

Automatic Sunspots Detection on Full-Disk Solar Images using Mathematical Morphology

J.J. Curto · M. Blanca · E. Martínez

Received: 20 June 2007 / Accepted: 2 June 2008 / Published online: 16 July 2008
© Springer Science+Business Media B.V. 2008

Abstract Sunspots are solar features located in active regions of the Sun, whose number is an indicator of the Sun's magnetic activity. Therefore accurate detection and classification of sunspots are fundamental for the elaboration of solar activity indices such as the Wolf number. However, irregularities in the shape of the sunspots and their variable intensity and contrast with the surroundings, make their automated detection from digital images difficult. Here, we present a morphological tool that has allowed us to construct a simple and automatic procedure to treat digital photographs obtained from a solar telescope, and to extract the main features of sunspots. Comparing the solar indices computed with our algorithm against those obtained with the previous method exhibit an obvious improvement. A favorable comparison of the Wolf sunspot number time series obtained with our methodology and from other reference observatories is also presented. Finally, we compare our sunspot and group detection to that of other observatories.

Keywords Sunspots · Automatic detection · Mathematical morphology

1. Introduction

Sunspots are dark areas that grow and decay on the photosphere, the lowest layer of the Sun visible from the Earth. Sunspots are darker than their surrounding area because they are cooler than the average temperature of the solar surface (about 6000 K). The appearance and disappearance of sunspots is due to underlying changes in the magnetic fields that exist throughout the Sun. The presence of these strong magnetic fields reveals the presence of large amounts of energy that potentially can be released. Generally, sunspots are first observed as tiny dark spots named pores. Some of them develop into fully-fledged sunspot

J.J. Curto (✉) · M. Blanca
Observatori de l'Ebre, CSIC – Universitat Ramon Llull, 43520 Roquetes, Spain
e-mail: jjcurto@obsebre.es

E. Martínez
Grup de recerca en processament multimodal, Enginyeria i Arquitectura La Salle, Universitat Ramon Llull, C. Quatre Camins, 2, 08022 Barcelona, Spain

regions, evolving over time scales from hours to days. Occasionally, when a spot becomes darker and larger, portions of it may break away from the original spot; hence, the reason for a total sunspot count and a group classification.

Any sunspot region of significant size can be observed to form a double-ended group, and with an instrument known as a solar magnetograph it can be determined that spots at opposite ends of the group have opposite magnetic polarity. This is in agreement with the idea of magnetic field lines emerging through the photosphere. A sunspot group is thus defined as the collection of sunspots that belong to the same magnetic flux tube.

In solar-terrestrial physics, and especially in geophysics, solar indices are of vital importance to evaluate the potential impact of solar activity on the Earth and spacecraft (Lanzetta, 2001), as measured by indices of the geomagnetic field and/or ionospheric parameters (Hargreaves, 1992). One of the most widely used solar indices is the Wolf sunspot number that is based on the number of sunspots and sunspot groups. Sunspot number counts are taken at several solar observatories (including the Ebro Observatory) and, once collected in the Sunspot Index Data Center (SIDC, located in the Royal Observatory of Belgium since 1981), they are used to calculate the International Sunspot Number, R_i , known also as relative Wolf sunspot number in honor of its creator (Clette *et al.*, 2007). SIDC acts as data analysis service of the FAGS (Federation of Astronomical and Geophysical Data Analysis Services) broadcasting the daily, monthly, and yearly international sunspot numbers, with middle range predictions of vital importance for space weather services (Vanlommel *et al.*, 2004). Given the great interest of these indices, they were reconstructed to 300 years back in time, and various periodicities have been found in these data. The most prominent are the 11-year solar cycle and the 27-day Bartels solar rotation (Bartels, 1934).

From its foundation, Ebro Observatory has been devoted to the study of solar-terrestrial relationships. As a result, a long series of photographs of the Sun's photosphere and chromosphere, which started in the beginning of the past century, are stored in the observatory. One of the main tasks of its solar section was a daily routine involving manual detection of faculae and sunspots. The task of detection was undertaken by skilled observers visually inspecting the plates. Then, using templates and an abacus, the position and the area of each sunspot group were determined. Although with time the material of the plates changed, the instrument, a Mailhat optical telescope, and the method for sunspot detection and evaluation remained constant. In epochs of high activity this task occupied a skilled observer several hours each day.

The availability of digital cameras with CCD sensors and digital image processing techniques opened new possibilities for automatic detection. In 2001 a new telescope with a Zeiss lens equipped with a solar filter and with a calibrated CCD camera is was installed at Ebro Observatory to obtain a daily photograph of the solar photosphere (see an example in Figure 1). The pictures have a resolution of 1024×1024 pixels and an intensity range of 256 levels (8 bits per pixel). The global resolution is 2.06 seconds of arc per pixel. The automatic detection system described here has been in operation at Ebro Observatory since 2001.

In this paper, we will discuss the procedure that we developed for the automatic detection of sunspots. The particular characteristics of solar images are described in Section 2. A review of the basis of the morphological tools used here to detect the sunspots is provided in Section 3. Section 4 is devoted to a detailed explanation of how the solar limb and the sunspots are detected. There, an explanation of how sunspots are classified in groups can be found, too. A comparison of our manual and automated results and with those from other observatories can be found in Section 5. Finally, concluding remarks and further work are drawn in Section 6.

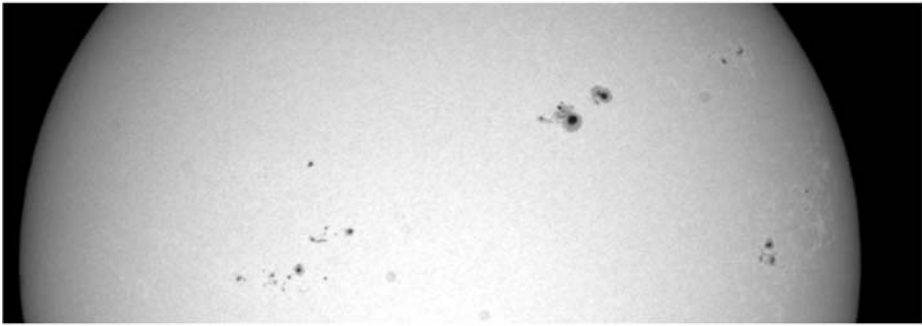


Figure 1 Fragment of the original image of the solar photosphere obtained at Ebro Observatory on 19 October 2001 at 9:21 hours.

2. Features of the Solar Images

There are particular difficulties in the analysis of white-light solar images. Some of these difficulties are permanent as the well-known limb darkening effect (Hubrecht, 1913; Neckel, 1994): intensity of the light coming from the center of the Sun is higher than in the extremes, thus contrast and background intensity of sunspots are not constant in the image but depend on their position on the solar disk. Standardization techniques were presented by Zharkova *et al.* (2003) but they are not required by the technique presented here. Also, the shape of the sunspots does not have a fixed pattern and varies during the sunspots' life. Other difficulties appear with the exposure time, which depends on the luminosity of the Sun and the sensitivity of the sensor. In each photograph it must be varied to optimize the contrast between the sunspots and the rest of the Sun. In general, one must avoid over- and underexposure. For those photographs taken at the Earth's surface, the luminosity depends on the day of the year (changes the Sun–Earth distance) and also on the hour of the day (changes the angle of incidence of the solar rays). But, also, it depends on irregular factors such as the transparency of the Earth's atmosphere. Other meteorological factors may produce dramatic results, too, such as the veiling effect of clouds, which can partially or, in an extreme case, totally mask the Sun. It can affect the whole image or some parts. Generally they produce irregular shadows on the image with structures that sometimes have the same size of the sunspots. Thus the possibilities of extracting sunspots contours on a cloudy day depend on the degree of transparency of those clouds. In general, clouds reduce the accuracy of sunspot detection and classification.

Another basic factor which influences the possibility of sunspot detection is the intrinsic quality of the image, degradable by many factors. The quality of the optical system and the substrate of the plate will determine the definition of the image and the appearance of details and the visibility of small structures. In our case, the superior resolution of the new optical system and CCD sensors posed a challenge because small sunspots not detectable by the old optical and the photographic development system became suddenly visible. Related to the clarity of the image is the focusing process which being manual is not always perfect. Unfocused images enlarge the sunspots size and reduce their intensity and contrast, provoking a dramatic decrease of the signal to noise ratio. Second, chromatic and spherical aberration on the optical lens produce distorted images with similar effects as unfocused images. On the other hand, dust and water on the optical elements or on the CCD sensors produce spots and shadows that could be interpreted as sunspots. Also damaged cells of the CCD produce

errors in the intensity values of their pixels. Therefore, a simple filtering process cannot be used to clear the background because we would remove information on sunspots. Neither can other selection criteria, such as size or shape, be used because one cannot forecast the characteristics of a particular sunspot.

Additionally to the position of the sunspots, we are interested in determining their relative area. There are several techniques for feature identification. Some give only rough estimates of the size and the shape of features, while more precise techniques are laborious to apply, rely on high resolution images, and often require visual inspection of images (Steinegger *et al.*, 1996). In general successful application of these methods requires very clean images (Preminger, Walton, and Chapman, 2001), which is not the case at Ebro.

The large amount of solar observations available stimulated the creation of automatic tools to treat them. Sunspots, filaments, active regions, flares, coronal mass ejections and magnetic neutral lines are the prominent features of the solar activity that can be analyzed (Zharkova *et al.*, 2005a). Many proposals have appeared, from Bayesian image segmentation (Turmon, Pap, and Mukhtar, 2002) to mean-field fast annealing (Bratsolis and Sigelle, 1998) and even neural networks (Fernandez Borda *et al.*, 2002).

Contrary to space-borne images, ground-based images have different levels of noise and distortion, so the method for the automatic detection must be robust and efficient. A method based solely on morphological tools provides the key aspects to achieve the required performance from the instrument at Ebro.

3. Morphological Tools

Mathematical morphology was born from an application of topology, geometry and algebra to investigate geometric structures (in binary or grayscale images) when applied to image processing and analysis (Serra, 1982). It provides powerful tools to extract the main features of a digitized image (González and Woods, 2002; Cordeiro, 2001). It allows the building of filters that ease region segmentation (Dougherty, 1992) and distinguishes meaningful shape information of geometric objects, measuring typical characteristics such as location, orientation, area or length of perimeter. Mathematical morphology has proved useful in solar images analysis of solar filaments (Qu *et al.*, 2005) and sunspot detection (Zharkov *et al.*, 2005).

This paper poses the complete set of procedures based on morphological tools to automatically detect sunspots on the solar plates of Ebro Observatory. The basic operations used in this work are:

3.1. Erosion and Dilation

Erosion $X \ominus B = \{h \mid B_h \subseteq X\}$. The result of an image, X , eroded by a structuring element (SE), B , consists of all the h points for which the translation of B by h fits inside X . The intensity of these points will be the minimum difference between the signal values and the translated-structuring-element values over the domain of the translated structuring element. *Erosion* has the effect of shrinking bright regions in the input image, so the eroded image is a subset of the input image (Figure 2).

Dilation $X \oplus B = \{h \mid B_h \cap X \neq \emptyset\}$. The dilated image results in an expansion of X . The intensity of these points will be the maximum difference between the signal values and the translated-structuring-element values over the domain of the translated structuring element.

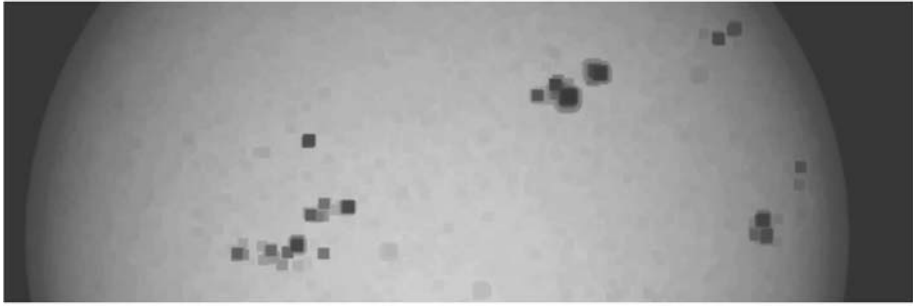


Figure 2 Eroded image using a square SE sized 11×11 once this operator is applied over the fragment of the original image obtained at Ebro Observatory on 19 October 2001.



Figure 3 Dilated image using a square SE sized 11×11 once this operator is applied over the fragment of the original image obtained at Ebro Observatory on 19 October 2001.

If we think of a disk structuring element, then dilation has an expanding effect filling in small (relative to the disk) dark holes in the images, whereas erosion has a shrinking effect eliminating small bright components of the image into its complement (Figure 3).

Whereas *erosion* results in a darker image and the dark objects become bigger, *dilation* obtains a brighter image, reducing, at the same time, the size of dark objects. A *dilation* applied after an *erosion* restores the objects to their original size.

3.2. Opening and Closing

Opening $(X \circ B) = \{(X \ominus B) \oplus B\} = \{B_h \mid B_h \subseteq X\}$. *Opening* of an input image, X , by a structuring element, B , is defined as an *erosion* concatenated with a *dilation*. It acts as a smoothing filter. The type of smoothing depends on the shape and the size of the structuring element. The *erosion* deletes all the high intensity pixels and turns the image darker. The next *dilation* results in a lighter image so the intensity of all the pixels is incremented. At the same time, an *opening* smoothes the corners of the objects (Figure 4). The overall effect is the deletion of small bright regions, smaller than the structuring element, while preserving the size of the rest of the regions.

Closing $(X \bullet B) = \{(X \oplus B) \ominus B\}$. *Closing* is the dual operation of *opening*. It is defined as a *dilation* followed by an *erosion*. *Dilation* erases the darker points of the image and



Figure 4 Resulting image after an opening transformation on the image in Figure 1.



Figure 5 Resulting image after a closing transformation on the image in Figure 1.

makes the image clearer. Then, the *erosion* turns the image darker. The effect of *closing* is the opposite of the *opening*: filters the image deleting small dark regions (Figure 5).

Both operations smooth the image but in two different ways: *opening* fills shape holes, whereas *closing* breaks wide lines and erases thin lines.

The most common applications are to erase salt noise (small interior holes) of an image by an *opening*, and the pepper noise (small surrounding objects) by a *closing*.

3.3. Top-Hat Transformation

Top-hat $top_hat(A, B) = X - (X \bullet B)$. *Top-hat* transformation consists in subtracting the original image X by the closing image. With *closing*, we erase small dark objects of the image and, in the next step, we subtract this image from its original. So, we obtain an image containing only the erased objects. This transformation leaves key markers such as sets of points with a sharp intensity gradient (Figure 6). As we use the *closing* operator, the transformation detects dark zones between bright zones so it achieves a valley detection.

4. Application

4.1. Solar Limb Detection

In order to determine the position of a sunspot in the solar disk, which should be expressed in spherical coordinates, first we must find where the solar center is.

This is not a trivial task because the solar disk is never completely centered in the plate, and, moreover, the apparent diameter of the Sun, seen from Earth, changes during the year.

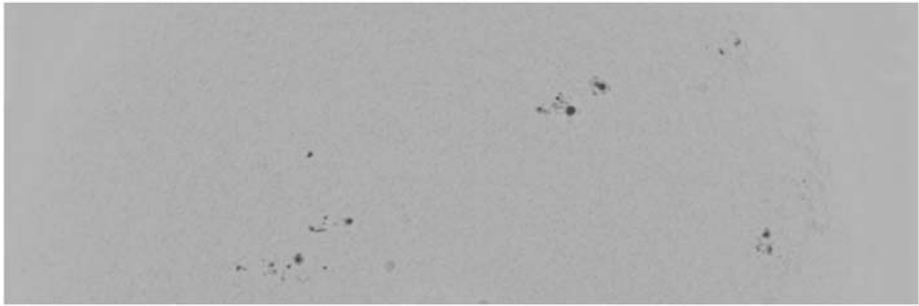


Figure 6 Fragment of the original image (solar photosphere obtained at Ebro Observatory on 19 October 2001) after a Top Hat transformation with $SE = 7 \times 7$.

Thus, an accurate measurement of the diameter of the solar sphere is required and this implies the correct detection of the solar limb.

However, the transition between the surface of the Sun and the image's background is not sharp but progressive, so the right threshold intensity level corresponding to the limb must be found. To achieve this, a special treatment to detect intense contrasts must be applied. While Zharkova *et al.* (2003) used "Canny edge" detection and later limb fitting, we used erosion, gradient transformation and threshold. To do this, first, an *erosion* must be applied to the original image by using a structuring element (SE) of 3×3 pixels size. This SE is the element that has its middle point centered, so that when applied, it does not displace the image. Obviously, in the eroded image, the solar disk is smaller and darker.

The following step is to apply an *eroded gradient transformation* by subtracting the original image from the eroded image. In the resulting image, the edges of the disk have the highest intensity.

To obtain a full contrasted edge, over the resulting image we retain only those pixels with higher intensity. So, from the inverted image, only those pixels under a threshold level are kept (Figure 7). This threshold level, of course, depends on the darkness of the image (exposure time). Statistically, the suitable value of threshold level usually is about 4% of the intensity range.

Due to air turbulence during windy days, the shape of the Sun seen with the telescope is not a perfect circumference but has minor irregularities (Figure 8). Therefore the Sun's center is obtained as the mean of the maximum and minimum points in horizontal and vertical directions.

The pole position is obtained by astronomical computation determining the tilt of the rotation axis of the Sun relative to the Earth axis for the moment the photograph was taken.

4.2. Sunspot Detection

Once the disk limit is determined, we can analyze its interior. The surface of the Sun presents a complex distribution of structures with different intensity levels and no regular patterns. Usually each dark structure is considered as an element or sunspot. Their identification involves what is called an image segmentation process. For an automatic system, it is desirable that no supervision, no previous parameters and no heuristics be needed to perform the final segmentation. There exists three basic ways for image segmentation: boundary based, region based and thresholding approaches. For the first class, boundaries between regions are

Figure 7 Inverted image after the eroded gradient transformation is applied over the original image taken at Ebro Observatory on 19 October 2001, now ready for solar limb detection.

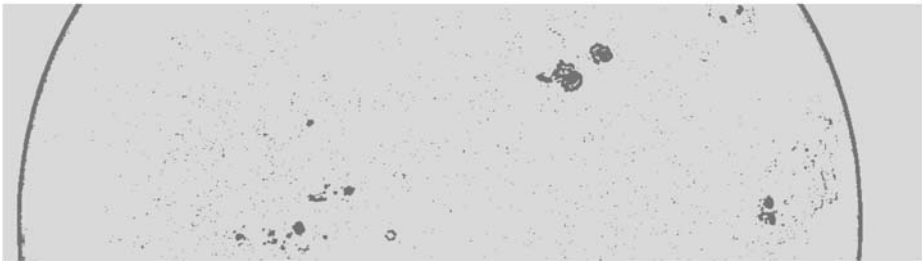
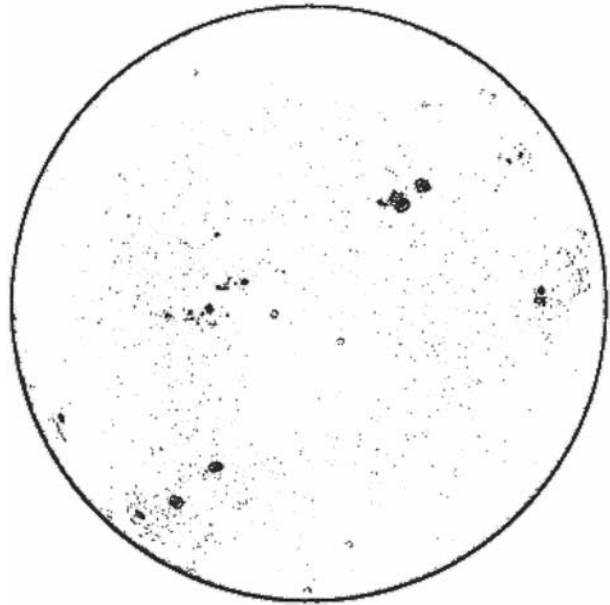


Figure 8 Detail of Figure 7 showing the irregular thickness of the solar limb.

found by looking for discontinuities in gray levels. Contours are determined by using operators sensitive to meaningful discontinuities in intensity level as Gauss smoothing and Sobel gradient in Zharkov *et al.* (2005a; 2005b). However, they need post-processing operations because of edge linking which leads to an unsatisfactory determination when structures are complex. Watershed and object markers belong to the second class. In the watershed method each catchment basin corresponds to a minimum of gradient. This has the tendency to oversegmentation. So an additional and complex selection of markers through homotopy modifications must be performed first. The third class, thresholding, is the simplest and quickest procedure. We chose that because it is very appropriate due to its fast computing and easy application and it will be optimal for future real-time treatment. However, a global threshold system is not sufficient because the uneven brightness of the solar disk (limb darkening effect) and the possible presence of clouds shadows produce histograms that cannot be partitioned. Normalization of the image in the pre-processing step is often necessary (Zharkova *et al.*, 2003). There are background regions of the solar disk which are darker than some sunspots in other regions. To overcome this, we process the original image with a Top Hat

operator acting as a local contrast detector which is very suitable for enhancing structures that have a detectable contrast against their local neighborhood.

Determining sunspots as individual entities is equivalent to determining which pixels belong to each sunspot. To do this, we have to find which pixels are active points. So, the critical factor to be found is the right threshold intensity level to distinguish active points from the background. This threshold varies from one image to another depending on the exposure time and should be computed in each case.

Our system has a detection procedure using an iterative threshold like Zharkov *et al.* (2005). As stated in Zharkov *et al.* (2005), the sunspot candidates are obtained first applying edge detection and then, an iterative threshold on the gradient image which has been previously normalized counting the number of the connected regions. These regions are filled in using morphological closing and watershed operators and finally a blob algorithm defines the regions. However, in our case, the sunspot candidates are obtained by applying a procedure involving only two iterative loops on the original image; one increasing the size of the structuring element of the closing operation and in the other increasing the level of intensity while the population growing of sunspots pixels, SP, is being controlled. Simple threshold methods will only provide a useless collection of points some of them belonging to sunspots but many others forming a disseminate of gray noisy points. So we must find aggregates of pixels of a certain size.

The way to detect the totality of the true sunspot pixels is using an iterative method increasing each time the scope scale and counting the detected pixels. Again a new difficulty arises: there is a large variety of sizes in sunspots. So a priori, nobody knows how big the biggest sunspot will be in a particular image. Thus, the iterations must go on until the population of the detected pixels stabilize. At that point, we have all the pixels belonging to detectable sunspots.

In practice this is achieved in the following way:

4.2.1.

First, a *closing transformation* over the original image is applied. Again a symmetric structuring element is chosen in order to keep centered the transformed image. With this process, sunspots smaller than the SE are detected. To make sure that no sunspot is lost, we start with an SE smaller than the smallest sunspot to be detected, 7×7 pixels in our case, increasing progressively its size in the following iterations. All this results in a (closed) image where the dark elements being smaller than the SE, disappear and the sunspots suffer a reduction in their extension.

4.2.2.

Then, a *Top Hat operator* – valley detector type – is applied by subtracting the closed image from the original image. This process gives the map of the points removed during the *closing transformation*. This includes few dark spots, fragments of the sunspots and a sea of gray noisy points coming from the background (Figure 9).

To separate them, an iterative process using threshold detection and counting is applied. One must start from a threshold (TL) of very low intensity level so at the beginning only the darkest points belonging to sunspots are detected. Then the threshold level is increased. At each step, the detected population must be counted. This process repeats until the population of the detected pixels increase dramatically indicating that the background level is reached. At that point, we are sure the previous threshold level is the correct value that will allow us to separate the true sunspot pixels from the background noise.

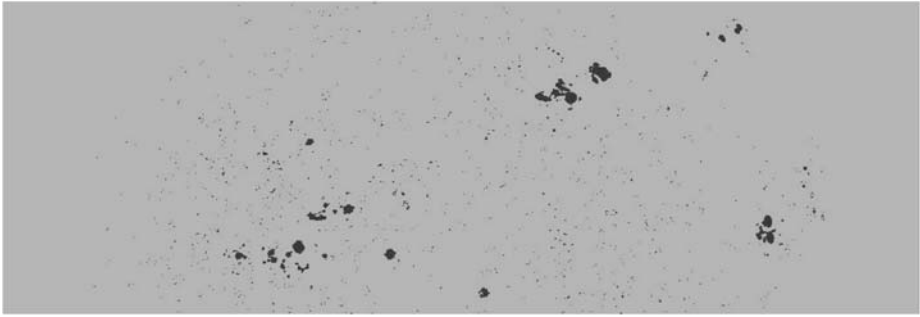


Figure 9 Image after the Top Hat operator with $SE = 25 \times 25$ was applied over the original image taken at Ebro Observatory on 19 October 2001.



Figure 10 Resulting image at the end of the process of treatment of the original image (Figure 1) with sunspots already isolated.

Then the whole process must be repeated again, starting from the *closing transformation*, but now, increasing the SE size. In general, the larger SE, the larger the number of true sunspot pixels is detected. But this iterative process finishes when all the points belonging to sunspots are detected and the pixel population remains stable – no more new pixels are detected – (Figure 10).

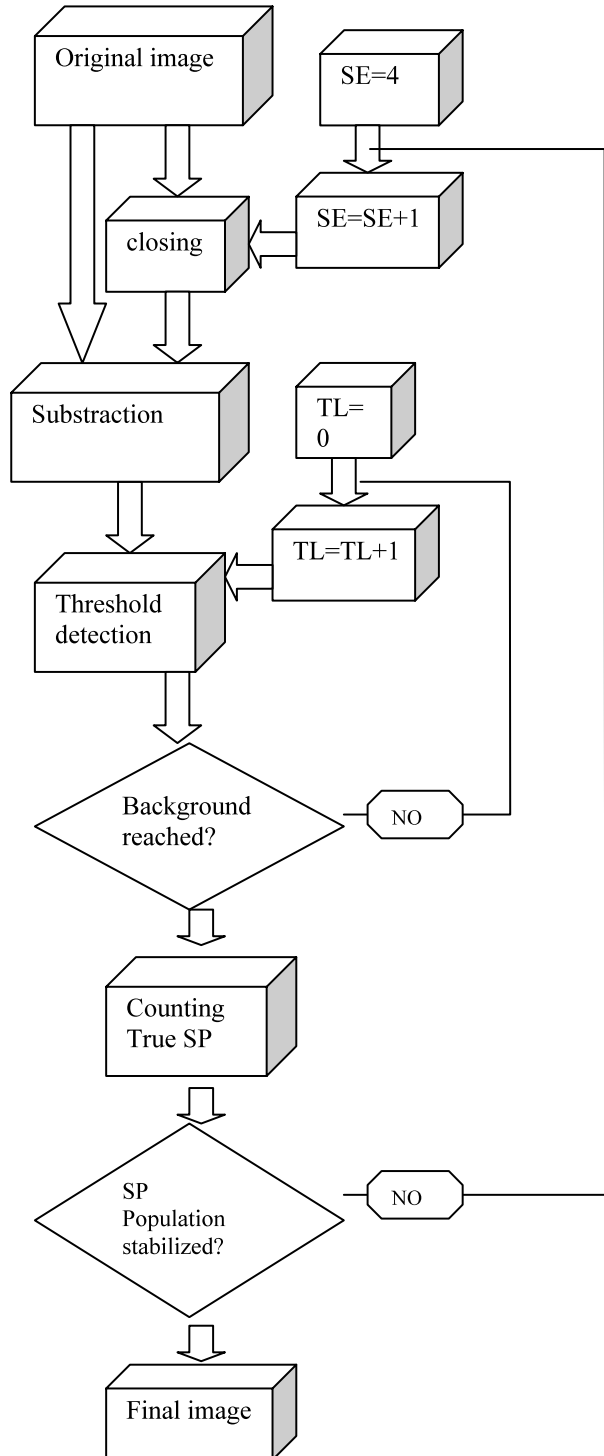
Obviously, the bigger the SE size is, the longer the time spent by the computer calculating all the values will be. This part of the detection is the most time consuming in the computational process. Figure 11 summarizes the whole process.

4.2.3.

Finally, each pixel belonging to the subset of true sunspot pixels (active points) must be assigned to its own sunspot. To do this we used a region growing procedure that groups pixels or subregions into larger regions based on the predefined criteria of neighborhood. The basic approach is to start with a set of seed points and from these grow regions by appending to each seed those neighboring pixels that have the property of being a true sunspot pixel.

Sweeping the image in one direction from a corner to the other, each active pixel is examined if it has some neighbors. In the affirmative case, the point is assigned to the sunspot of its neighbors. In the negative case, it is assigned as the beginning of a new sunspot as a seed. Due to the irregular shape of the sunspots, neighboring has a variable scale for border pixels. So the process must be iterated until the whole population is assigned to their own sunspot.

Figure 11 Chart diagram of the sunspot detection procedure.



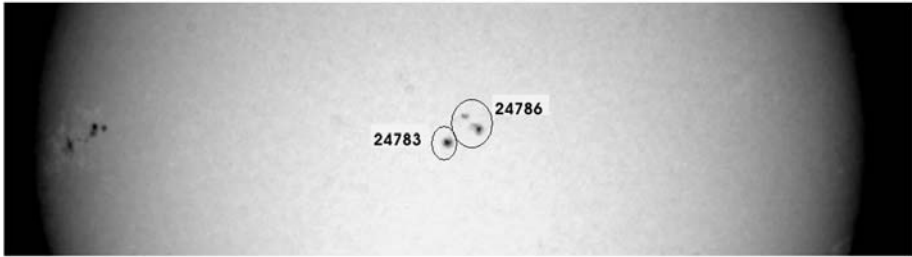


Figure 12 Fragment of the original image of the solar photosphere obtained at Ebro Observatory on 3 February 2003 at 9:05 hours. The groups #24786 and #24783 are so close together that they were initially taken as a single group by the automatic classification.

4.2.4. Pattern Recognition

Finally, the main characteristics of each sunspot are extracted. Thus position in the image, mean intensity and area are computed and stored in an associated file. Our method, trying to be coherent with the former manual method of detection, considers both, umbra and penumbra, as a unity in the sunspot and does not separate them.

4.3. Sunspot Groups

As explained above, sunspots are formed in groups that share physical properties such as belonging to the same magnetic flux loops. Both the number of individual sunspots and the number of sunspot groups are needed to determine the Wolf solar activity index. Thus, the final task of our algorithm determines the groups as individual entities. This is equivalent to joining the sunspots belonging to the same solar group. The procedure is similar to that used to assign each pixel to its sunspot: here sunspots play the role of the pixels and groups, that of the sunspots.

From previous statistics, we know that the sunspots neighboring scale is 6 heliographic degrees (sunspots which are separated less than 6° form part of the same solar group). However, to compute this, we have to convert the image coordinates into heliographic coordinates (Green, 1985). To do this conversion, several astronomical parameters as the coordinates of the apparent center of the Sun and the tilt of the rotation axis of the Sun relative to the Earth axis – which depend on the time and date and the place on the Earth where the photograph was taken – and the value of the apparent solar diameter on the plate – computed above – are needed. Operating in the same way as for sunspot detection, solar groups are determined.

Although a crosscheck with the synchronized solar magnetogram gives very reliable criteria for verification (Zharkov *et al.*, 2005), neighborhood is a simple and good enough criterion to classify groups and only in very few cases, when groups are placed in a very close space, mostly in the maximum of the solar cycle, the operator who supervises the whole process at Ebro must redefine the classification (Figure 12).

At the end of the whole process, we have a comprehensive list with each sunspot and the group it belongs to (Table 1). The program has a graphic interface which allows the operator to control the whole process (Figure 13). Results are daily send to the SIDC center and annually published in a solar bulletin available on the web: http://www.obsebre.es/php/meteosol/boletin_solar.php.

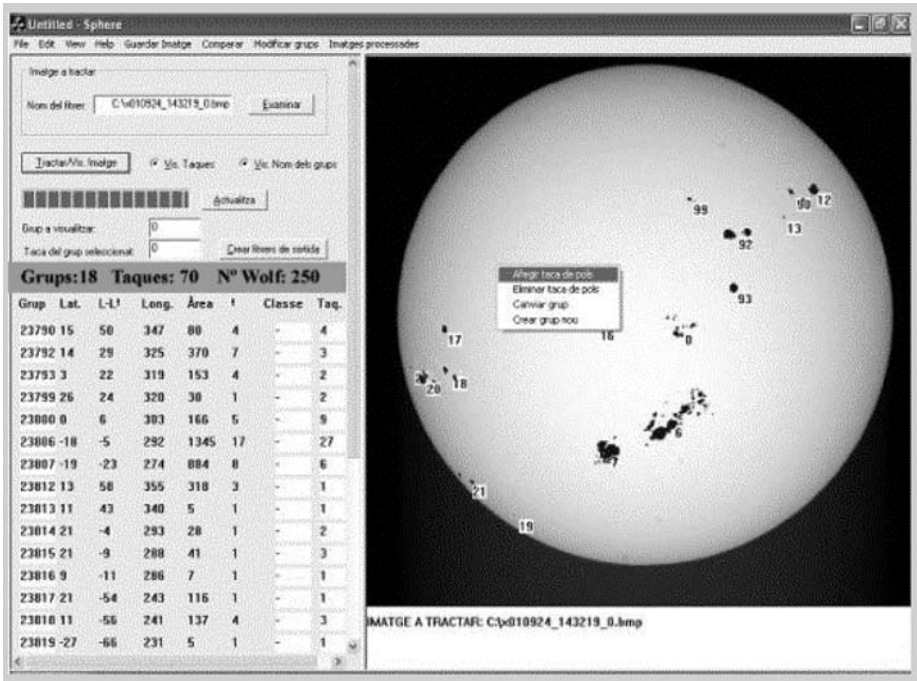


Figure 13 Visual interface with the outputs of the treatment: List of groups with their characteristics (left) and their position on the solar disk (right).

Table 1 Example of the output list of groups for 1 January 2003.

Year: 2003, Month: 1, Day: 1, Hour: 10, Minute: 10, Day_of_the_year: 1, Rotation: 1998
 Solar limit: X1: 25, X2: 981, Y1: 986, Y2: 28
 Quality of the photograph 1
 NUMBER OF GROUPS: 2 NUMBER OF SUNSPOTS: 4 WOLF NUMBER: 24
 BASIC INFORMATION OF SOLAR GROUPS

GROUP	Lat.	L-L°	Long.	Area	Degrees	T. Act.	Sunspots	Pixels	T. Ant	X	Y
24739	20	26	323	12	1	J	1	14	J	697	318
24741	-9	-4	292	12	4	C	3	17	C	470	555

DETAILED INFORMATION OF SOLAR GROUPS

GROUP	Sunspot	Pixels	Area	Degrees	X	Y	Intensity
24739	1	14	12	1	697	318	201
24741	1	1	1	1	496	552	234
24741	2	11	8	1	466	555	215
24741	3	5	4	1	473	556	220

5. A Comparison of Wolf Numbers

The level of solar activity is usually measured with indices. One of the most popular is that known as the relative Wolf number, R , which was defined as $R = k(f + 10G)$,

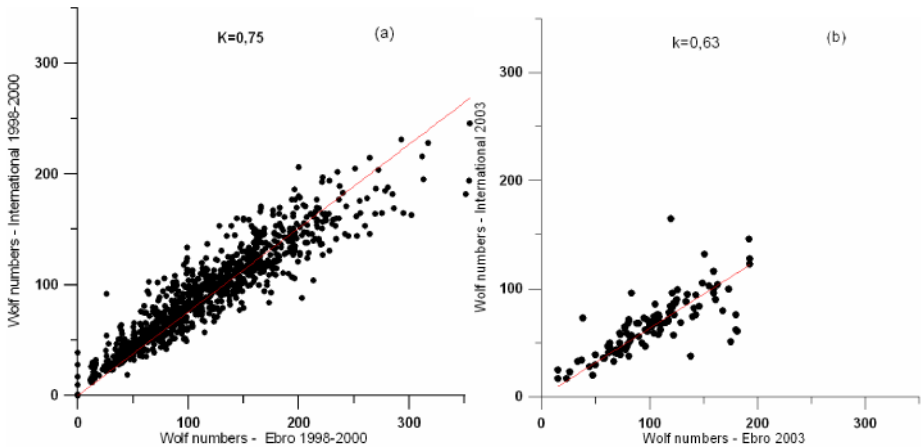


Figure 14 International Wolf number against Ebro Wolf number for the period 1998–2000 (a) and for the period 2003 (b). The parameter k was computed as the slope of the linear fit and its values confirm that more sunspots are detected in the actual configuration.

where f is the total number of sunspots, G the number of groups, and k corresponds to an observatory factor which depends on actual conditions of each observatory (the seeing conditions due to the clarity of the air, and the instrument (telescope) used). The original intention of introducing the correction factor k in about 1882 by Wolf's successors at the Zürich Observatory was to convert the actual daily measurements to the scale originated by Wolf (Waldmeier, 1961). Thus, nowadays, the International Wolf number is elaborated in the SIDC so that it is coherent with the long series of data extending to the last centuries. A system in Locano reproducing the original Zürich old detection method is used as reference, and the values given by each solar observatory collaborating in the production of the international index is weighted with a k parameter to normalize its results. The final daily international index is then computed as a weighted average of all of them.

To check the improvement achieved with our method of sunspot and group detection, we computed the k parameter by comparing the daily Ebro Wolf number with the International Wolf number. We did it first for the period 1998–2000 when images came from the old telescope and the manual method of detection was used, and then for the year 2003 when the new telescope and the automatic detection system was fully operative. In principle, the closer its results fit to the Zürich observation, the higher the k is (obviously with the limit of 1).

In our case, Ebro took the value of $k = 0.75$ for the period 1998–2000, which is very close to the International Wolf Number (Figure 14). For the year 2003, its lower value, $k = 0.63$, makes clear that more sunspots are detected in the actual configuration. These results confirm that the older the equipment and the method are, the closer its results are to the Zürich values.

Let us now check the performance of Ebro with regards to other observatories, which had already modernized of their observation system. We choose three of the most prestigious solar observatories: Mt. Wilson (34° N, 118° W) at 1742 m high in the west of USA, Sacramento Peak (33° N, 106° W) at 2821 m high in New Mexico (USA), and Catania (38° N, 16° E) at 3340 m high, close to Etna volcano, in Sicily (south of Italy). All three were installed at sites with very good air conditions and have powerful telescopes.

Again we chose 2003 as the test year and we took those days when simultaneous data were available for all four observatories (111 out of 365 days) checking the position of the sunspots and groups in each case. Although there is a large amount of information in their online databases, not all the information that we needed (the number of sunspots, groups and Wolf numbers) was explicit, so we had to determine it from their auxiliary files and images. In a primary inspection of the original photographs, their high quality stands out.

Figure 15 presents graphically the evolution of the number of sunspots, groups and solar indices for the four observatories. Sacramento Peak and Mt. Wilson observe more sunspots. The latter is the most regular observatory presenting fewer maxima and minima. Catania has more fluctuation with several maxima and minima. Ebro, although regular, usually observes fewer sunspots than the rest which, after direct comparison of the original images, can be attributed to the inferior quality of its images. However, the results change when one considers the number of groups. Ebro has the highest rate followed closely by Mt. Wilson. Again Ebro is the most regular observatory. Globally, Mt. Wilson has the highest Wolf numbers and prominent maxima. Catania is more variable and has the highest absolute value while Ebro has the lower values but its minima are not too far from those of the other observatories.

Proceeding in the same way as we did for Ebro, for the period 2003 we found that the k parameter for Mt. Wilson is 0.53, for Sacramento Peak is 0.55 and for Catania is 0.59. All three values are lower than that of Ebro. Mt. Wilson has the lowest value, probably due to the excellent quality of its images. Sacramento Peak has a k value similar to that of Mt. Wilson but with a wider dispersion which could be attributed to irregularities in its database. Catania has the highest value and closer to Ebro (belonging to the same climatic area).

To check the performance of the sunspot detection in more detail, we proceeded as Zharkov, Zharkova, and Ipson (2005) where the authors carefully verified their detection results with detections from other observatories (NOAA and SIDC datasets) providing the accuracy of the applied technique.

We use data from the Solar Feature Catalogue (SFC) developed in the European Grid of Solar Observatories (Zharkova *et al.*, 2005b) and available at <http://solar.inf.brad.ac.uk>.

They are very suitable because processed data not manually but with an automated technique for sunspot recognition as we do.

Although in the SFC catalog there are large number of sunspots, most of them correspond to very small features. Ebro with images having poorer quality detects few small features and the comparison could not be done. So we opted to compare bigger features. Results for sunspots with an area equal or greater than one degree are shown in Table 2. The total amount of detections is comparable. Additionally to the absolute number of detections, it is worthy to compute the false acceptance rate (FAR) (where we detect a feature and they do not) and the false rejection rate (FRR) (where they detect a feature and we do not). Both get close to 30%. Although most of them could be explained by external factors, limiting the quality of the image (clouds, dust, not taken synchronously), or internal factors, counting method, multi-umbra detection (Zharkov *et al.*, 2005), this indicates that there is still space to improve the techniques.

An additional verification is to compare sunspot areas from our method with those from a standardized catalog as that of NOAA/USAF as in Zharkov *et al.* (2005) (Figure 16). Here the coefficient of correlation is 70%.

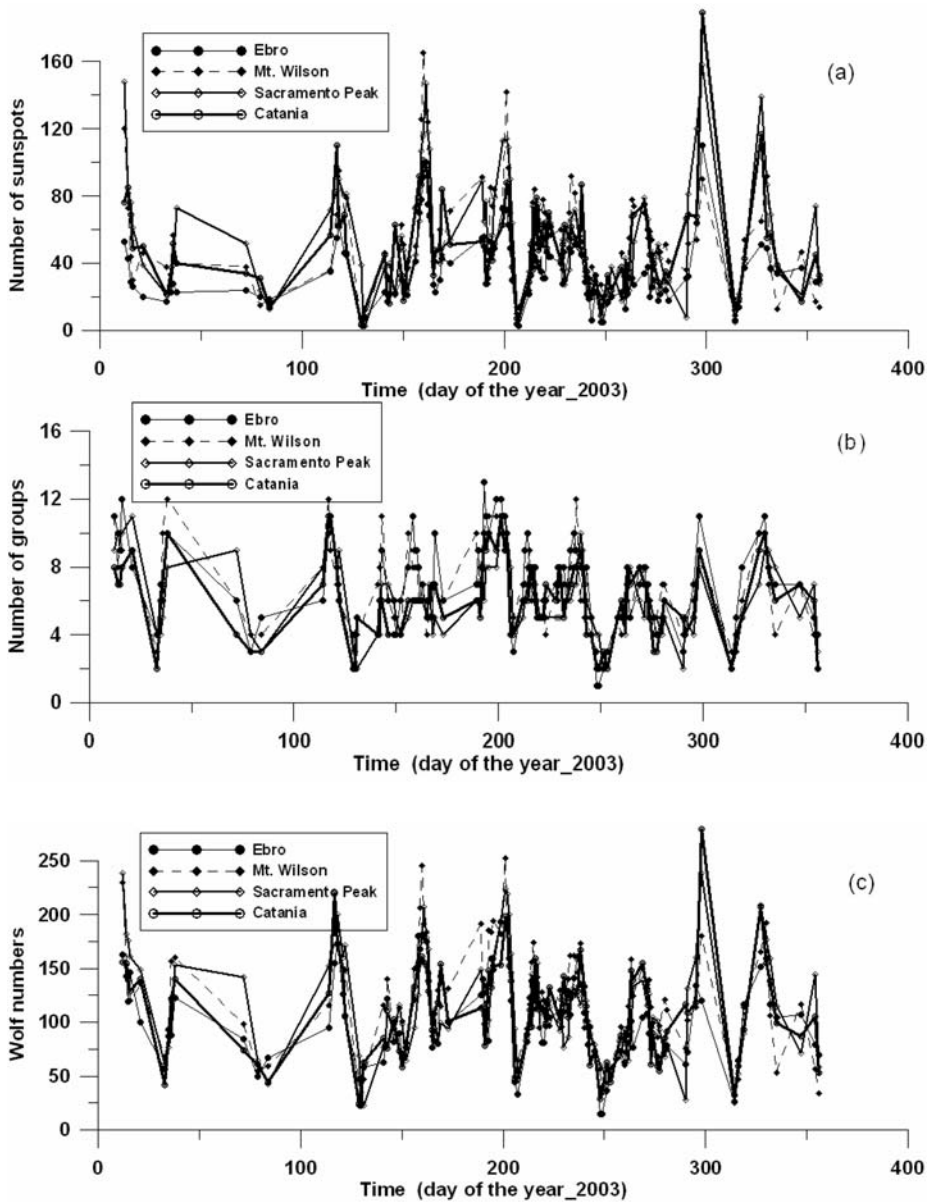


Figure 15 Daily variation of (a) sunspots, (b) groups and (c) Wolf numbers of the observatories Ebro, Mt. Wilson, Sacramento Peak and Catania for the year 2003.

Finally, we compared group classifications by Ebro and Mt. Wilson (Table 3). Although our system is very simple (only uses neighboring criteria and no information on magnetic field polarities) we have 80% of exact coincidence.

Table 2 A comparison of sunspots with area equal or greater than one degree detected by Ebro and SFC.

<i>Date</i>	Number of sunspots (<i>Ebro</i>)	Number of sunspots (<i>SFC</i>)	FAR	FRR
04-May-2004	9	9	3	3
05-May-2004	9	3	6	0
06-May-2004	9	4	5	0
07-May-2004	4	7	1	4
08-May-2004	3	7	0	4
09-May-2004	4	6	2	4
10-May-2004	9	7	4	2
12-May-2004	21	4	17	0
13-May-2004	10	16	1	7
14-May-2004	20	18	5	3
15-May-2004	16	26	2	12
16-May-2004	20	22	3	5
17-May-2004	23	33	5	15
18-May-2004	17	24	2	9
19-May-2004	17	17	3	3
20-May-2004	15	17	4	6
21-May-2004	15	13	2	0
24-May-2004	28	13	15	0
27-May-2004	27	12	15	0
28-May-2004	12	6	6	0
29-May-2004	11	18	3	10
30-May-2004	12	22	0	10
31-May-2004	13	10	0	3
TOTAL	314	304	103	100

6. Conclusions

Automatic detection of sunspots in photographic plates is possible with procedures based on mathematical morphological tools. These procedures have been implemented with a robust and simple technique which has been used routinely at Ebro Observatory for automatic detection since 2001. The resulting dataset is very consistent with those from the most sophisticated solar observatories. Position and area of individual sunspots and sunspot groups are obtained, and the Wolf numbers are computed by the same program. In this way, Ebro Observatory, as a solar observatory, compiles essential information which is delivered regularly to the scientific community through the WDC in Brussels.

The development of our automatic algorithms, a part of saving manpower, is able to provide an immediate evaluation of solar parameters that are very valuable for real-time space weather predictions. New software developments are now in course aiming to the automatic determination of other geometric parameters of the sunspots (additionally to the position and area), which will help improve the classification of the sunspots groups. Other improvements as umbra/penumbra separation as in Zharkov *et al.* (2005) and Zharkov, Zharkova, and Ipson (2005) will be considered.

Figure 16 A comparison of temporal variation of daily sunspot areas extracted in 2005 from (a) NOAA/USAF archive, US, and from (b) Ebro catalog with the technique we present here.

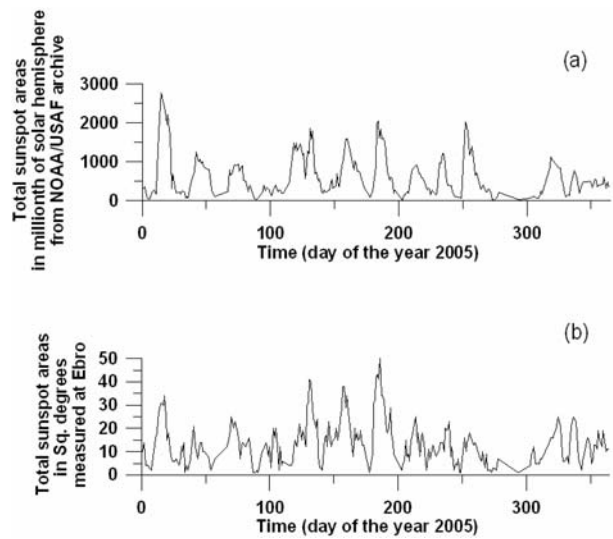


Table 3 A comparison of group classification by Ebro and Mt. Wilson.

Date	Number of groups (Ebro)	Number of groups (Mt. Wilson)	FAR	FRR
11-Sep-2001	9	8	3	2
12-Sep-2001	11	10	1	0
13-Sep-2001	11	10	1	0
14-Sep-2001	10	10	1	1
15-Sep-2001	14	14	5	5
17-Sep-2001	9	12	0	3
18-Sep-2001	14	14	3	3
20-Sep-2001	18	16	4	2
21-Sep-2001	12	12	1	1
24-Sep-2001	13	16	1	4
25-Sep-2001	14	17	1	4
26-Sep-2001	14	18	1	5
TOTAL	158	157	22	30

Acknowledgements The authors thank the referee for his/her fruitful comments and J. Casanovas, M. Vázquez and R. Casas for their advice in the instrumental and astronomical fields. Also we thank D. Esmel for his fruitful work in Wolf number comparison. We are grateful to SFC and NOAA/USAF for providing us open access to their catalogues. Finally we would like to thank Dr. H. McCreddie for her language revision.

References

- Bartels, J.: 1934, Twenty-seven day recurrences in terrestrial magnetic and solar activity, 1923-33. *Terr. Magn. Atmos. Electr.* **39**, 201 – 202.
- Bratsolis, E., Sigelle, M.: 1998, Solar image segmentation by use of mean field fast annealing. *Astron. Astrophys.* **131**, 371 – 375.

- Clette, F., Berghmans, D., Vanlommel, P., Van der Linden, R.A.M., Koeckelenbergh, A., Wauters, L.: 2007, From the Wolf number to the International Sunspot index: 25 years of SIDC. *Adv. Space Res.* **40**, 919–928. doi:[10.1016/j.asr.2006.12.045](https://doi.org/10.1016/j.asr.2006.12.045).
- Cordeiro, M.: 2001, *Algorithmic Patterns for Morphological Image Processing*, Prentice-Hall, Englewood Cliffs.
- Dougherty, E.R.: 1992, *An Introduction to Morphological Image Processing*, SPIE Optical Engineering Press.
- Fernandez Borda, R.A., Mininni, P.D., Mandini, C.H., Gómez, D., Bauer, O., Rovira, M.: 2002, Automatic solar flare detection using neural network techniques. *Solar Phys.* **206**, 347–357.
- González, R.C., Woods, R.E.: 2002, *Digital Image Processing*, Prentice-Hall, Englewood Cliffs.
- Green, R.M.: 1985, *Spherical Astronomy*, Cambridge University Press, Cambridge.
- Hargreaves, J.K.: 1992, *The Solar-Terrestrial Environment*, University Press.
- Hubrecht, J.B.: 1913, Some problems of astronomy: X. The darkening of the Sun's limb. *The Observatory* **36**, 398–401.
- Lanzerotti, L.J.: 2001, Space weather effects on communications. In: Daglis, I.A. (ed.) *Space Storms and Space Weather Hazards*, Kluwer, Dordrecht, 313–334.
- Neckel, H.: 1994, Solar limb darkening 1986–1990, Hamburger Sternwarte, Germany.
- Preminger, D.G., Walton, S.R., Chapman, G.A.: 2001, *Solar Phys.* **202**, 53–62.
- Qu, M., Shih, F.Y., Jing, J., Wang, H.M.: 2005, Automatic solar filament detection using image processing techniques. *Solar Phys.* **228**, 119–135.
- Serra, J.: 1982, *Image Analysis and Mathematical Morphology*, Academic, New York.
- Steiniger, M., Vazquez, M., Bonet, J.A., Brandt, P.N.: 1996, On the energy balance of solar active regions. *Astrophys. J.* **461**, 478–498.
- Turmon, M., Pap, J.M., Mukhtar, S.: 2002, Statistical pattern recognition for labelling solar active regions: Application to SOHO/MDI imagery. *Astrophys. J.* **568**, 396–407.
- Vanlommel, P., Cugnon, P., Van der Linden, R.A.M., Berghmans, D., Clette, F.: 2004, The SIDC: World data centre for the sunspot index. *Solar Phys.* **224**, 113–120.
- Waldmeier, M.: 1961, *The Sunspot-Activity in the Years 1610–1960*, Schultheiss Publisher, Zürich.
- Zharkov, S., Zharkova, V., Ipson, S.: 2005, Statistical properties of sunspots in 1996–2004: I. Detection, North–South asymmetry and area distribution. *Solar Phys.* **228**, 377–397.
- Zharkov, S., Zharkova, V., Ipson, S., Benkhalil, A.: 2005, Technique for automatic recognition of sunspots on full-disk solar images. *EURASIP J. Appl. Signal Process.* **15**, 2573–2584.
- Zharkova, V., Ipson, S., Zharkov, S., Benkhalil, A., Abouadarham, J., Bentley, R.: 2003, A full-disk image standardisation of the synoptic solar observations at the Meudon observatory. *Solar Phys.* **214**, 89–105.
- Zharkova, V., Ipson, S., Benkhalil, A., Zharkov, S.: 2005a, Feature recognition in solar images. *Artif. Intell. Rev.* **23**, 209–266.
- Zharkova, V., Abouadarham, J., Zharkov, S., Ipson, S., Benkhalil, A., Fuller, N.: 2005b, Solar feature catalogue in EGSO. *Solar Phys.* **228**, 361–375.