Automated Recognition of Sunspots on the SOHO/MDI White Light Solar Images

S.Zharkov, V.Zharkova, S.Ipson and A. Benkhalil

[1]Department of Cybernetics, University of Bradford, BD7 1DP, UK
S.Zharkov@brad.ac.uk

Abstract. A new technique is presented for automatic identification of sunspots on the full disk solar images allowing robust detection of sunspots on images obtained from space and ground observations, which may be distorted by weather conditions and instrumental artefacts. The technique applies image cleaning procedures for elimination of limb darkening, intensity noise and non-circular image shape. Sobel edge-detection is applied to find sunspot candidates. Morphological operations are then used to filter out noise and define a local neighbourhood background via thresholding, with threshold levels defined as a function of the quiet sun intensity and local statistical properties. The technique was tested on one year (2002) of full disk SOHO/MDI white light (WL) images. The detection results are in very good agreement with the Meudon manual synoptic maps as well as with the Locarno Observatory Sunspot manual drawings. The detection results from WL observations are cross-referenced with the SOHO/MDI magnetogram data for verification purposes.

1 Introduction

With a substantial increase in the size of solar image data archives, the automated detection and verification of various features of interest is becoming increasingly important for, among other applications, data mining and the reliable forecast of the solar activity and space weather. However, this raises the accuracy and reliability required of the detection techniques used for automated recognition which have to be significantly improved in comparison with the existing manual ones in order to create a fully automated Solar Feature Catalogue. Manual sunspot catalogues in various formats are produced in various locations all over the world such as the Meudon Observatory (France), the Locarno Solar Observatory (Switzerland), the Mount Wilson Observatory (US) and many others. Sunspot studies play an essential part in the modelling of the total solar irradiance and in determining variations of sunspot properties with latitude and/or the solar cycle phase.

The compilation of the Zurich relative sunspot numbers, or since 1981 the Sunspot Index Data (SIDC), is one of the most commonly used measures of solar activity (Hoyt & Schatten[1] and Temmer, Veronig, and Hanslmeier[2]). As integral component of Solar Active Regions, sunspot behaviour is also used in the study of Active Region evolution and in the forecast of solar flare activity (see Steinegger et al [4]).

A sunspot is a dark cooler part of the Sun's surface photosphere and is characterised by a strong magnetic field, formed below the photosphere, which extends out into the solar atmosphere and corona. Sunspots are best observed in the visible spectrum also known as 'white light'. Sunspots can also be observed in CaII K1 absorption line images. Sunspots generally consist of the two parts: a darker, roughly circular central disk called the umbra, and a lighter outer area called the penumbra.

2 The existing Methods of Detection.

From the point of view of digital imaging the sunspots (as represented on white light, CaII k1, CaII k3 and H-alpha line images) can be generally characterised by the following two properties: they are considerably darker than the surrounding photosphere and they have well-defined borders, i.e. the intensity change occurs over reasonably short distance from photospheric value to the spot value. The existing techniques for sunspot detection can be divided into the three basic classes. A number of existing methods [1-9], called thresholding methods but also including region-growing techniques, rely on sunspots lower intensity variations. There are also methods, called border methods developed by Győri [14], Pettauer and Brandt [8], making use of the intensity gradient of the sunspot image. In addition, substantial work has been carried out on Active Region detection and labelling by means of Bayesian Pattern Recognition methods by Turmon, Pap and Mukhtar [13] that also incorporates sunspot detection (penumbra only).

The above mentioned methods, with the exception of Bayesian ones, can be described as semi-automatic techniques since they require a user participation of some kind (for verification, threshold choice, choice of input image for instance). At the same time, all these methods are data specific in the sense that they were developed for specific data sets and, hence, make a number of assumptions about the data photometric properties, image resolution and presence of image artefacts.

3 Data Description and Preprocessing

It can be observed that, in general, the results of sunspot detection on digital images depend on the following: seeing conditions (for ground-based observatories); wavelength (more spots in white light than Ca II k1); instrumental accuracy and artefacts; image resolution (smaller sunspots/pores may not be seen in smaller resolution images).

MDI images taken aboard the SOHO satellite with continuous (4 synoptic images per day) coverage since May 1995 are also characterised by extensively descriptive and precise header information and absence of seeing (weather) condition effects. This makes this dataset very attractive for our purposes, notwithstanding its relatively low resolution of 2 arc seconds per pixel. To improve the data coverage and provide catalogue continuity we have extended our methods to sunspot detection on the Meudon Observatory daily Ca II k1 line images. Both data sets cover the time period spanning April 1, 2002 to April, 30, 2002. The SOHO/MDI data for the entire year 2002 have been also processed.

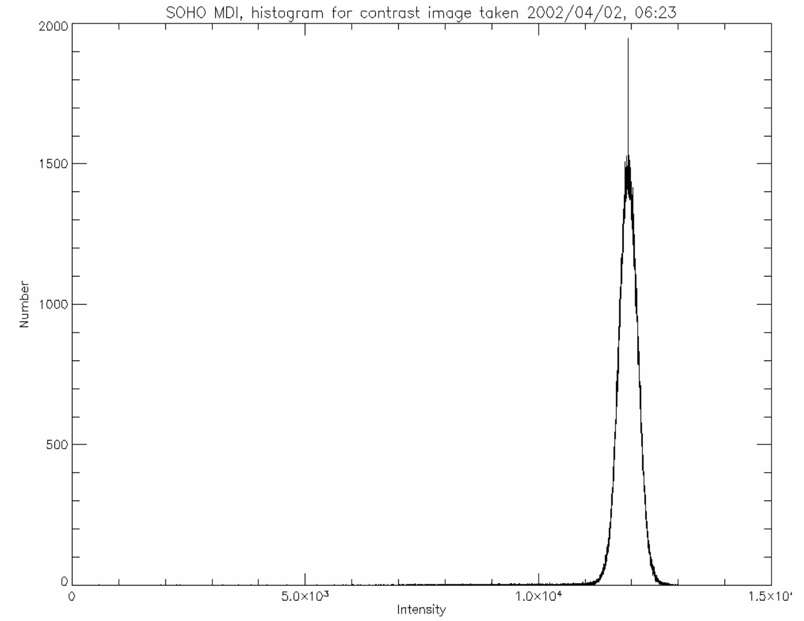

Fig. 1. Determination of the Quiet Sun Intensity as the highest pixel count from the histogram of a “flattened” photospheric image

The SOHO/MDI instrument provides almost continuous observations of the Sun in the white light continuum in the vicinity of the Ni I 676.7 nm line with resolution comparable to the ground-based telescopes. For the SOHO/MDI intensity spectrograms the pre-processing stage was as follows. For the pixels outside of the solar disk, as specified by the FITS file header information, the intensity values were set to zero, thus taking image dynamic intensity range into the non-negative integers set. Parameters such as disk centre, resolution, date of observation, disk radius were extracted from the FITS file header and solar disk intensity was then flattened by compensating for the limb darkening curve, see Zharkova et al [11] for details, thus producing a flat white light image, on which sunspot detection is run. For a flat image, the Quiet Sun Intensity value, I_{QSun}, is defined as the most populated non-zero intensity (i.e. as the intensity with the highest pixel count, see Figure 1).

While our detection method relies mainly on intensity properties of the white-light image, SOHO/MDI magnetogram data were used for verification purposes. In most cases, it is possible to locate a magnetogram virtually simultaneous (observations made within 30 seconds of each other) with the continuum observations. Since time difference between these observations is always under 2 hours, we synchronise both observations to a 'continuum point of view' by rotating the magnetogram data to the 'continuum' time of observation using SolarSoft procedures.

4. The Sunspot Detection Method for the SOHO/MDI images

Following the discussion in the Introduction, in order to extract as much information as possible from the chosen data we combine the thresholding and border methods. In order to avoid dependency on the choice of a global intensity threshold, edge detection is used instead. By examining the features with well defined borders we are then able to apply thresholding methods locally. Good results are achieved for noisy ground-based images using either the morphological gradient or Sobel operators.

The detection code is applied to a SOHO MDI continuum "flattened" full disk image, Δ (Figure 2, top left image), with determined quiet Sun intensity, I_{QSun} (Figure 1). image size, solar disk center pixel coordinates, disk radius, date of observation, and resolution (in arc seconds per pixel). A SOHO/MDI magnetogram, M, is synchronised to the continuum image via a (temporal) rotation and a spatial displacement to obtain the same point of view as the continuum.

The detection algorithm is described below in the pseudo-code.

1. Apply histogram equalization to increase a contrast (if required, optional)
2. Apply Gaussian smoothing with sliding window 5x5 followed by a Sobel operator to a copy of Δ ;
3. Using the initial threshold value, T_0, threshold the edge map and apply the median filter to the result. Count the number of connected components (Feature Candidates, Figure 2, top right). If it is too large, increase T_0 and repeat step 3 from the beginning.
4. Remove the edge corresponding to the limb from Candidate Map and fill the possible gaps in the feature outlines using IDL's morphological closure and watershed operators to define a candidate feature, F_i, as a set of pixels representing a connected component on the resulting binary image, B_Δ (Figure 2, second row, left).
5. Create an empty Sunspot Candidate Map – a byte mask for the original input image indicating detection results with pixels belonging to umbra marked as 2, penumbra as 1.
6. For every F_i extract a cropped image containing F_i and define the

   i.if F_i =< 5 pixels assign the thresholds:

   for penumbra $\mathbf{T_s = 0.91}$ I_{QSun} ; for umbra $\mathbf{T_u = 0.6}$ I_{QSun}

   ii.if F_i > 5 pixels assign the thresholds:

   for penumbra: $\mathbf{T_s = 0.93}$ I_{QSun} ;

for umbra: $\mathbf{T_u} = \mathbf{max}\ \{\ \mathbf{0.55}\ I_{QSun}\ ;\ (\langle \mathbf{P_i} \rangle - \Delta\ \mathbf{P_i}\ )\}$,

where $\langle \mathbf{P_i} \rangle$ is mean intensity value and $\Delta\ \mathbf{P_i}$ a standard deviation for F_i

7. Threshold a cropped image at this value to define the candidate umbral and penumbral pixels and insert the results back into B_Δ (Figure 2, second row, right).

8. To verify the detection results, cross check B_Δ with M, as follows:

for every connected component S_i of B_Δ extract

$$B_{\max}(S_i) = \max(M(p) \mid p \in M)$$

$$B_{\min}(S_i) = \min(M(p) \mid p \in M)$$

9. if $\max(abs(B_{\max}(S_i)), abs(B_{\min}(S_i))) < 100$ then disregard S_i as noise.

10. For each S_i extract and store the following parameters: gravity center coordinates (Carrington and projective), area, diameter, umbra size, number of umbras detected, maximum-minimum-mean photometric intensity (as related to flattened image), maximum-minimum magnetic flux, total magnetic flux and total umbral flux.

4. The Results and Conclusion

The results of sunspot detection on the image taken on the 2nd April are presented in Figure 2, (second row, right) with a closer view of a particular sunspot group presented in the third and fourth rows of Figure 2. The technique has been used to process one year of SOHO/MDI data with results stored and shared over the Internet. The ability to verify the detection results by checking the magnetic data helped to increase the detection accuracy (both in terms of False Acceptance Rates and False Rejection Rates) on the SOHO/MDI images and to detect the presence of the smaller pores, which are normally detectable on the images with a higher spatial resolution.

Since daily Wolf Numbers are primary indicators of sunspot activity extracted manually, in order to be able to statistically compare our detection results with manual catalogues we have to develop our method further to classify detected sunspots into sunspot groups, thus generating Wolf numbers.

A manual comparison of the sunspots detected on the SOHO/MDI images using the above technique with the sunspot drawings for June-July, 2002, produced in Locarno Solar Observatory, revealed an excellent (about 98%) agreement between the data sources and detected features with minor differences naturally present due to the differences in spatial resolution, observation time and seeing conditions for the ground-based observatory.

The accuracy of the technique developed for the sunspot detection on the Meudon Observatory Ca II k1 images, with more noisy backgrounds, was tested by a comparison with the manual synoptic maps for sunspots generated at the Meudon Observatory (Zharkov at al., 2003). Comparison was based on the data for April ’02 with the false rejection and acceptance rates, FRRs and FARs, calculated for daily observations that did not exceed in average 0.8 and 1.3, respectively. The discrepancies occurred because the Meudon images have more noise from the atmospheric conditions as well as because of the usage of extra knowledge in their manual technique, i.e. placing the sunspots on the map when they are not actually seen on the image. The SOHO/MDI data was concluded to be a preferable data source for the same period of observations since it provides higher accuracy and a much better coverage while the features visible in Ca II k1 are a subset of those seen in 'white light' .

In summary, the new technique for automated sunspot detection on full disk white light SOHO/MDI images achieved the detection of sunspots with an excellent accuracy, the extraction of sunspot locations, umbral/penumbral areas, diameters, irradiance and their correlation with magnetic field variations.

This technique is used for building the Sunspot Feature Catalogue for the European Grid of Solar Observations (EGSO) funded by the European Commission within the IST Fifth Framework, grant IST-2001-32409.

References

1. Hoyt, D.V., & Schatten, K.H. Solar Physics, 1998b; 181: 491-512
2. Temmer, M., Veronig, A., Hanslmeier, A. Hemispheric Astronomy and Astrophysics. 2002; 390, 707-715
3. Steinegger, M., Vazquez, M., Bonet, J.A. and Brandt, P.N. Astrophysical Journal. 1996; 461: 478.
4. Chapman, G. A. and Groisman, G. Solar Physics, 1984, 91: 45
5. Steinegger, M., Brandt, P. N., Pap, J., and Schmidt, W. Astrophysics and Space Science. 1990; 170: 127-133
6. Brandt, P. N., Schmidt, W., and Steinegger, M., Solar Physics 1990; 129: 191
7. Beck, J.G. Chapman, G.A. A study of the contrast of sunspots from photometric images. 1993, Solar Physics, 146: 49
8. Pettauer, T., Brandt, P.N. Solar Physics, 1997, 175: 197
9. Steinegger, M., Bonet J., A., Vazquez, M.. Solar Physics, 1997; 171: 303
10. Preminger, D. G., Walton, S.R., Chapman, G.A. Solar Physics, 2001, 202: 53
11. Zharkova, V.V., Ipson, S. S., Zharkov and 3 other authors Solar Physics, 2003; 214: 89-105
12. Győri, L., Solar Physics, 1998; 180:109-130
13. Turmon, M., Pap, J. M., Mukhtar S., ApJ, 2002; 568:396-407
14. Zharkov, S., Zharkova V.V., Ipson, S.S., Benkhalil A., Proc. of the AISB'03 Symposium on Biologically-inspired Machine Vision, Theory and Application, University of Wales, Aberystwyth, 7th-11th April 2003, pp. 74-84, ISBN 1-902956-33-1, 2003

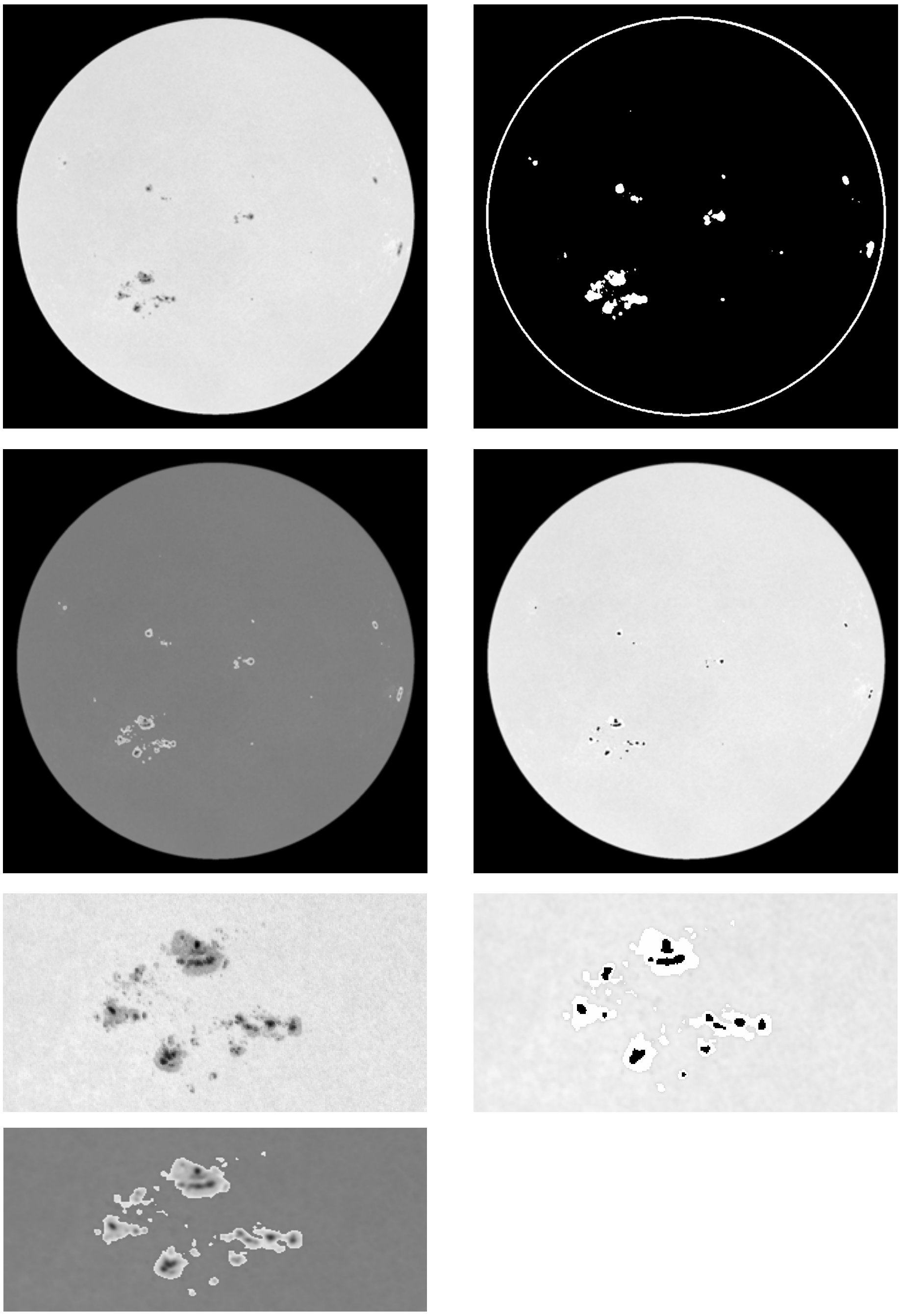

Fig. 1. An example of sunspot detection on a SOHO/MDI white light continuum image taken on 2nd April 2002.