone or combined with other features (*e.g.*, shape, brightness, motion, *etc.*), is commonly applied in computer vision to distinguish, characterize, and eventually extract objects in digital images. Although several techniques to detect and track CMEs have been exercised on white-light coronagraph images (Robbrecht and Berghmans, 2004; Boursier *et al.*, 2005; Liewer *et al.*, 2005; Qu *et al.*, 2006; Olmedo *et al.*, 2008) none of them is based on texture analysis.

Briefly, image texture is defined as a function of the spatial variation of the pixel intensities (gray values). The problem that texture analysis research attempts to solve is that of texture segmentation, where the goal is to obtain a boundary map separating the differently textured regions in an image.

Goussies *et al.* (2009) first attempted to detect and track CMEs using texture segmentation on LASCO-C2 running-difference images. Their work showed that texture analysis can provide valuable insights into the characterization of coronal events in the C2 field of view and thus help discriminate them, on a quantitative basis, from the background. In this work, we extend the algorithm to work with any kind of coronagraph image, and improve it to discriminate between concurrent events, as well as between the bulk of the CME events and the associated shock, if any. We test this new CORonal SEgmentation Technique (hereafter CORSET) based on texture on LASCO-C2, C3, and SECCHI-COR2 coronagraph observations.

The paper is organized as follows: In Section 2 we briefly address the problem of CME tracking from the computer-vision point of view as well as the solution proposed. In Section 3 we show the modifications that we introduced to the procedure with respect to the original work to improve its performance. Section 4 shows the data-preprocessing scheme devised to reveal the texture of the event while minimizing side-effects due to conventional base – difference techniques. Section 5 describes the sample of the events analyzed and the results obtained. The discussion of the results is conveyed in Section 6. Finally, we conclude in Section 7.

2. Segmentation

Our objective is to isolate a certain coronal feature from the background in a given coronagraph image. In computer-vision terminology, this is called a bipartitioning segmentation problem. To address this problem we developed a supervised segmentation algorithm inspired by the region competition model using a level-sets technique for the CME front evolution. In more common terms, we developed a quasi-automated procedure to separate the CME from the background. The method uses an initial user input to delineate (*i.e.*, to give an initial estimate of) the CME boundary in the first image in a series. The algorithm then traces this boundary as it evolves in a sequence of images, by assigning pixels with common texture characteristics to the CME and the rest to the background. A detailed description of the technique as originally developed can be found in Goussies *et al.* (2009). Here, we briefly outline the algorithm, terminology, and background concepts.

2.1. The Segmentation Algorithm

Let $I : \Omega \rightarrow \Re$ be an image defined over $\Omega \subseteq \Re^2$. The goal of the bipartitioning segmentation process is to find a set of two regions $\mathcal{R} = \{\mathbf{R}_1, \mathbf{R}_2\}$ such that:

- Each region is a subset of the image domain $\mathbf{R}_i \subseteq \Omega$.
- The regions are pairwise disjoint $\mathbf{R}_i \cap \mathbf{R}_j = \phi \; \forall i \neq j$.
- The union of all the regions covers the image domain $\bigcup_{i=1}^{N} \mathbf{R}_i \subseteq \Omega$.
- The points in each region share common image characteristics.

The set of regions found by the segmentation process should be a representation of a meaningful feature, in our case, the white-light CME we are interested in. Since the meaningfulness of the regions depends on the problem at hand, many algorithms have been developed for image segmentation: graph-based (Shi and Malik, 2000), front evolution using level sets (Chan and Vese, 2001), and region growing (Gonzalez and Woods, 2002).

Front evolution is a popular technique in image segmentation. It consists of evolving a curve towards the lowest potential of a so-called cost function. The definition of the cost function reflects the task to be addressed, and imposes certain smoothness constraints. For the bi-partitioning segmentation problem, Zhu, Lee, and Yuille (1995) propose the following cost function:

$$\begin{aligned} E^{\mathrm{ZY}}(\mathbf{R}_1, \mathbf{R}_2, \alpha_1, \alpha_2) = &- \int_{\mathbf{R}_1} \log P\big(v(\mathbf{x}) : \alpha_1\big)\, \mathrm{d}\mathbf{x} \\ &- \int_{\mathbf{R}_2} \log P\big(v(\mathbf{x}) : \alpha_2\big)\, \mathrm{d}\mathbf{x} \\ &+ \lambda \oint_{\partial \mathbf{R}_1} \mathrm{d}\mathbf{s}, \end{aligned} \tag{1}$$

where $\partial \mathbf{R}_1$ is the interface between $\mathbf{R}_1$ and $\mathbf{R}_2$, and $P(v(\mathbf{x}) : \alpha_i)$ is the probability of the feature vector $v(\mathbf{x})$, *e.g.*, the gray level, at each position $\mathbf{x}$, α_i being the parameters of a pre-specified probability distribution.

The regions that minimize the function are the desired set of regions $\mathcal{R}$. The first and second term are the sum of the cost for coding the intensity of every $\mathbf{x}$ pixel inside the $\mathbf{R}_i$ according to its distribution. The third term is a regularization term, which penalizes large boundaries. The parameter $\lambda > 0$ is a weighting constant controlling the regularization.

Minimizing the function in Equation (1) is a difficult task and a lot of algorithms have been developed for it. In this work, we chose the level sets method, which has been proposed by Osher and Sethian (1988). The idea is to represent the evolving contour using a signed function, where the zero level corresponds to the actual contour. Hence, using the level-sets formalism one can describe the evolution of Equation (1) by means of the following equation:

$$\frac{\partial \Phi}{\partial t}(\mathbf{x}, t) = \left\| \nabla \Phi(\mathbf{x}, t) \right\| \left[\lambda \nabla \cdot \left(\frac{\nabla \Phi(\mathbf{x}, t)}{\| \nabla \Phi(\mathbf{x}, t) \|} \right) - \log \left(\frac{P(v(\mathbf{x}) : \alpha_1)}{P(v(\mathbf{x}) : \alpha_2)} \right) \right], \tag{2}$$

where $\Phi : \Re^2 \to \Re$ is the level-set function representing the regions $\mathbf{R}_1$ for $\Phi > 0$ and $\mathbf{R}_2$ for $\Phi < 0$, as well as the contour $\partial \mathbf{R}_1$ for $\Phi = 0$.

Several methods have been proposed to approximate the curve evolution. In particular in this work, we make use of the fast-evolution algorithm developed by Shi and Karl (2005).

2.2. The Gray Level Co-Occurrence Matrix (GLCM)

To discriminate between different features in the images (*e.g.*, foreground, white-light feature, background), it is necessary to distinguish, on a quantitative basis, the pixels that belong to the foreground from those that do not. The gray-level intensity of each pixel alone is generally insufficient to discriminate between them. An additional way to discriminate between foreground and background features is by taking into account their spatial complexity, as given by the local variations of intensity (hereafter "texture"). Haralick, Shanmugam, and Dinstein (1973) proposed the use of the so-called Gray Level Co-occurrence Matrix (hereafter GLCM) as a means to capture and therefore characterize the texture of the different regions.

Formally, the GLCM can be defined as follows: Let $\{J_k\}_{k=1\ldots K}$ be a limited set of image intensity levels, the GLCM $M_\Omega^{d,\phi}$ associated with a region Ω of an image I is defined at position $1 \leq i, j \leq K$ as the frequency of pairs of pixels $p, q \in \Omega$ such that $I(p) \in J_i$ and $I(q) \in J_j$ for a given displacement d and a given angle ϕ between both pixel positions.

In other words, the GLCM associated with a given region of an image is a matrix that contains information about the distribution of the intensity levels inside the region. The elements of the GLCM are simply the relative frequency of occurrence of pairs of gray-level values of pixels separated by a certain distance in a given direction. Or, from another point of view, if one looks at a pixel P_1 with intensity I inside the region, the GLCM tells how probable it is to find a pixel P_2 with intensity J at a distance d in the ϕ direction from P_1. For a deeper insight on the GLCM the reader is referred to Wagner (1999).

We computed the GLCM (for $d = 1$ and $\phi = 45°$) of several white-light features observed in the C2, C3, and COR2 coronagraph images by manually selecting the region of interest and found that the GLCM for the CME features does not follow exactly a known probability density function. In other words, we do not know either $P(v(\mathbf{x}) : \alpha_1)$ or α_1 [see Equation (1)]. The same is applicable to the background. Therefore, we had to generalize Equation (2) to circumvent the lack of knowledge of $P(v(\mathbf{x}) : \alpha_1)$ and α_1. For that task, we make use of the χ^2 statistical test, which allows us to evaluate if the observed events follow a well-known distribution or if two sets of events have the same distribution up to certain level of significance (Press *et al.*, 1992). The test for a second-order histogram is defined as

$$\chi^2\left(M_{\Omega_1}^{d,\phi}, M_{\Omega_2}^{d,\phi}\right) = \sum_{ij} \frac{((M_{\Omega_1}^{d,\phi})_{ij} - (M_{\Omega_2}^{d,\phi})_{ij})^2}{(M_{\Omega_1}^{d,\phi})_{ij} + (M_{\Omega_2}^{d,\phi})_{ij}}, \tag{3}$$

where $M^{d,\phi}_{\Omega_1}$ and $M^{d,\phi}_{\Omega_2}$ are the GLCM for the regions Ω_1 and Ω_2, respectively. The value of $\chi^2(M^{d,\phi}_{\Omega_1}, M^{d,\phi}_{\Omega_2})$ is zero when the histograms are equal and increases when the histograms differ.

In this way, we can modify Equation (2) to include the non-parametric method described above. Hence, Equation (2) becomes

$$\frac{\partial \Phi}{\partial t}(\mathbf{x}, t) = \left\|\nabla \Phi(\mathbf{x}, t)\right\| \left[\lambda \nabla \cdot \left(\frac{\nabla \Phi(\mathbf{x}, t)}{\|\nabla \Phi(\mathbf{x}, t)\|}\right) - \log\left(\frac{\chi^2(M^{d,\phi}_{W(\mathbf{x})}, M^{d,\phi}_{\Omega_1})}{\chi^2(M^{d,\phi}_{W(\mathbf{x})}, M^{d,\phi}_{\Omega_2})}\right)\right], \tag{4}$$

where $M^{d,\phi}_{W(x)}$ is the GLCM of an $m \times m$ window $W(\mathbf{x})$ centered at the pixel $\mathbf{x}$, and $M^{d,\phi}_{\Omega_1}$, $M^{d,\phi}_{\Omega_2}$ are the GLCM that captures the foreground and background texture, respectively. The size of the GLCM $[s]$ will depend upon the number of gray-intensity intervals chosen. For details on the derivation of Equation (4) the reader is referred to Goussies *et al.* (2009) and references therein. For the present work we used $d = 1$ and $\phi = 45°$.

Both, the window size $[m]$ and the GLCM size $[s]$ change the way in which the segmentation algorithm works. The bigger the window size, the less precise the algorithm to find the curve that best separates the event from the background. The reason for this is that the GLCM $M^{d,\phi}_{W(\mathbf{x})}$ of the pixels in the edge of the event are a mix of the GLCM for the event $M^{d,\phi}_{\Omega_1}$ and the background $M^{d,\phi}_{\Omega_2}$. On the other hand, if m is not big enough then there will not be enough samples to have the GLCMs populated. Therefore, there is a compromise between the size of the GLCM, the size of window, and the precision of the segmentation. By experimenting with different pairs (m, s) we found that, for an image size of 1024×1024 pixels, $m = 7$ and $s = 4$ gave the most reliable results for all of the cases considered.

3. Tracking

To track a coronal event in a video sequence means to detect (*i.e.*, segment) it on each frame. The way to do this is to see whether the texture that characterizes a given pixel $\mathbf{x}$ (*i.e.*, $M^{d,\phi}_{W(\mathbf{x})}$) resembles that of the foreground ($M^{d,\phi}_{\Omega_1}$) or that of the background ($M^{d,\phi}_{\Omega_2}$). This decision is made by comparing the values of $\chi^2(M^{d,\phi}_{W(\mathbf{x})}, M^{d,\phi}_{\Omega_2})$ and $\chi^2(M^{d,\phi}_{W(\mathbf{x})}, M^{d,\phi}_{\Omega_1})$ subject to the likelihood of neighboring pixels also belonging to the region.

Both, $M^{d,\phi}_{\Omega_1}$ and $M^{d,\phi}_{\Omega_2}$ must be computed at the beginning of the process. It is at this stage that the supervision aspects of the process present themselves. To create the foreground model, the user must identify the region of interest (*i.e.*, an approximate area comprising the CME feature) in the image when the event is first seen. Similarly, a region excluding the CME feature must be defined to be used for the background model.

Since there may be overlapping events, it is then required for the operator to identify again, in the frame where the event is first seen, the region where the event of interest is located by delineating with the mouse a contour around the feature. Note that the definition of the region of interest does not have to exactly follow the boundaries of the feature. The important issue is that the region so defined must include the main texture characteristics of the event. This region is then evolved to discriminate the feature defining the event from the background as explained in Section 2.2. The region obtained in this first frame is then used as the initial region to be evolved in the following frame. The procedure continues until the texture of the foreground becomes similar to the background texture.

However, reality imposes some caveats to this simple scheme. Firstly, the resulting segmentation may have holes, *i.e.*, excluded areas inside the extracted region. These holes

would correspond to regions inside the feature that are very faint and similar to the background. Since our aim is to delineate the outer boundaries of the event, its interior structure is of secondary importance so we applied morphological operations to suppress the holes.

Moreover, coronal events show up in a great variety of flavors, from both a kinematic and morphological point of view. For example, gradual events (*e.g.*, those generally associated with prominence eruptions), propagating close to the plane of the sky, generally maintain their morphology in the FOV of the coronagraphs. In other words, the texture of the feature remains similar during its development. On the other hand, explosive events, *i.e.*, those that usually start with large speeds, are not only morphologically different from the gradual events but they also do not keep a similar texture, especially during their early development. Moreover, the texture of the shock-like feature in those events greatly differs from that of the bulk of the event.

To solve those issues and, in particular, to be able to select which particular texture to follow, we introduced a modification to the original scheme, which consists of the re-evaluation of the foreground's GLCM depending upon the characteristics of the target coronal feature. This re-evaluation is performed by recomputing the GLCM of the foreground in the new frame $[i+1]$ in a region corresponding to an expanded version of the region found in the previous frame $[i]$, the extent of the expansion being Z/Q pixels in all directions. Z denotes the size of the segmented region found in frame i and Q a user-defined parameter between 1 and 60. For $Q = 0$ no expansion is performed. It is important to note that the value of this parameter is heuristically defined and passed as an input keyword to the program. Once the outcome is obtained, the user decides upon the success of the tracking. If the user is satisfied, the set of parameters used is archived in a database with the type of event tracked. Events with similar morphological characteristics will be analyzed with the same set of input parameters. Otherwise, the program is run again with a different Q. As the user gains experience, the most appropriate Q value for each class of event quickly becomes evident. See Sections 5 and 6 for details.

4. Data Preprocessing

While our segmentation algorithm performs well when applied to background-ratio images, events with poor contrast are practically impossible to track above certain heights regardless of the instrument used. To increase the contrast, we tried base difference images but we found that the output is strongly dependent on the artifacts introduced as a result of the processing. In particular, the further a given frame from the base frame, the more noticeable the streamer deflections (if any), which result in false detections. To avoid these false detections, it would be better to apply the algorithm on running differences. However, in this case, the inner part of the event keeps changing, altering its texture constantly and therefore losing the original meaning of tracking events by their texture.

Therefore, we developed a technique that seeks to combine the advantages of both difference-based techniques, namely the suppression of streamers (running difference) and the preservation of the inner morphological structure of the event (base difference). The technique consists of performing a running difference outside the boundaries of the feature and a base difference inside it according to

$$K_j(\mathbf{x}) = \begin{cases} I_1(\mathbf{x}) - I_{\text{base}}(\mathbf{x}) & \text{if } j = 1 \\ I_j(\mathbf{x}) - I_{\text{base}}(\mathbf{x}) & \text{if } j > 1 \text{ and } \mathbf{x} \in \mathbf{R}_{j-1}\,, \\ I_j(\mathbf{x}) - I_{j-n}(\mathbf{x}) & \text{if } j > 1 \text{ and } \mathbf{x} \notin \mathbf{R}_{j-1} \end{cases} \tag{5}$$

where $\mathbf{R}_j$ is the region of interest found at the frame j, and I_{base} is the base frame, *i.e.*, a frame prior to the first appearance of the event in the field of view of the instrument. For the nominal C2 cadence (20 minutes), C3 cadence (40 minutes), COR2 cadence (30 minutes), $n = 1$ is a suitable choice for fast events (*i.e.*, events exhibiting speeds above 1500 km s^{-1}). For slower events, $n \geq 2$ (the slower the event, the bigger the n to achieve better contrast at the outermost parts of the event). This choice was initially based on visual criteria and later confirmed by experimenting on a sample of carefully-chosen cases.

This differencing scheme is not only useful for streamer removal. A careful choice of n allows the isolation of concurrent events with different velocities. If n is large then both fast and slow events can be discerned. But on the other hand, if n is small slow event will be hard to distinguish, while a fast event will not be.

5. Results

We selected a sample of events recorded by the three coronagraphs already mentioned, *i.e.*, LASCO-C2 and C3, and SECCHI-COR2. The events were chosen based on their morphological complexity and kinematic characteristics to test the algorithm for a variety of common CME morphologies. Namely, a three-part structure CME on 14 March 2005 (C2 and C3), a slow halo-like event on 5 January 2005 (C3), a fast, asymmetric halo event with a clear associated shock on 15 – 16 January 2005, contaminated by the presence of a previous (slower) event (C3), and a limb event with an associated shock on 31 December 2007 (COR2A).

The LASCO C2 and C3 observing program ran on the nominal temporal cadence, *i.e.*, 20 minutes and 40 minutes, respectively. For LASCO we used level-05 images (although the algorithm also works without modifications on calibrated images). For COR2 we used the total brightness images [Tb], which were taken with a 30-minute cadence. Table 1 shows the time of first appearance, central position angle [PA], and angular width [ΔW] for the LASCO and SECCHI events selected as obtained from the CACTUS, CDAW, and SEEDS catalogs. Note that the halo event on 5 January 2005 was detected by both CACTUS and SEEDS as two different events. Likewise, the COR2 event on 31 December 2007 was detected in two parts.

In the following subsections we briefly describe the events and show the segmentations, *i.e.*, morphological characterization, obtained.

Table 1 Selected LASCO and SECCHI events. Time: time of first appearance in the C2/COR2 FOV, PA: position angle, and ΔW: angular width as obtained from the respective event catalogs.

Event	CACTUS			CDAW			SEEDS		
	Time	PA	ΔW	Time	PA	ΔW	Time	PA	ΔW
14 March 2005	07:48	258	104	08:00	259	105	08:36	260	99
05 January 2005	15:30	093	082	15:30	Halo	360	15:30	115	10
05 January 2005	15:54	162	032		–		15:30	92	72
15 January 2005	14:54	243	098	14:54	222	271	14:54	255	84
15 January 2005	23:06	322	352	23:06	Halo	360	23:30	243	60
31 December 2007	01:37	153	014		-COR2A-			-COR2A-	
31 December 2007	01:37	090	096		-COR2A-			-COR2A-	

5.1. Event on 14 March 2005

This event was a typical three-part structure CME. It was first seen in LASCO-C2 at 07:48 UT on the west limb. It developed as a bright and wide loop front, followed by a cavity and a bright, twisted, inner core. Some small ragged fronts followed later during the day.

Figure 1 shows a sequence of C2 and C3 frames with a pre-event image subtracted, the pre-event image (I_{base}) being the one at 07:48 UT and 08:18 UT, respectively. The first two rows show a sample of the development of the event in the C2 FOV (08:00, 08:24, 08:48, and 09:12 UT, respectively), while the last two rows show it in the C3 FOV (09:18, 10:42, 12:42, and 14:18 UT, respectively). The black line delineates the outermost edge of the event as determined by our algorithm applied on the differencing scheme depicted by Equation (5) for $n = 1$, *i.e.*, consecutive frames (first and third row), and $n = 5$, *i.e.*, five frames apart (second and fourth row). Comparison of rows 3 and 4, in particular the last two columns, shows the discrepancy that arises in the detection of this particular event when

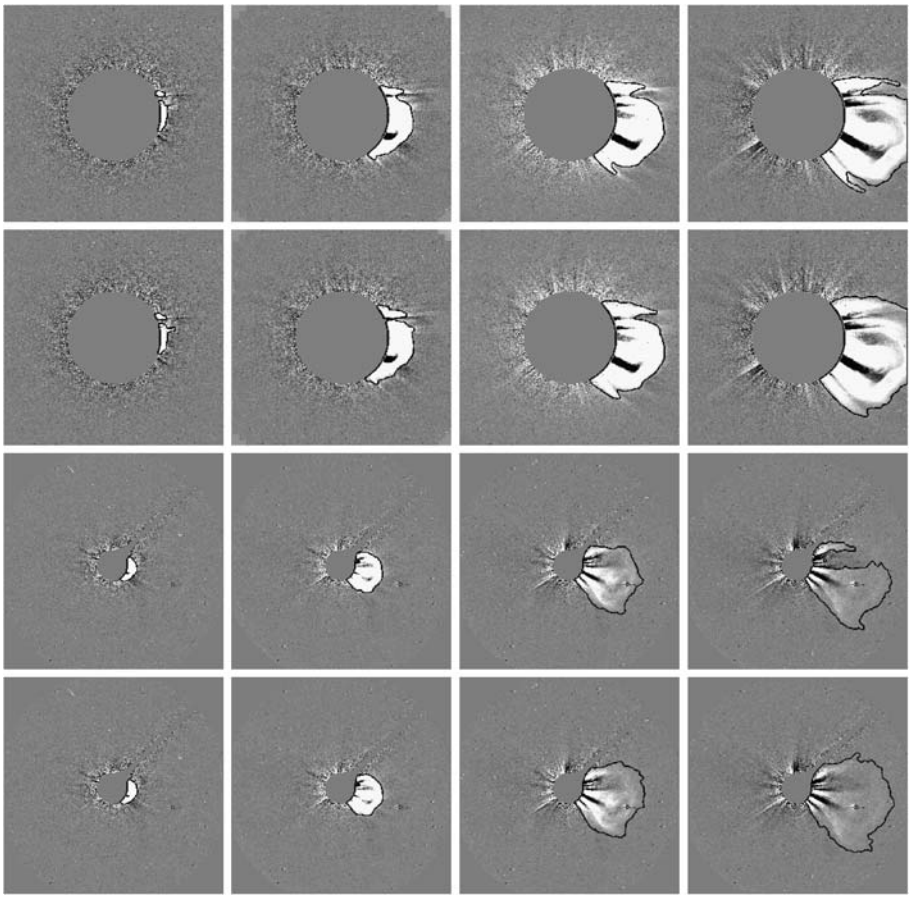


Figure 1 Event on 14 March 2005. C2 and C3 difference frames processed according to Equation (5). For each image we subtracted the nth previous frame. Top row: C2, $n = 1$. Second row: C2, $n = 5$. Third row: C3, $n = 1$. Bottom row: C3, $n = 5$.

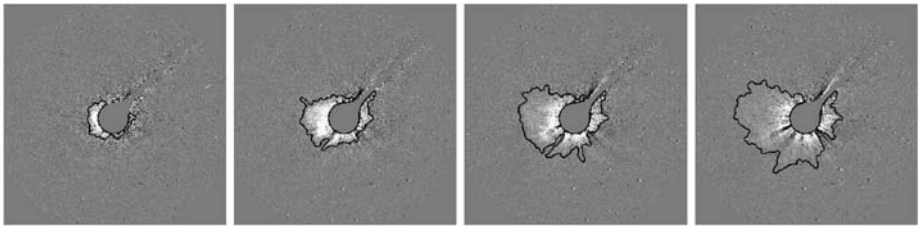


Figure 2 Event on 5 January 2005. LASCO C3 frames at 16:18, 17:42, 18:42, and 19:42 UT (from left to right). The black line delineates the outer envelope of the event as determined by our algorithm.

different n are used for the running-differences scheme. The segmentation obtained with $n = 5$ captures the event entirely.

5.2. Event on 5 January 2005

We chose this event as an example of a very faint and relatively slow halo-like CME in the LASCO-C3 field of view. Briefly, the event shows up in the LASCO-C2 FOV at 15:30 UT above the east limb as a diffuse, expanding, loop-like front followed by core material. By around 16:54 UT the C2 occulting disk seems to be completely covered. The event was first seen in C3 at 16:18 UT above the east limb, the C3 occulting disk being fully surrounded by 18:18 UT (signal barely visible to North). The mean plane-of-sky speed of the leading edge at PA ≈ 85 (where the feature seems to expand fastest) was about 725 km s^{-1} (slightly accelerated), based on C3 data (computed manually by point-and-click).

Figure 2 shows a C3 sequence of the event, the black line delineating the outer limits of the event as determined by our algorithm for $n = 5$ and I_{base} taken at 15:42 UT [see Equation (5)]. Note that the event, as defined by the black line, seems to expand faster along PA 80 – 85, which matches the PA chosen by the human observer to measure the maximum speed (see the description of the event in previous paragraph).

Note also that neither of the automated catalogs (see Table 1) reports the event correctly. They detect the halo as two separate events.

5.3. Event on 15 January 2005

This event was selected as an example of a fast asymmetric halo event in the C3 FOV with a clear associated shock, contaminated by the presence of a previous (slower) event. The halo event was first seen in LASCO-C2 at 23:06 UT above the north – northwest limb as a very bright loop front covering practically the whole northwest quadrant. By 23:30 UT the C2 occulting disk appears completely covered. Early on next day, a proton storm was detected in the C2 and C3 FOV. The event was first seen in LASCO-C3 at 23:18 UT above the northwest limb extending from the West up to past the North Pole. It is shown as a bright loop front surrounded by a fainter and diffuse envelope (most likely an associated shock). By 23:42 UT, the C3 occulting disk is fully surrounded. The mean plane-of-sky speed of the leading edge (measured manually at the outermost front of the diffuse envelope ahead of

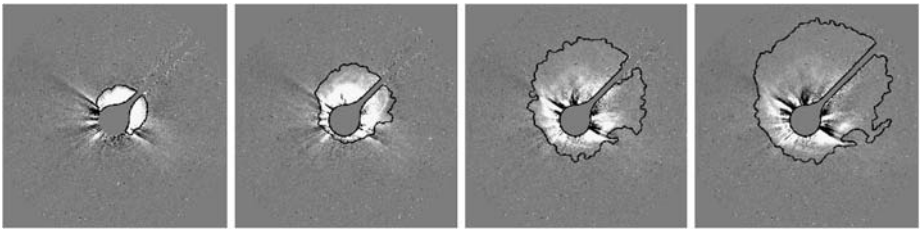


Figure 3 Halo event on 15 – 16 January 2005. LASCO C3 frames at 23:18, 23:42, 00:18, and 00:42 UT (from left to right). Same as Figure 2.

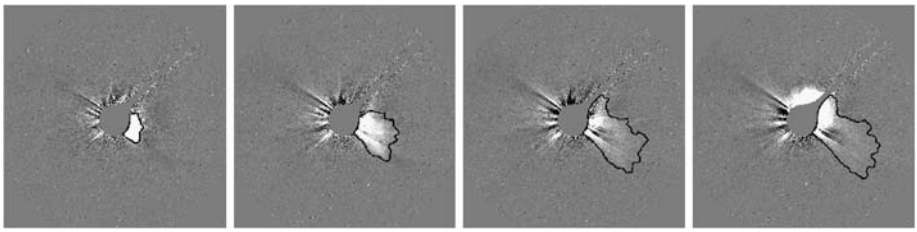


Figure 4 Event on 15 January 2005 that occurred prior to the halo event depicted in Figure 3. LASCO C3 frames at 17:18, 20:18, 22:18, and 23:18 UT (from left to right). Same as Figure 2.

the loop-like feature in the C3 FOV) at PA 329 was ≈2890 km s^{-1}. The mean plane-of-sky speed of the leading edge of the loop-like feature at the same PA was ≈2400 km s^{-1}.

The slower event mentioned above appeared in the C2 FOV above the west – southwest limb at 14:54 UT as a bright, ragged loop front leading a twisted trailing structure.

Figure 3 shows the development of the halo event in the C3 FOV, illustrating with the black line the outer limits of the associated shock. For this case we use [Equation (5)] $n = 2$ and I_{base} at 15:18 UT, and $Q = 16$ (see Section 3). Such an early I_{base} was used in order to have a pre-event image taken prior to the first appearance of both events in the C3 FOV. Figure 4 shows the segmentation of the slower event for comparison. We want to point out that the last frame already shows the commencement of the halo event, which our algorithm correctly identifies as a separate event because of its different texture. Note that although both automated catalogs (see Table 1) reports the event as a very wide one, none of them catalogues it as a halo.

5.4. Event on 31 December 2007

We chose this event as an example of a limb event exhibiting an associated shock in the FOV of the SECCHI-COR2A instrument. The event was first seen at 01:22 UT above the east – southeast limb. It develops as a big bright loop front. Starting at 01:52 UT, the deflection of the streamer on the North is noticeable. By 02:37 UT, a diffuse envelope starts developing ahead of the bright loop front, a likely signature of the associated shock responsible for the deflection of the streamer. Trailing material (including several ragged fronts) follows all throughout the day, with some inflows. The CACTUS catalog reports a median speed [V] of 781 km s^{-1}, a central position angle [PA] at 90°, and an angular extent [ΔW] of 96° for this event.

Figure 5 shows a sequence of the event in the COR 2A FOV. The frames were processed according to Equation (5) with $n = 5$ and I_{base} taken at 00:37 UT. The top row shows the

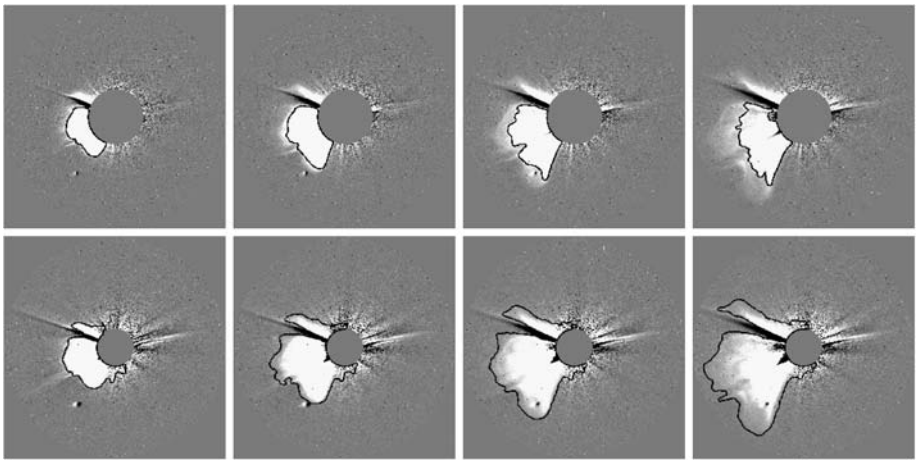


Figure 5 Event on 31 December 2007. SECCHI-COR2 A frames at 02:07, 02:37, 03:07, and 03:37 UT (from left to right). The black line delineates the segmentation obtained for the bulk of the event (top row), and the associated shock (bottom row).

segmentation of the bulk of the event. In this case, the deflected streamer at PA $\approx 75^\circ$ has been excluded by the tracking algorithm. The bottom row shows the segmentation of the event with the associated shock included. Unlike the previous case, the deflected streamer was included as part of the event. To achieve the distinction between the different segmentations, we varied the extent of the expansion of the segmented region found in a prior given frame to compute the new GLCM. In particular, $Q = 0$ for the bulk of the event (top row), *i.e.*, no expansion, and $Q = 16$ for the event including the shock (bottom row). See Section 3 for details.

For the SECCHI-COR2 data, only the CACTUS catalog measurements are available online. Table 1 shows CACTUS segmented the event in two parts; the main CME, and one at PA 90° which includes the streamer deflection in the North.

6. Discussion

As an example which highlights the pros and cons of CORSET, we discuss in more detail the quantitative measurements obtained from the morphological characterization of the 14 March 2005 event. The kinematical characterization is addressed in a forthcoming paper.

The left panel of Figure 6 shows the height – time diagram for the event at PA 261° as obtained from CORSET and the CDAW catalog. One immediately notices that the CORSET measurements diverge from the human-derived ones with time and that the CORSET heights are consistently lower (by $\approx$10%) to the CDAW heights. The reason for this discrepancy is demonstrated in the right panel of Figure 6, where we see that the manually-selected points do not follow a single PA. This example demonstrates the advantages of our technique. Unlike human determinations, each CORSET determination is derived from a unique and quantitative gauge driven by the texture changes between the event and its background. Moreover, CORSET allows an automatic, and consistent, determination of the height – time profile along any position angle without the uncertainty inherently introduced by manually selecting the front.

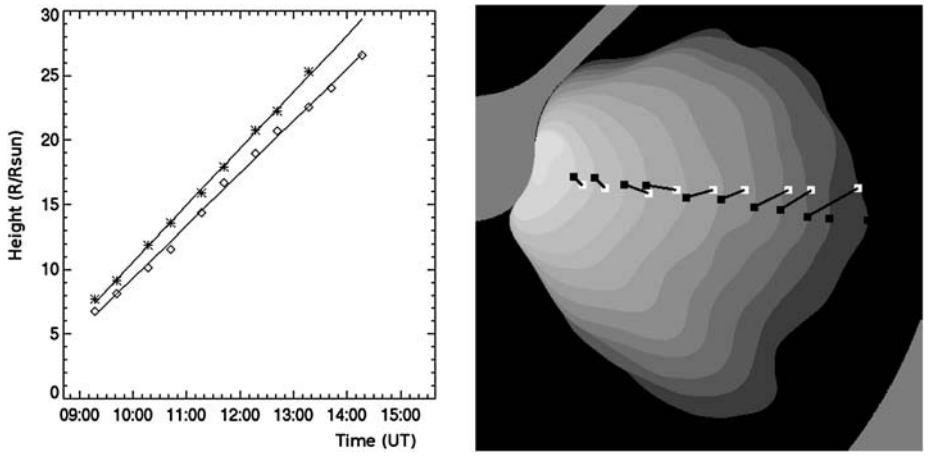
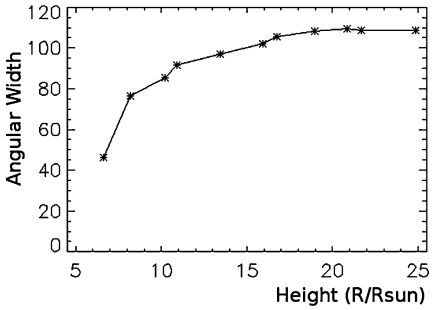


Figure 6 Left panel: Height – time plot for the 14 March 2005 event. The diamonds (asterisks) correspond to the CORSET (CDAW) determination of the leading edge at position angle 261. Right panel: Motion history of the event (outer edge low-pass filtered) in the C3 FOV as obtained with CORSET at (from left to right): 08:42, 09:18, 09:42, 10:18, 10:42, 11:18, 11:42, 12:18, 12:42, 13:18, 13:42, and 14:18 UT. The black (white) points depict the CORSET (CDAW) determinations of the leading edge at PA 261°. The measurements start at 09:18 UT and extend up to 14:18 UT (CORSET) and 13:18 UT (CDAW). The black lines joining the black and corresponding white dots are intended to show the correspondence between respective measurements in a given frame.

Figure 7 Angular width [degrees] of the 14 March 2005 event as a function of height.



Unlike other CME-tracking techniques (*e.g.*, Olmedo *et al.*, 2008; Byrne *et al.*, 2009) CORSET does not attempt to fit the edges of the event to a given geometrical model. An inherent advantage of our approach is that varying height – time plots of the leading edge at different position angles (not shown here) could be due to subtle differences in the propagation of the different parts of the event, and could reveal components or a variable solar-wind background. We think that this seems to be the case in this 14 March 2005 event. The time sequence of the boundary maps, in the right panel of Figure 6, shows the gradual appearance of a lobe along the CME southern edge by 11:18 UT (PA 230°, sixth map starting from the innermost part).

Since we obtain measurements at several PAs, we can easily construct the evolution of the angular width of the event as a function of height (Figure 7). We see that this parameter tends to ≈102° matching the angular width reported in the three catalogs (see Table 1). One obvious advantage of the segmentation obtained with CORSET is the possibility to study the rate of growth of this parameter, something that is very difficult to do with other catalogs.

The rate of growth roughly determines the location of the source region without the aid of any other instrument (*e.g.*, based on the projection onto the plane of sky of a cone-shaped CME as a function of the CME angular width and source region longitude.

7. Conclusions and Outlook

We have developed a new technique to automatically determine the boundaries of events in white-light coronagraphs based on their texture (Figures 1 – 5). Applying the technique to a time sequence, it can track the event as it transits through the field of view of the coronagraph. In this paper we have outlined the technique's main characteristics, and demonstrated its performance *i*) on following a variety of events, *ii*) on distinguishing the main event from the associated shock (when present), and *iii*) on separating simultaneous events. The keys to its success are a careful selection of the pre-event image [I_{base}], the time separation used for the difference [n], and the extent of expansion of the GLCM [Q]. Furthermore, the algorithm is able to work with data from any coronagraph and is relatively fast ($\approx$30 s per 1024^2 image). Some parameters of the method, in particular Q and n, are not fixed, but are modified based on the characteristics of the particular CME. However, events that share common morphological characteristics may be treated with a single given set of parameters which changes for different event types. This is intriguing as it mimics the adaptability of the human observer to adjust their eyes and their interpretation of individual CMEs. The boundaries of some CMEs can be easily defined and all observers (or automatic techniques) will determine the same or very similar definitions of the CME extent. But determining the boundaries of other CMEs can be very difficult for a variety of reasons such as the faintness of the CME, multiple overlapping CMEs, or pre-existing structures that get deflected or affect the propagation of the event. Our technique has proven to be an important step forward, as it does quantitatively and automatically what a human observer does qualitatively and interactively.

The algorithm, as currently developed, requires the selection of the region of interest (*i.e.*, a rough estimation of the initial location of the event) at the time of its first appearance. Goussies *et al.* (2009) have shown that this step can be automatized for detection purposes. Our focus in this paper was the tracking of the event and its morphological characterization, which is the first step before determining its kinematical and dynamical parameters to compare, for examples with theoretical models. The algorithm has the potential to provide a significant improvement to the consistency and completeness of CME measurements.

Acknowledgements SOHO is a project of international cooperation between ESA and NASA. The STEREO/SECCHI data are produced by a consortium of NRL (US), LMSAL (US), NASA/GSFC (US), RAL (UK), UBHAM (UK), MPS (Germany), CSL (Belgium), IOTA (France), and IAS (France).

References

Boursier, Y., Llebaria, A., Goudail, F., Lamy, P., Robelus, S.: 2005, Automatic detection of coronal mass ejections on LASCO-C2 synoptic maps. In: Fineschi, S., Viereck, R.A. (eds.) *Soc. of Photo – Optical Instr. Eng. (SPIE) Series, CS* **5901**, 13 – 24. doi:10.1117/12.616011.

Brueckner, G.E., Howard, R.A., Koomen, M.J., Korendyke, C.M., Michels, D.J., Moses, J.D., Socker, D.G., Dere, K.P., Lamy, P.L., Llebaria, A., Bout, M.V., Schwenn, R., Simnett, G.M., Bedford, D.K., Eyles, C.J.: 1995, The large angle spectroscopic coronagraph (LASCO). *Solar Phys.* **162**, 357 – 402. doi:10.1007/BF00733434.

Byrne, J.P., Gallagher, P.T., McAteer, R.T.J., Young, C.A.: 2009, The kinematics of coronal mass ejections using multiscale methods. *Astron. Astrophys.* **495**, 325 – 334. doi:10.1051/0004-6361:200809811.

Chan, T.F., Vese, L.A.: 2001, Active contours without edges. *IEEE Trans. Image Process.* **10**(2), 266 – 277.
Gonzalez, R.C., Woods, R.E.: 2002, *Image Processing*, 2nd edn. Addison-Wesley, Boston. ISBN 0-201-18075-8.
Goussies, N., Mejail, M., Jacobo, J., Stenborg, G.: 2009, Detection and tracking of coronal mass ejections based on supervised segmentation and level sets. *Pattern Recog. Lett.* doi:10.1016/j.patrec.2009.07.011
Haralick, R.M., Shanmugam, K., Dinstein, I.: 1973, Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **SMC3**, 610 – 621.
Howard, R.A., Moses, J.D., Vourlidas, A., Newmark, J.S., Socker, D.G., Plunkett, S.P., Korendyke, C.M., Cook, J.W., Hurley, A., Davila, J.M., Thompson, W.T., St Cyr, O.C., Mentzell, E., Mehalick, K., Lemen, J.R., Wuelser, J.P., Duncan, D.W., Tarbell, T.D., Wolfson, C.J., Moore, A., Harrison, R.A., Waltham, N.R., Lang, J., Davis, C.J., Eyles, C.J., Mapson-Menard, H., Simnett, G.M., Halain, J.P., Defise, J.M., Mazy, E., Rochus, P., Mercier, R., Ravet, M.F., Delmotte, F., Auchere, F., Delaboudiniere, J.P., Bothmer, V., Deutsch, W., Wang, D., Rich, N., Cooper, S., Stephens, V., Maahs, G., Baugh, R., McMullin, D., Carter, T.: 2008, Sun Earth connection coronal and heliospheric investigation (SECCHI). *Space Sci. Rev.* **136**, 67 – 115. doi:10.1007/s11214-008-9341-4.
Liewer, P.C., Dejong, E.M., Hall, J.R., Lorre, J.J.: 2005, Automatic CME detection from coronagraph image pairs. In: *AGU Fall Meeting Abstracts*, SH14A-01.
Olmedo, O., Zhang, J., Wechsler, H., Poland, A., Borne, K.: 2008, Automatic detection and tracking of coronal mass ejections in coronagraph time series. *Solar Phys.* **248**, 485 – 499. doi:10.1007/s11207-007-9104-5.
Osher, S., Sethian, J.A.: 1988, Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulations. *J. Comput. Phys.* **79**, 12 – 49.
Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T.: 1992, *Numerical Recipes: The Art of Scientific Computing*, 2nd edn. Cambridge University Press, Cambridge. ISBN 0-521-43064-X.
Qu, M., Shih, F.Y., Jing, J., Wang, H.: 2006, Automatic detection and classification of coronal mass ejections. *Solar Phys.* **237**, 419 – 431. doi:10.1007/s11207-006-0114-5.
Robbrecht, E., Berghmans, D.: 2004, Automated recognition of coronal mass ejections (CMEs) in near-real-time data. *Astron. Astrophys.* **425**, 1097 – 1106. doi:10.1051/0004-6361:20041302.
Robbrecht, E., Berghmans, D., Van der Linden, R.A.M.: 2009, Automated LASCO CME catalog for solar cycle 23: Are CMEs scale invariant? *Astrophys. J.* **691**, 1222 – 1234. doi:10.1088/0004-637X/691/2/1222.
Shi, J., Malik, J.: 2000, Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(8), 888 – 905.
Shi, Y., Karl, W.: 2005, Real-time tracking using level sets. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* **2**, 34.
Wagner, T.: 1999, Texture analysis. *Handbook of Computer Vision and Applications* **2**, 276 – 308. ISBN 0-12-379772-1.
Yashiro, S., Gopalswamy, N., Michalek, G., St. Cyr, O.C., Plunkett, S.P., Rich, N.B., Howard, R.A.: 2004, A catalog of white light coronal mass ejections observed by the SOHO spacecraft. *J. Geophys. Res.* **109**, 7105. doi:10.1029/2003JA010282.
Zhu, S., Lee, T., Yuille, A.: 1995, Region competition: Unifying snakes, region growing, energy/bayes/mdl for multiband image segmentation. In: *Proc. Fifth IEEE Int. Conf. Computer Vision (ICCV)*, 416 – 423.