

Combining STEREO SECCHI COR2 and HI1 images for automatic CME front edge tracking

Vladimir Kirnosov^{1,*}, Lin-Ching Chang¹, and Antti Pulkkinen²

¹ Electrical Engineering & Computer Science Dept., The Catholic University of America, 620 Michigan Ave, N.E., Washington, DC 20064, USA

*Corresponding author: 57kirnosov@cua.edu

² NASA Goddard Space Flight Center, Code 674, Greenbelt, MD 20771, USA

Received 3 December 2015 / Accepted 31 October 2016

ABSTRACT

COR2 coronagraph images are the most commonly used data for coronal mass ejection (CME) analysis among the various types of data provided by the STEREO (Solar Terrestrial Relations Observatory) SECCHI (Sun-Earth Connection Coronal and Heliospheric Investigation) suite of instruments. The field of view (FOV) in COR2 images covers 2–15 solar radii (Rs) that allow for tracking the front edge of a CME in its initial stage to forecast the lead-time of a CME and its chances of reaching the Earth. However, estimating the lead-time of a CME using COR2 images gives a larger lead-time, which may be associated with greater uncertainty. To reduce this uncertainty, CME front edge tracking should be continued beyond the FOV of COR2 images. Therefore, heliospheric imager (HI1) data that covers 15–90 Rs FOV must be included. In this paper, we propose a novel automatic method that takes both COR2 and HI1 images into account and combine the results to track the front edges of a CME continuously. The method consists of two modules: pre-processing and tracking. The pre-processing module produces a set of segmented images, which contain the signature of a CME, for both COR2 and HI1 separately. In addition, the HI1 images are resized and padded, so that the center of the Sun is the central coordinate of the resized HI1 images. The resulting COR2 and HI1 image set is then fed into the tracking module to estimate the position angle (PA) and track the front edge of a CME. The detected front edge is then used to produce a height-time profile that is used to estimate the speed of a CME. The method was validated using 15 CME events observed in the period from January 1, 2008 to August 31, 2009. The results demonstrate that the proposed method is effective for CME front edge tracking in both COR2 and HI1 images. Using this method, the CME front edge can now be tracked automatically and continuously in a much larger range, i.e., from 2 to 90 Rs, for the first time. These improvements can greatly help in making the quantitative CME analysis more accurate and have the potential to assist in space weather forecasting.

Key words. STEREO SECCHI – CME – Front edge tracking – COR2 – HI1

1. Introduction

Coronal mass ejections (CMEs) are large-scale expulsions of plasma and magnetic field from the solar atmosphere (Tousey 1973; Brueckner 1974; MacQueen et al. 1974). It is well known that they play a crucial role in disturbing the space environment as they propagate into the interplanetary medium and interact with the Earth's magnetic field producing strong geomagnetic storms (Allen 1944; Burlaga & Lepping 1977; Gosling et al. 1990; Crooker & McAllister 1997; Srivastava & Venkatakrishnan 2002, 2004; Srivastava et al. 2009). An ability to track a CME through the heliosphere would give a far better estimate of the arrival time, the mass, and the structure of the event as the CME passes over the Earth (Harrison et al. 2007). To study the nature of CMEs, NASA (National Aeronautics and Space Administration) and ESA (European Space Agency) have launched several missions. One of the missions is the twin Solar Terrestrial Relations Observatory (STEREO) (Kaiser et al. 2008). To accomplish scientific goals, the STEREO mission was divided into the prime and extended periods containing a total of four phases. The STEREO prime mission covers the first 2 years after the spacecraft reached the heliospheric orbit on January 21, 2007. Phase 1 occurs approximately the first year when the spacecraft are less than 50° apart. Phase 2 is centered on a quadrature between the two

spacecraft with separations between 50° and about 110°, corresponding to days 400–800 counted from the launch of the spacecraft.

STEREO uses two spacecraft to capture CMEs from two different viewpoints, which allow for three-dimensional (3D) observation of the events. The images of CMEs are captured by the two coronagraphs (COR1 and COR2) and two heliospheric imagers (HI1 and HI2) included in the Sun-Earth Connection Coronal and Heliospheric Investigation (SECCHI) package of instruments (Howard et al. 2008) on board the spacecraft. Each instrument provides a two-dimensional (2D) representation of a three-dimensional (3D) CME structure projected into the sky plane. As a consequence, the estimated CME propagation parameters, such as position angle, angular width, height, and speed, are also projected into the same plane (Mierla et al. 2010). These CME parameters can be estimated using COR and/or HI images commonly used to assist in space weather forecasting. The propagation parameters extracted from a single point of view (i.e., one spacecraft) do not give reliable information on the actual CME propagation path. Extraction of the parameters from two vantage points (i.e., both spacecraft), however, makes reconstruction of a true CME propagation path in 3D space possible (Howard & Tappin 2008; Mierla et al. 2008, 2010; Boursier et al. 2009; Colaninno & Vourlidas 2009; de Koning et al. 2009;

de Koning & Pizzo 2011). The reconstruction can be accomplished using a technique of geometric triangulation, such as that proposed by Liu et al. (2010) and Mays et al. (2015). Prior to using a geometric triangulation technique, one needs to determine the location of a CME front edge in each image in stacks from STEREO A and B. The method proposed in this study allows for tracking a CME front edge in STEREO A/B COR2 and HI1 image stacks and can be integrated with the triangulation technique to provide a fully automated algorithm to estimate a CME propagation path in 3D space.

The CME front edge and propagation parameters can be estimated semi-manually. For example, to analyze COR2 data, a Stereoscopic CME Analysis Tool (StereoCAT) (<http://ccmc.gsfc.nasa.gov/analysis/stereo/>; Mays et al. 2015) can be used. This tool requires a user to manually identify and trace a CME front edge in each image in a stack, and based on that, it automatically derives the parameters. Although such tools are of great help to researchers, the manual processing part is a very time-consuming task. The result of manual visual analysis is also subjective to the experience and perception of the observer. In addition, the volume of STEREO SECCHI data is huge and has been increasing constantly. All these facts underline the need for an automatic method that provides a faster, systematic, and accurate way to estimate CME kinematic properties.

Several CME detection methods have been proposed and most of them use white-light coronagraph images from SOHO (Solar and Heliospheric Observatory) LASCO (Large Angle and Spectrometric Coronagraph). For example, Computer Aided CME Tracking (CACTus) has been used to automatically construct a CME catalog (Robbrecht & Berghmans 2004; Robbrecht et al. 2009). Other methods include the Solar Eruptive Event Detection System (SEEDS), presented by Olmedo et al. (2008), Boursier et al.'s (2009) ARTEMIS: Automatic Recognition of Transient Events and Marseille Inventory from synoptic maps, and Coronal Image Processing (CORIMP), presented by Byrne et al. (2012). In addition, the CACTus and SEEDS are also capable of processing STEREO SECCHI COR2 data and deriving CME kinematic parameters such as height and speed. However, the method used in CACTus is not designed to track the front edge of a CME (Yashiro et al. 2008). On the other hand, the SEEDS catalog is capable of tracking the front edge, however, their module to process STEREO SECCHI COR2 images was added to the SEEDS recently and has not yet been rigorously tested. As noted on the SEEDS website (<http://spaceweather.gmu.edu/seeds/>), the STEREO module has not been rigorously tested and the data should be treated carefully.

Young & Gallagher (2008) have proposed a multiscale edge detection method that was further studied by Byrne et al. (2009, 2010, 2012), Gallagher et al. (2011), Pérez-Suárez et al. (2011), and Morgan et al. (2012). This method uses a set of wavelet- and gradient-based techniques to locate the front edge of a CME and has been employed in the CORIMP catalog for SOHO LASCO data. The CORIMP catalog for STEREO has not been developed and reports contain STEREO processing results with only a few CME events. Tappin et al. (2012) have also proposed a method to detect CMEs in heliospheric images captured by the Coriolis SMEI (solar mass ejection imager) instrument. They have reported that their method is effective as a region of interest highlighter and a promising first step in autonomous heliospheric imager CME detections. To characterize the 3D structure of solar transients,

Pulkkinen et al. (2009) have introduced a method that allows for the automatic determination of cone model parameters from white-light coronagraph images. They reported that the proposed method shows reasonable agreement in CME parameter estimations using SOHO LASCO C3 difference images and can consequently become a part of a fully automatic CME detection and 3D characterization system.

Despite these earlier works, there are currently no automatic, reliable methods to track a CME front edge and estimate the propagation parameters from combining the STEREO SECCHI COR2 and HI1 images.

In comparison to the above-mentioned automatic techniques, our method is a more generic algorithm that allows for tracking a CME in both COR2 and HI1 fields of view (FOVs). As a main difference to SEEDS and CACTus catalogs, our detection method is not based on traditional running-difference images where CME features and background noise both present in the images. Our method first pre-processes data to suppress background noise and to segment moving CME features. Only segmented CME patterns are then used to detect front edge and estimate the propagation parameters. Speeds estimated by CACTus represent the motion of the brightest material of CMEs; therefore, CACTus speeds are lower and assumed to be approximately constant (Yashiro et al. 2008; Tappin et al. 2012; Byrne 2015). On the contrary, in our method, the speed of a CME is estimated using the furthestmost (fastest) point in the front edge. Comparing the capabilities of our method with those of CORIMP in STEREO COR2, we noted that the workable field of CORIMP is limited to 3–11 Rs for COR2A and 4.5–11 Rs for COR2B (Byrne et al. 2012) (potential FOV is 2–15 Rs). Our method does not have such limitations and allows for tracking the front edge of a CME covering the whole FOV of a COR2 image (i.e., 2–15 Rs).

COR2 images are the most commonly used data for CME front edge detection among the various data types provided by the STEREO SECCHI suite of instruments. The FOV in COR2 images is from 2 to 15 Rs (Kaiser et al. 2008) that allows for larger lead-time forecasting of further CME developments. Such an early forecast, however, can be associated with greater uncertainty. To reduce the uncertainty, we herein propose to continue tracking the front edge of a CME beyond the COR2 FOV, i.e., by including the HI1 data. The HI1 FOV is from 15 to 90 Rs (Eyles et al. 2009).

Using both the COR2 and HI1 images for CME tracking is beneficial. Readers may refer to the work by Harrison et al. (2007) for more information on the advantages of tracking a CME using COR2 and HI1. In the combined COR2 and HI1 FOVs, a CME can be tracked continuously up to the distance of 90 Rs. Therefore, the parameters of the CME can be determined more precisely, which can lead to increasingly accurate forecasts of a CME propagation path.

In this paper, we propose a novel automatic pipeline for efficient CME front edge detection and propagation parameter estimation using COR2 and HI1 science quality images. The images were captured in the ecliptic plane and their FOVs are in the plane of the sky. The proposed method allows for tracking the CME front edge and estimating propagation parameters in these FOVs across the range of 2–90 Rs. Validation was performed using 15 CME events captured in the period of the primary phase of the STEREO mission, from January 1, 2008 to August 31, 2009. Each data set consists of both series of STEREO A COR2 and HI1 RGB images in science quality.

2. STEREO SECCHI COR and HI data

STEREO SECCHI has two separate telemetry data streams coming down from each spacecraft: the space weather beacon telemetry and the science recorder playback telemetry. The beacon telemetry contains the most recent data and images which have been transmitted 24 h per day. The beacon data transmission rate is very low, the images are compressed by large factors and are thus of much lower quality than the actual science data that reach the STEREO Science Center (SSC) with a delay of several days. All data and images obtained by the STEREO project are available to the general public within a short period of time (Kaiser et al. 2008).

Both science and beacon data are freely available for researchers in two types of data formats: Flexible Image Transfer System (FITS) and JPEG. FITS files are the calibrated versions of raw images captured by the STEREO A/B spacecraft. The FITS files are 16-bit gray scale images (single channel, unsigned integer with 65,536 possible intensity levels). The feature of the FITS format is that the header of each file contains photometric and calibration information, along with metadata that describe the file and its origin. By looking into the file's header, one may determine the instrument that captured the image, coordinates of the spacecraft at that moment, telemetry type, photometric details, any processing steps that were done to the image, and much more.

The STEREO SECCHI FITS files can be processed using routines included in a SolarSoft package and publicly available in Interactive Data Language (IDL) (<http://exelisvis.com/ProductsServices/IDL.aspx>). The SolarSoft (<http://www.lmsal.com/solarsoft/>) contains a set of software libraries, utilities, and databases, which all together provide a comprehensive environment for programming support and solar data analysis. For more information on the SECCHI software included into the SolarSoft package, the reader can refer to the SECCHI Wiki online page (<http://secchi.nrl.navy.mil/wiki/pmwiki.php?n=Main.DataProcessingAndAnalysis>). This software has been used at SSC at NASA to process and calibrate SECCHI FITS data upon receiving them from the spacecraft. It has also been used at SSC to convert FITS images to Red, Green, Blue model (RGB) browse images in JPEG format. The reader, if interested, can refer to the conversion procedure written in IDL (http://darts.jaxa.jp/pub/ssw/sterео/ssc/idl/browse/ssc_browser_secchi_jpeg.pro). The process of conversion includes adding a color to the browse images. This is done on purpose as each instrument is associated with a particular color. The color maps used are the standard IDL's color tables. For COR2 images the color table #3 ("red temperature") is used and for HI1 – color table #1 (blue-white linear). The resulting colored JPEG RGB browse images contain 24-bits-per-pixel (8 bits per channel; each channel has 256 intensity levels).

FITS images have traditionally been used for automatic CME analysis. The colored JPEG browse images, on the other hand, have been used for visual CME inspection by space weather forecasters and scientists. In this study, we show that in addition to visual inspection by a human operator, the JPEG browse images can also be used for automatic CME detection and tracking. This allows for using a single set of data for both: visual and automatic CME analysis, thus avoiding possible discrepancy during comparison of the results that may take place when the two data sets are used (16-bit gray FITS and 24-bit colored JPEG). The conversion from single-channel 16 bits to three-channel 24 bits may lead to modification of

features, so the representation of a CME in raw FITS files and JPEG browse images may vary slightly. Although the colored JPEG browse images are used in this study, there is no doubt that for more sophisticated CME analysis, FITS files should be employed.

3. Method

In this section, we provide an overview of the proposed method for CME front edge tracking and propagation parameters' estimation. The proposed algorithm is designed to process two types of STEREO SECCHI data: COR2 and HI1 JPEG browse images. The proposed method consists of two modules: pre-processing and tracking modules. The pre-processing module converts input COR2 and HI1 JPEG browse images into a gray scale format, removes noise, and segments CME material. HI1 images are further padded in this module to include the position of the Sun as the central coordinate. The resulting set of binary images is then fed into the tracking module for estimating the CME propagation parameters, which include the position angle, the speed, and the front edge of the CME.

3.1. Images used in this study

The data used in this paper are the STEREO SECCHI COR2 and HI1 science quality images captured in the period from January 1, 2008 to August 31, 2009. The images were obtained via SSC's File Transfer Protocol (FTP) site (<ftp://stereoftp.nascom.nasa.gov/pub/browse/>). To download images from the SSC's FTP site, one can use any FTP client software (e.g., Total Commander, FAR, etc.). All images used in this study are in JPEG RGB format with a color depth of 24-bits-per-pixel (8 bits per channel; each channel has 256 possible intensity levels) and with a dimension of 256×256 pixels. The COR2 images are red-colored and HI1 images are blue-colored. These images were generated at SSC from the calibrated versions of raw image files in FITS format. For more information, please refer to Section 2 of this manuscript.

3.2. Pre-processing module

The workflow of the pre-processing module is similar to the pre-processing approach in Kirnosov et al. (2015). Kirnosov et al.'s (2015) method was designed to process only COR2 JPEG science quality images. In this study, we improve the method and propose a more generic pre-processing pipeline, which allows for using both COR2 and HI1 JPEG data sets in science quality.

Given an image stack (COR2 or HI1), each color image is first decomposed into its three channels (red, green, and blue). The format of an RGB color image is 24-bits-per-pixel, which represents three color channels with 8 bits for each channel. Therefore, decomposition of an RGB color image into three channels is straightforward.

When processing COR2 data, we extract and use only the red channel images (e.g., Kirnosov et al. 2015). For HI1 data, however, we use "green with red subtracted" images which we produce by subtracting red channel images from green channel images. Such subtraction is beneficial as it gives the effect of applying a noise filter, i.e., it removes stars, planets, and other bright non-CME objects captured in the HI1 green component.

The blue and green channels are ignored in COR2 data, as well as blue in HI1 data since they are not representative of CME signature in its full extent and do not contain information useful for CME detection (e.g., Kirnosov et al. 2015). As depicted in the ‘‘FITS to JPEG’’ conversion procedure, used at SSC at NASA to process data upon receiving them from the spacecraft (<http://darts.jaxa.jp/pub/ssw/stereo/ssc/idl/browse/ssc Browse secchi jpeg pro>), the color maps used to colorize images are the standard IDL color tables. For COR2 images, color Table #3 (so-called ‘‘red temperature’’) is used, and most of the information kept in FITS images is assigned to the red channel, instead of the blue and green channels. Similarly, for HI1 images, the color Table #1 (blue-white linear) is applied, making the green channel have more intensity levels and better contrast, representing CME in its full extent.

The second step is to reduce the noise in the images produced in the previous step, by applying a 3×3 median filter to the images. The kernel size was experimentally determined to be effective for noise reduction in our application.

The third step is to remove the constantly bright areas (e.g., streamers) and to segment the CME material. To perform these tasks, the mean image of the stack is computed. This mean image is then subtracted from each image in the stack, thus producing a set of images, which contain pixels that differ from the mean image. The streamers and other constant non-CME regions are eliminated by this procedure.

The fourth step is to extract the moving features that represent the CME material and remove any unwanted regions that might still be present in the images. The moving features are detected by subtracting the preceding image from the current one, producing a set of running-difference images. Then, a temporary binary version of this set of images is generated (i.e., all pixels greater than 0 are set to 1) and morphological erosion with a structuring element (SE) 3×3 is applied to these binary images. This binary set is used as a mask. The pixels that were set to zero during the erosion in the binary images are set to zero in the relevant running-difference images as well. To restore the size of the segments, the set of running-difference images is converted into the binary format once again and a morphological closing with the same structuring element SE is applied. These binary images are the output of this step. Note that if the input images are from the STEREO B, the order of columns in COR2 or HI1 images would be reversed at this point, so the Sun-Earth line is on the left as in STEREO A FOV. This allows use of the same algorithm for both STEREO A and B images.

For COR2 data, the output of step four is the final result of the pre-processing module. Note that in a COR2 image, the Sun is located within the FOV. The center of Sun is in the central coordinate of the image. For HI1 data, the Sun is located outside the FOV. Therefore, an additional procedure is performed to adjust the size of the HI1 image, so the central coordinate of the image is referred to the center of the Sun. The FOV in an HI1 image is from 15 to 90 Rs (Eyles et al. 2009) and the size is 256×256 pixels. The total FOV of 75 Rs would correspond to 256 columns in the scale of digital images, i.e., one solar radius is 3.4 columns. The distance from the center of the Sun to the beginning of the FOV in HI1 images is 15 Rs, which corresponds to 51 columns in an image. To produce a new HI1 image with a central coordinate referring to the center of the Sun, we append 358 columns to the right side of the images (see Fig. 1). The resized HI1

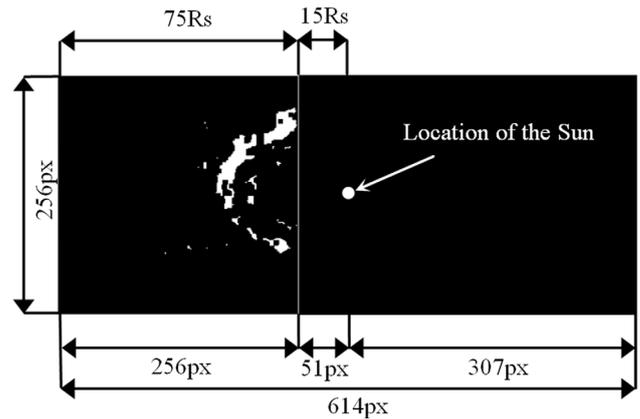


Fig. 1. An example of the resized HI1 image. The central point with the coordinates [307, 128] in the image refers to the center of the Sun.

images are of size 614×256 where the central coordinate [307, 128] would refer to the center of the Sun. Such adjustment is required in order to properly rotate the images for propagation parameter estimation in the tracking module. Figure 2 shows an example of the final result of the pre-processing module using a selected CME event (event #2 in Table 1); we refer to the set of these pre-processed images as the SegmSet for the remainder of the paper.

3.3. Tracking module

The input to the tracking module is the set of binary segmented images (SegmSet) produced by the pre-processing module. The tracking module is designed based on the fact that a CME can propagate only in the outward direction. When looking through the pixels along each degree in a sequence of segmented images, one can notice that only a particular angular span is representative of a CME. Moreover, in some specific directions (degrees), the CME material can be tracked further than in others. Therefore, an angular span of a CME, a position angle, and a distance that it moves can be determined.

In this module, we first estimate a primary propagation direction of CME central material, i.e., position angle (PA), which is then used to detect and track a CME front edge. Corresponding to the PA, a maximum number of images can be found in which the CME material can be tracked. To estimate PA, we look through the pixels along each degree on the left half of each SegmSet, i.e., the area shown in Figure 3a for COR2 images and Figure 3b for HI1 image. PA is measured counterclockwise starting from solar north at 0° . For COR2 we iterate from 0° to 180° ; for HI1 from 30° to 150° . For each degree, we go over all the images in a SegmSet stack and count the number of images (ImCount) that contain non-zero pixels along each degree. The result for each degree is stored in an associative array (AR) as an entry such as the current degree-ImCount pair. After that, from the AR array, we select all non-zero ImCount values and calculate their standard deviation, mean, and mean minus standard deviation. Then, again from the AR array, we select the pairs that have ImCount equal to or greater than mean minus standard deviation. From this new set of pairs, we select values of degrees and calculate their mean. The computed mean degree would be the estimated PA of the CME.

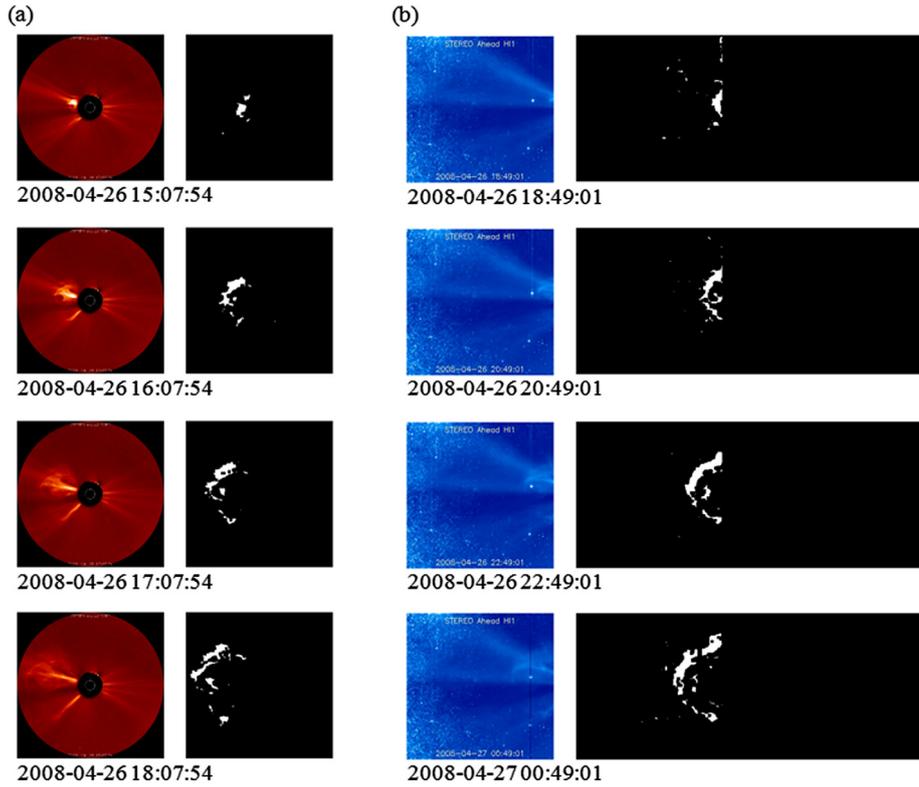


Fig. 2. An example of the final segmented images from the pre-processing module, for (a) COR2 and (b) HI1 images.

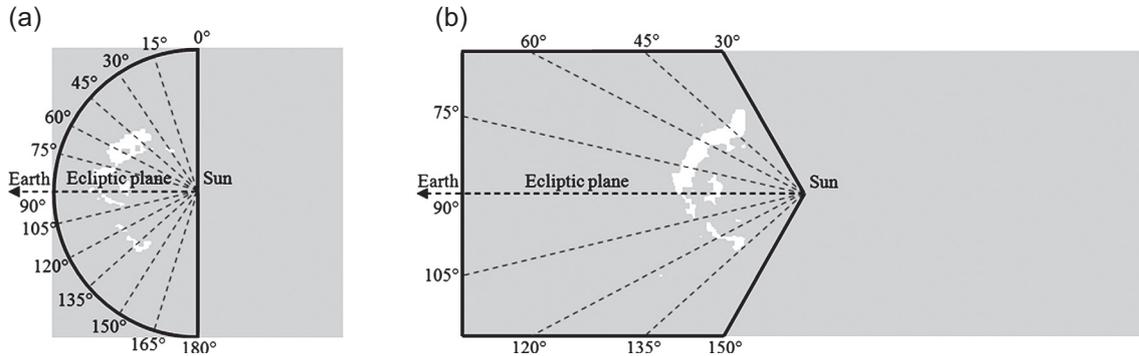


Fig. 3. The areas enclosed within the solid outlines from (a) COR2 image and (b) HI1 image are used to estimate the PA of the CME (event #2 in Table 1).

Once the degree of PA is determined, all SegmSet images are rotated by PA degrees in order to detect and track the CME front edge. After rotation, the original PA line becomes the horizontal line in the images, i.e., after the rotation, the pixels that fall onto the PA line become the middle horizontal row in the images. Next, the pixels along each column are examined within each image. If a column contains non-zero pixels (CME material), we store value 1 in the [column number, image number] entry in a 2D array of size $[V, N]$, where V is the half width of an image and N is the number of images in the stack. The generated 2D array is then analyzed with a connected-component labeling technique to segment isolated regions and estimate their sizes. The largest region will be identified as the material of the CME. Figure 4 shows the binary images of the resulting 2D arrays extracted from CME event #2 (see Table 1). In this example, the COR2 SegmSet image stack has 46 images and the HI1 SegmSet image stack has 71 images.

Note that each row in this 2D array represents a corresponding image in the SegmSet stack. For each row, we select the most left non-zero element and identify its index (i.e., column number). This index indicates the column representing the furthest (fastest) point in the front edge in the current image. This approach allows for tracking the CME front edge in the stack of images and estimating its projected height and speed using the most advanced point in the front edge.

The projected height of a CME represents the distance from the center of the Sun to the furthest position of the front edge in a particular image. The FOV in the COR2 and HI1 images is equal to 15 and 75 Rs (Eyles et al. 2009), which correspond to 128 and 256 columns in an image. The projected height, denoted as R_{Image} in Eq. (1), derived in the scale of the digital image, can be converted into the physical unit of measure – solar radii (Rs) using Eq. (1):

$$R_s = R_{\text{Image}} / k, \quad (1)$$

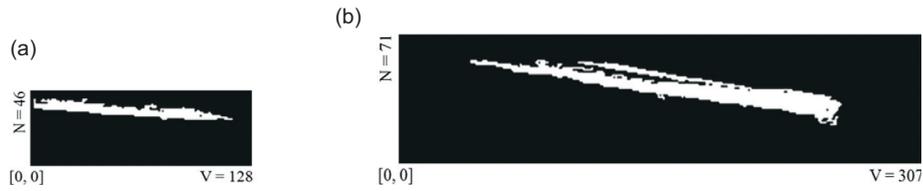


Fig. 4. The extracted binary CME patterns using (a) COR2 images and (b) HI1 images from CME event #2 in Table 1.

Table 1. The selected CME events used for the validation.

#	Date	COR2			HI1		
		First image	Last image	Number of images	First image	Last image	Number of images
1	2008-02-13	00:07:54		47	00:09:01		72
	2008-02-14		23:37:54			23:29:01	
2	2008-04-26	00:07:54		47	00:09:01		72
	2008-04-27		23:37:54			23:29:01	
3	2008-05-17	00:07:54		46	00:09:01		36
4	2008-06-02	00:07:54		48	00:09:01		72
	2008-06-03		23:37:54			23:29:01	
5	2008-07-07	00:07:54		43	21:29:01		40
	2008-07-08		23:37:54			23:29:01	
6	2008-08-30	00:07:54		96			36
	2008-08-31		23:37:54		00:09:01	23:29:01	
7	2008-12-12	00:07:54		48	00:09:01		72
	2008-12-13		23:37:54			23:29:01	
8	2008-12-27	00:07:54			00:09:01	29:29:01	35
9	2009-01-14	00:07:30		62	00:49:01		45
	2009-01-15		23:52:30			23:29:01	
10	2009-02-13	00:07:30		52	00:09:01		72
	2009-02-14		23:37:30			23:29:01	
11	2009-02-18	00:07:30		62	00:09:01		70
	2009-02-19		23:37:30			23:29:01	
12	2009-03-18	00:07:30		48	00:09:01		72
	2009-03-19		23:37:30			23:29:01	
13	2009-04-14	00:07:30		47	00:09:01		72
	2009-04-15		23:37:30			23:29:01	
14	2009-08-04	18:07:30		49			36
	2009-08-05		18:07:30		00:09:01	23:29:01	
15	2009-08-08	00:07:30		105			36
	2009-08-09		23:52:30		00:09:01	23:29:01	

where $k = 128/15 = 8.5$ for COR2 data and $k = 256/75 = 3.4$ for HI1 data.

The speed of a CME is calculated from the ratio of the distance (projected height) of the CME front edge traveled to the amount of time it took. Linear regression is used to estimate the CME speed which is the slope of the fitted line. Figures 5a and 5b show the “CME projected height versus time” plots (i.e., height-time profiles), extracted using COR2 and HI1 images of the same event.

3.4. Validation

To validate our method, we used a set of 15 CME events captured by the STEREO A COR2 and HI1 instruments in science quality. The chosen events propagate in the STEREO A COR2 and HI1 FOVs to the east of the Sun and were recorded in the period from January 1, 2008 to August 31, 2009. To obtain the reference propagation parameters, the COR2 set was manually inspected using a space weather analysis tool called StereoCAT (<http://ccmc.gsfc.nasa.gov/analysis/stereo/>; Mays et al. 2015). We used this tool only to

obtain the reference PA and speed of each CME event in the COR2 set used for validation. The parameters for the HI1 set were obtained from the Solar Stormwatch CME catalogue (<http://www.met.reading.ac.uk/~spate/solarstormwatch/>; Barnard et al. 2014). The reference CME parameters (PA and speed) were compared against the CME parameters computed using the proposed automatic method.

The image stacks used in our method were created from non-overlapping blocks of images and represent 24–48 h of observation. The COR2 image stacks contain 43–105 images. The HI1 stacks contain 35–72 images. The number of images in a stack is not critical in the proposed method, as long as there are no major losses in CME observation. Table 1 lists the chosen CME events and their corresponding timestamps for both the first and the last images of a stack and the total number of images within a stack.

4. Results

Figure 6a shows that the estimated PAs using the proposed automatic and the manual methods are in good agreement.

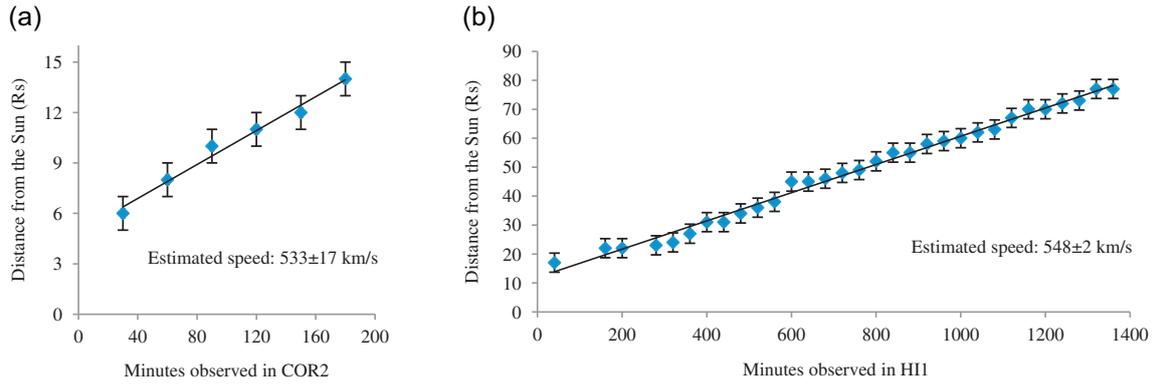


Fig. 5. The estimated height-time profile using the science quality (a) COR2 images and (b) HI1 images for CME event #2 in Table 1. The error bars were estimated as standard error of the mean.

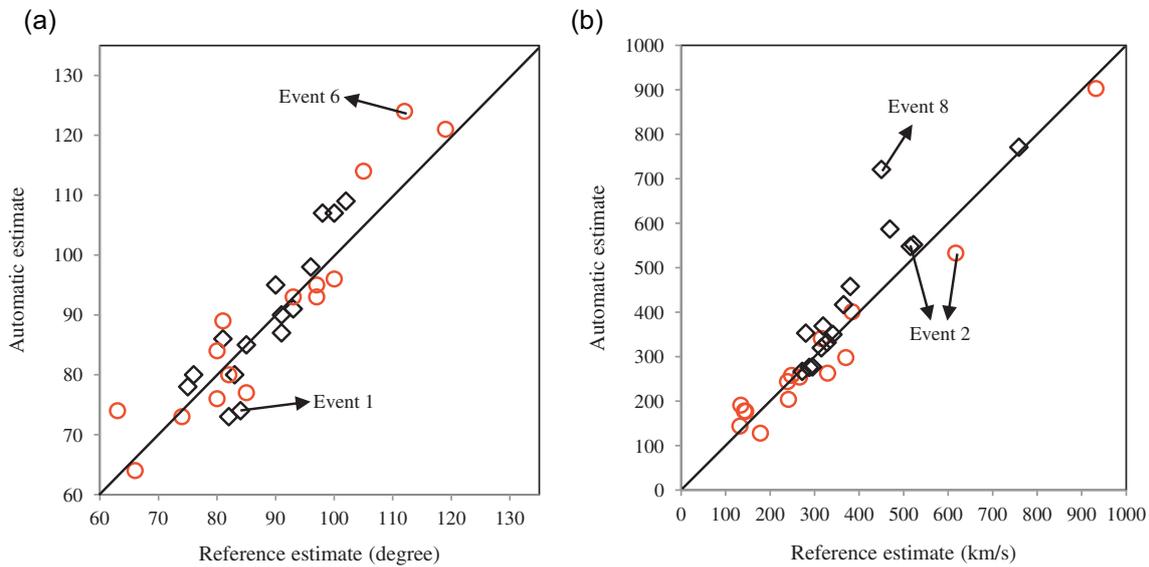


Fig. 6. The estimated (a) PA and (b) speed of the 15 CME events using the proposed automatic method versus the manual method. Red circle – COR2 data set, black diamond – HI1 data set.

It also shows that the PA estimates extracted from COR2 images do not exactly mirror the PA estimates made in HI1 images. We will discuss this outcome later in Section 5. Broad disagreement on the estimated PA parameters is found in event #6 for COR2 and event #1 for HI1 (see their corresponding timestamps in Table 1). Visual inspection of these two events reveals that CME analysis is a very subjective procedure. For event #1, the automatic method was able to detect some low intensity CME parts with eastward propagation. These parts which could be missed by human observers, but taken into account by our automatic method, result in different values of angular span of a CME and lead to the disagreement in the PA estimate. The disagreement in the estimated PA for event #6 is due to the slightly displaced core part of the CME. The expert may have identified this feature and estimated the PA more accurately.

Figure 6b shows that the CME speed estimated by the automatic method is also in accordance with the reference Stormwatch CME catalogue. Based on our 15 tested CME events, the estimated CME speeds are slower in COR2 FOV

Table 2. The computed absolute percentage errors (APEs) between the manual and the automatic methods.

Event #	COR2		HI1	
	PA (%)	Speed (%)	PA (%)	Speed (%)
1	3.1	15.4	11.9	1.2
2	9.4	13.6	6.2	6.4
3	4.0	3.1	5.5	1.6
4	4.1	28.1	1.1	2.6
5	1.7	20.1	7.0	5.7
6	10.7	25.4	6.8	25.2
7	1.4	4.0	1.5	14.2
8	17.5	4.4	3.6	60.2
9	5.0	19.5	2.2	26.1
10	1.2	8.3	4.4	1.6
11	5.0	4.5	5.3	15.7
12	9.9	22.1	4.0	4.2
13	2.4	42.5	11.0	1.8
14	2.1	2.1	9.2	20.5
15	3.0	9.1	2.1	6.1
MAPE	5.4	14.8	5.5	12.9

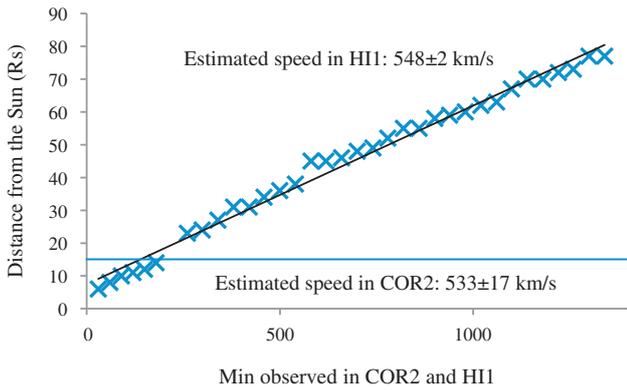


Fig. 7. Automatically estimated combined COR2 and HI1 height-time profile for CME event #2 in Table 1.

Table 3. The estimated CME speed using three different methods for CME Event 2008-04-26 (Event #2 in Table 1).

	Reference estimate	Wood and Howard estimate (km/s)	Automatic estimate (km/s)
COR2	617	676	533
HI1	515	543	548

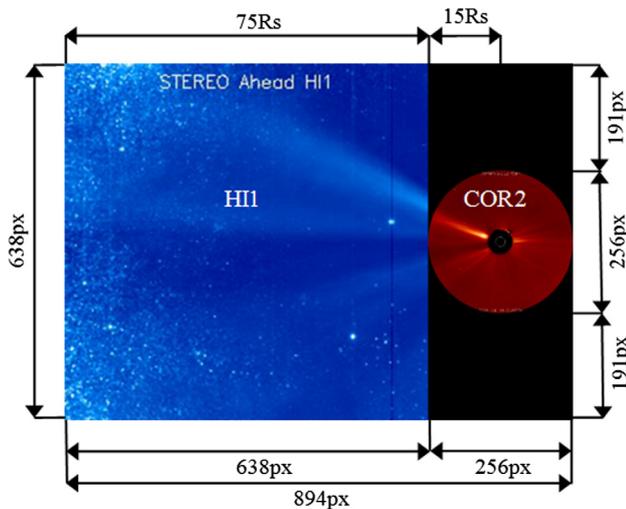


Fig. 8. An example of a combined COR2 and HI1 image.

compared to the speeds derived in HI1 FOV. It is also noted that the estimated CME speed in one HI1 case stands out from the average – event #8 (see corresponding timestamp in Table 1). The disagreement is due to the fact that the front edge of this CME diffuses as it propagates in HI1 FOV, becoming less distinguishable from the background. The reason for disagreement between the reference and automatic estimates is, most likely, because different CME features were used to track the front edge. We will discuss this further in the Section 5.

To evaluate the average level of disagreement between the manual and automatic methods, we computed the Mean absolute percentage errors (MAPEs) using the 15 CME events.

Table 2 shows the computed MAPEs for both PA and speed parameters. In general, the proposed automatic method showed a 94.6% agreement in the derivation of the PA parameter and 86.2% agreement in the estimation of the speed parameter.

Figure 7 shows the estimated height-time profiles using the proposed method combining COR2 and HI1 for event #2 (see Table 1). The proposed method was able to track the CME front edge in a much wider distance, starting from 6 Rs and up to 77 Rs. Such combined height-time profiles can reveal the speed of a CME and could be used to adjust or correct the earlier forecasts that use only COR2 images.

5. Discussion and conclusion

Space weather forecasts are critically important for many industries such as electric power grids, airlines, precision geo-location services, military, and satellite manufacturing and operations. Advance warning about the arrival of an earthward-directed CME is vital in many applications in order to provide time for proper mitigation. For example, electric power grid operators use geomagnetic storm detection and warning systems to ensure power grid stability and prevent large-scale blackouts (Graham et al. 2008). The initial forecasts of a CME propagation path can be issued based on COR observations. In turn, such forecasts are usually associated with large uncertainty. Therefore, HI data can be used to reduce the uncertainty and update early predictions. We have demonstrated that the proposed method can correctly track the CME front edge and estimate CME parameters in both COR2 and HI1 images (as shown in Fig. 5), and compute the height-time profile with the combined COR2 and HI1 images.

Among the selected CME events, event #2 in Table 1 and Figure 6b is a good example that demonstrates the benefit of tracking a CME beyond the COR2 FOV, i.e., continue the tracking in the HI1 FOV. Wood & Howard (2009) have performed a detailed kinematic analysis of this event. We compared their outcomes (Wood & Howard 2009) against the reference and our automatic estimates. As shown in Table 3, the estimated speeds using the COR2 images have larger disagreement, whereas HI1 estimates agree very well. This example is given solely to demonstrate that estimates issued using the COR2 data can have larger uncertainty and that HI1 estimates can potentially reduce such uncertainty. The results also demonstrate that manual CME analysis can be subjective and depends on the perception and experience of the human observer, as well as the tools used in the CME analysis.

We would like to note that after the detections have been made separately for COR2 and HI1, we combine the image stacks into a single stack and browse detection results in appropriately scaled COR2 and HI1 FOV images side-by-side. It is worth mentioning that there is no timestamp matching between COR2 and HI1 images. Therefore, one cannot combine COR2 and HI1 images simply by timestamp matching. To select images for a combined image stack, we select COR2 images starting from the first CME image in a stack and stopping at the image where the front edge was last detected. Next, we select HI1 images such that the timestamp is right after the last timestamp of the selected COR2 image.

The scale of FOVs in COR2 and HI1 images is obviously different. To combine images with different scales, we resize the HI1 images using a linear interpolation to the dimension

of 638×638 pixels and maintain the size of COR2 as 256×256 pixels. The dimension of images in a combined stack is equal to 894×638 pixels. Figure 8 demonstrates how the COR2 and HI1 images can be combined and placed side-by-side.

The results have demonstrated the efficiency of the proposed algorithm to track CME material in both COR2 and HI1 data sets. Good agreement in the PA assessment (Fig. 6a) indicates that the automatic method is robust in detecting CME central material and provides a reliable estimate about the main direction of propagation. In our algorithm, the evaluation of the speed parameter is fully dependent on the correctness of detection and tracking the CME front edge. The results in Figure 6b show that, in general, the estimates of the speed parameter are in close agreement. Note that HI1 speed estimates have better agreement between automatic and reference. This can be due to the fact that HI1 stacks contain more frames and that CMEs are visible in the HI1 FOV longer than in COR2. Consequently, CMEs can be tracked longer and further, leading to improved precision. In addition to that, this might be partly due to the sensitivity of our method while tracking CMEs in HI1 images.

CME events can have different structure, signal intensity, speed, and direction of propagation. The proposed automatic method has proven to be effective in processing most of them. As noted in Section 5, the estimates of the same CME event may differ between COR2 and HI1 data sets. This is because a CME can deflect from the initial path estimated low in the corona, i.e., COR2 FOV. It can also change its visual appearance (side edges become diffusive) or look different due to optical distortions caused by COR2 and HI1 detectors. Due to any of these reasons, CME features used to determine a CME propagation path may be different in COR2 and HI1 images, leading to different estimates. A CME feature captured in the FOV of one instrument may have a different appearance or may not be present in the FOV of another. Also, when comparing the estimates for the same CME event computed by the automatic and the manual methods, the results may vary depending on features and tools used to estimate the CME parameters.

In this study, we assumed that the FOV of a COR2 image is 2–15 Rs as described in Kaiser et al. (2008); and the FOV of an HI1 image is 15–90 Rs as described in Eyles et al. (2009). However, we found that there is some disagreement among previous works (e.g., Howard et al. 2008; Kaiser et al. 2008; Eyles et al. 2009) regarding the exact FOV of COR and HI images. This disagreement is possible since the estimation of the FOV in terms of Rs is not a trivial procedure. Due to orbital eccentricity, an instrument's FOV can change as the STEREO spacecraft moves on its orbit. The orbital distance of STEREO Ahead varies a little, while that of STEREO Behind varies more. This means that the instruments' FOVs for each spacecraft need to be estimated separately based on their positions in the orbit. Therefore, for precise estimation of a CME in STEREO SECCHI images, it is important to know the exact value of the FOV while an image is captured. In addition, the FOV information is required when combining images from different instruments, such as COR and HI.

The proposed method is not restricted to images with a resolution of 256×256 pixels. The median filter kernel and the size of structuring element (SE) used in morphological erosion and closing may need to be adjusted accordingly when a different image resolution is used. The values can be

determined experimentally based on the resolution and the quality of images that form the data set. In addition, one will have to estimate the solar radius in the scale of the new images and determine sizes of intermediate images as described in the description of the pre-processing module.

The proposed method is based on algorithms derived from data analysis with further application of image processing and pattern recognition techniques. The proposed method includes two individual modules: one unique approach to segment images in a stack (pre-processing module) and a novel approach to track a CME front edge and estimate propagation parameters in the stack of segmented images (tracking module). The pre-processing module allows for effective background removal and CME segmentation. The output of the pre-processing module is a set of segmented binary images, which can be fed into the tracking module to track a CME front edge and to estimate the propagation parameters. The method was validated using the selected CME events captured in the period from January 1, 2008 to August 31, 2009. The results demonstrate that the proposed pipeline is effective for CME front edge tracking and CME propagation parameter estimation. Integration of these innovative approaches into a technique for geometric triangulation can provide a necessary tool for automatic estimation of 3D CME properties. These improvements can greatly help in making quantitative CME analysis more accurate and readily available to assist in space weather forecasting.

Acknowledgements. The STEREO/SECCHI data used here are produced by an international consortium of the Naval Research Laboratory (USA), Lockheed Martin Solar and Astrophysics Laboratory (USA), NASA Goddard Space Flight Center (USA), Rutherford Appleton Laboratory (UK), University of Birmingham (UK), Max-Planck-Institut für Sonnensystemforschung (Germany), Centre Spatial de Liège (Belgium), Institut d'Optique Théorique et Appliquée (France), and Institut d'Astrophysique Spatiale (France). The authors would like to thank Dr. William Thompson at NASA for providing helpful details on the STEREO mission. The editor thanks Nandita Srivastava and an anonymous referee for their assistance in evaluating this paper.

References

- Allen, C.W. Relation between magnetic storms and solar activity. *Mon. Not. R. Astron. Soc.*, **104**, 13–21, 1944.
- Barnard, L., C. Scott, M. Owens, M. Lockwood, K. Tucker-Hood, et al. The solar stormwatch CME catalogue: results from the first space weather citizen science project. *Space Weather*, **12**, 657–674, 2014, DOI: [10.1002/2014SW001119](https://doi.org/10.1002/2014SW001119).
- Boursier, Y., P. Lamy, and A. Llebaria. Three-dimensional kinematics of coronal mass ejections from STEREO/SECCHI-COR2 observations in 2007–2008. *Sol. Phys.*, **256**, 131–147, 2009.
- Brueckner, G.E. The behavior of the outer solar corona ($3 R_{\odot}$ to $10 R_{\odot}$) during a large solar flare observed from OSO-7 in white light. In: G. Newkirk, Editor. *Coronal Disturbances*, IAU, D. Reidel Publishing Company, Dordrecht, Holland, p. 333, 1974.
- Burlaga, L.F., and R.P. Lepping. The causes of recurrent geomagnetic storms. *Planet. Space Sci.*, **25**, 1151–1160, 1977, DOI: [10.1016/0032-0633\(77\)90090-3](https://doi.org/10.1016/0032-0633(77)90090-3).
- Byrne, J.P., P.T. Gallagher, R.T.J. McAteer, and C.A. Young. The kinematics of coronal mass ejections using multiscale methods. *A&A*, **495**, 325–334, 2009, DOI: [10.1051/0004-6361/200809811](https://doi.org/10.1051/0004-6361/200809811).
- Byrne, J.P., S.A. Maloney, R.T.J. McAteer, J.M. Refojo, and P.T. Gallagher. Propagation of an Earth-directed coronal mass

- ejection in three dimensions. *Nat. Commun.*, **1**, 74, 2010, DOI: [10.1038/ncomms1077](https://doi.org/10.1038/ncomms1077).
- Byrne, J.P., H. Morgan, S.R. Habbal, and P.T. Gallagher. Automatic detection and tracking of coronal mass ejections. II. Multiscale filtering of coronagraph images. *Astrophys. J.*, **752**, 145, 2012, DOI: [10.1088/0004-637X/752/2/145](https://doi.org/10.1088/0004-637X/752/2/145).
- Byrne, J.P. Investigating the kinematics of coronal mass ejections with the automated CORIMP catalog. *J. Space Weather Space Clim.*, **5**, A19, 2015, DOI: [10.1051/swsc/2015020](https://doi.org/10.1051/swsc/2015020).
- Colaninno, R., and A. Vourlidas. First determination of the true mass of coronal mass ejections: a novel approach to using the two STEREO viewpoints. *Astrophys. J.*, **698**, 852–858, 2009.
- Crooker, N.U., and A.H. McAllister. Transients associated with recurrent storms. *J. Geophys. Res.*, **102** (A7), 14041–14047, 1997, DOI: [10.1029/96JA03431](https://doi.org/10.1029/96JA03431).
- de Koning, C.A., V.J. Pizzo, and D.A. Biesecker. Geometric localization of CMEs in 3D space using STEREO beacon data: first results. *Sol. Phys.*, **256**, 167–181, 2009.
- de Koning, C.A., and V.J. Pizzo. Polarimetric localization: a new tool for calculating the CME speed and direction of propagation in near-real time. *Space Weather*, **9**, S03001, 2011, DOI: [10.1029/2010SW000595](https://doi.org/10.1029/2010SW000595).
- Eyles, C.J., R.A. Harrison, C.J. Davis, N.R. Waltham, B.M. Shaughnessy, et al. The heliospheric imagers onboard the STEREO mission. *Sol. Phys.*, **254**, 387–445, 2009, DOI: [10.1007/s11207-008-9299-0](https://doi.org/10.1007/s11207-008-9299-0).
- Gallagher, P.A., C.A. Young, J.P. Byrne, and R.T.J. McAteer. Coronal mass ejection detection using wavelets, curvelets and ridgelets: applications for space weather monitoring. *Adv. Space Res.*, **47**, 2118–2126, 2011.
- Gosling, J.T., S.J. Bame, D.J. McComas, and J.L. Phillips. Coronal mass ejections and large geomagnetic storms. *Geophys. Res. Lett.*, **17**, 901–904, 1990.
- Graham, S.J., T.M. Fisher, V. Swisher, and C.A. Gruber. *Severe space weather events – understanding societal and economic impacts – a workshop report*, National Research Council, Division on Engineering and Physical Sciences, Space Studies Board, Committee on the Societal and Economic Impacts of Severe Space Weather Events, National Academies Press, Washington, DC, USA, ISBN: 0-309-12769-6, 2008.
- Harrison, R.A., C.J. Davis, C.J. Eyles, D. Bewsher, S.R. Crothers, et al. First imaging of coronal mass ejections in the heliosphere viewed from outside the Sun-Earth line. *Sol. Phys.*, **247**, 171–193, 2007, DOI: [10.1007/s11207-007-9083-6](https://doi.org/10.1007/s11207-007-9083-6).
- Howard, R.A., J.D. Moses, A. Vourlidas, J.S. Newmark, D.G. Socker, et al. Sun Earth Connection and Heliospheric Investigation (SECCHI). *Space Sci. Rev.*, **136**, 67–115, 2008, DOI: [10.1007/s11214-008-9341-4](https://doi.org/10.1007/s11214-008-9341-4).
- Howard, T.A., and S.J. Tappin. Three-dimensional reconstruction of two solar coronal mass ejections using the STEREO spacecraft. *Sol. Phys.*, **252**, 373–383, 2008.
- Kaiser, M.L., T.A. Kucera, J.M. Davila, O.C.St. Cyr, M. Guhathakurta, and E. Christian. The STEREO mission: an introduction. *Space Sci. Rev.*, **136**, 5–16, 2008, DOI: [10.1007/s11214-007-9277-0](https://doi.org/10.1007/s11214-007-9277-0).
- Kirnosov, V., L.-C. Chang, and A. Pulkkinen. Automatic CME front edge detection from STEREO white-light coronagraph images. *Space Weather*, **13** (8), 469–483, 2015, DOI: [10.1002/2015SW001190](https://doi.org/10.1002/2015SW001190).
- Liu, Y., J.A. Davies, J.G. Luhmann, A. Vourlidas, S.D. Bale, and R.P. Lin. Geometric triangulation of imaging observations to track coronal mass ejections continuously out to 1 AU. *Astrophys. J. Lett.*, **710**, L82–L87, 2010, DOI: [10.1088/2041-8205/710/1/L82](https://doi.org/10.1088/2041-8205/710/1/L82).
- Mays, M.L., A. Taktakishvili, A. Pulkkinen, P.J. MacNeice, L. Rastätter, et al. Ensemble modeling of CMEs using the WSA-ENLIL+Cone model. *Sol. Phys.*, **290** (6), 1775–1814, 2015, DOI: [10.1007/s11207-015-0692-1](https://doi.org/10.1007/s11207-015-0692-1).
- MacQueen, R.M., J.A. Eddy, J.T. Gosling, E. Hildner, R.H. Munro, G.A. Newkirk Jr., A.I. Poland, and C.L. Ross. The outer solar corona as observed from SKYLAB: preliminary results. *Astrophys. J.*, **187**, L85–L88, 1974, DOI: [10.1086/181402](https://doi.org/10.1086/181402).
- Mierla, M., J. Davila, W. Thompson, B. Inhester, N. Srivastava, M. Kramar, O.C. StCyr, G. Stenborg, and R.A. Howard. A quick method for estimating the propagation direction of coronal mass ejections using STEREO-COR1 images. *Sol. Phys.*, **252**, 385–396, 2008.
- Mierla, M., B. Inhester, A. Antunes, Y. Boursier, J.P. Byrne, et al. On the 3-D reconstruction of coronal mass ejections using coronagraph data. *Ann. Geophys.*, **28**, 203–215, 2010.
- Morgan, H., J.P. Byrne, and S.R. Habbal. Automatically detecting and tracking coronal mass ejections. I. Separation of dynamic and quiescent components in coronagraph images. *Astrophys. J.*, **752**, 144, 2012, DOI: [10.1088/0004-637X/752/2/144](https://doi.org/10.1088/0004-637X/752/2/144).
- Olmedo, O., J. Zhang, H. Wechsler, A. Poland, and K. Borne. Automatic detection and tracking of coronal mass ejections in coronagraph time series. *Sol. Phys.*, **248**, 485–499, 2008, DOI: [10.1007/s11207-007-9104-5](https://doi.org/10.1007/s11207-007-9104-5).
- Pérez-Suárez, D., P.A. Higgins, D.S. Bloomfield, R.T.J. McAteer, L.D. Krista, J.P. Byrne, and P.T. Gallagher. Automated solar feature detection for space weather applications. In: R. Qahwaji, R. Green, and E. Hines, Editors. *Applied signal and image processing: multidisciplinary advancements*, IGI Global, Hershey, PA, 207–225, 2011, DOI: [10.4018/978-1-60960-477-6.ch013](https://doi.org/10.4018/978-1-60960-477-6.ch013).
- Pulkkinen, A., T. Oates, and A. Taktakishvili. Automatic determination of the conic coronal mass ejection model parameters. *Sol. Phys.*, **261**, 115–126, 2009, DOI: [10.1007/s11207-009-9473-z](https://doi.org/10.1007/s11207-009-9473-z).
- Robbrecht, E. and D. Berghmans. Automated recognition of coronal mass ejections (CMEs) in near-real-time data. *A&A*, **425**, 1097–1106, 2004, DOI: [10.1051/0004-6361:20041302](https://doi.org/10.1051/0004-6361:20041302).
- Robbrecht, E., D. Berghmans, and R.A.M. Van Der Linden. Automated LASCO CME catalog for solar cycle 23: Are CMEs scale invariant. *Astrophys. J.*, **691**, 1222–1234, 2009, DOI: [10.1088/0004-637X/691/2/1222](https://doi.org/10.1088/0004-637X/691/2/1222).
- Srivastava, N., and P. Venkatakrishnan. Relationship between CME speed and geomagnetic storm intensity. *Geophys. Res. Lett.*, **29**, 1287, 2002, DOI: [10.1029/2001GL013597](https://doi.org/10.1029/2001GL013597).
- Srivastava, N., and P. Venkatakrishnan. Solar and interplanetary sources of major geomagnetic storms during 1996–2002. *J. Geophys. Res.*, **109**, A010103, 2004, DOI: [10.1029/2003JA010175](https://doi.org/10.1029/2003JA010175).
- Srivastava, N., B. Inhester, M. Mierla, and B. Podlipnik. 3D Reconstruction of the front edge of the 20 May 2007 partial halo CME. *Sol. Phys.*, **259**, 213–225, 2009, DOI: [10.1007/s11207-009-9423-9](https://doi.org/10.1007/s11207-009-9423-9).
- Tappin, S.J., T.A. Howard, M.M. Hampson, R.N. Thompson, and C.E. Burns. On the autonomous detection of coronal mass ejections in heliospheric imager data. *J. Geophys. Res.*, **117**, A05103, 2012, DOI: [10.1029/2011JA017439](https://doi.org/10.1029/2011JA017439).
- Tousey, R. The solar corona. *Space Res.*, **13**, 713–730, 1973.
- Wood, B.E., and R.A. Howard. An empirical reconstruction of the 2008 April 26 coronal mass ejection. *Astrophys. J.*, **702**, 901–910, 2009, DOI: [10.1088/0004-637X/702/2/901](https://doi.org/10.1088/0004-637X/702/2/901).
- Yashiro, S., G. Michalek, and N. Gopalswamy. A comparison of coronal mass ejections identified by manual and automatic methods. *Ann. Geophys.*, **26**, 3103–3112, 2008.
- Young, C.A., and P.T. Gallagher. Multiscale edge detection in the corona. *Sol. Phys.*, **248**, 457–469, 2008.