Automatic Detection of Active Regions on Solar Images

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

Department of Cybernetics, University of Bradford, BD7 1DP, UK
a.k.benkhalil@brad.ac.uk

Abstract. In this paper techniques are described for the automated detection of solar Active Regions (ARs). AR detection is achieved using intensity thresholds and a region growing procedure. These procedures have been tested on full-disk solar images from the Meudon observatory for the months of April and July 2002 and compared with their manually generated synoptic maps. Comparisons were also made with AR data published by the National Oceanic and Atmospheric Administration observatory (NOAA) and very good correspondence was found.

1 Introduction

There are a growing number of archives of digitized images of the Sun taken from ground-based and space-based instruments in various wavelengths. These archives are available from different locations and are to be unified by the European Grid of Solar Observations (EGSO) project [1].

There are three different approaches identified in the literature for the automatic identification of bright ARs (plages). The first is based on the selection of a threshold to separate the object from a background and is straightforward if the intensity histogram is bimodal, but otherwise can be difficult [2]. The second approach is based on region growing techniques segmenting images into bright and dark regions [3, 4] and is applied to solar images in various wavelengths, including Hα, from a number of sources. Finally, the third approach uses the Bayesian inference method for automatically identifying various surface structures on the Sun [5]. All these approaches can give a reasonable accuracy of detection with suitable images, but the Bayesian based methods are the most computationally expensive of the three. The intensity threshold-based methods are simple and fast, but are relatively sensitive to noise which affects the reliability of the segmentation results obtained.

In order to replace the existing manual detection methods, the current paper presents techniques combining elements of the first two approaches above for the automated detection of ARs (plages) at different heights in the solar atmosphere, which are revealed at different wavelengths. The types of solar images which have been used are Hα and CaIIK3 line spectroheliograms from the Meudon observatory. The two types of image are processed using the same region growing technique but with different intensity thresholds in the initial processing stage, to find seed locations, and in the subsequent AR segmentation stage. The methods used for AR detection and

comparison of the resulting identified regions are discussed in Section II. Conclusions are given in Section III.

2 The Techniques for Active Regions Recognition

There are two basic assumptions made about the solar images when applying the developed techniques. The first is that the input images are standardized with size 1024 pixel×1024 pixel, solar disk radius 420 pixel, solar centre 511.5 pixel ×511.5 pixel, and free of a radial limb darkening. In order to comply with this assumption, the techniques are applied to full-disk high- resolution solar images which have been standardized using procedures [6] for limb fitting, shape correction and limb darkening removal. At the same time the images are also transferred to the required format. The second assumption is related to the properties of ARs, and is simply that they are the brightest features on the solar disk. This means that the intensity values inside the detected regions of interest are greater than the intensity values of the local background.

2.1 The Initial segmentation

In order to define a suitable local threshold all images were first remapped into polar coordinates with origin at the solar disc centre. After remapping, localized intensity thresholds, with values derived as explained below, are used for an initial segmentation of the bright plages. Pixels whose intensity values are over this intensity threshold have their values set to 1 and all other pixels have their values set to zero. The choice of these initial intensity threshold values is very importation because a value that is too high may lead to real features being missed, whereas a value that is too low may lead to noisier binary images and, hence, spurious features. The optimum global threshold value also varies with the image brightness levels and the non-radial large scale intensity variations, which are present in some Meudon images, are a particular problem. To overcome these problems optimized local intensity threshold values (T) are calculated for quarter-sized regions of an image as follows:

$$T_i = \mu_i + (1+\Delta_i)\times\sigma_i \tag{1}$$

where μ_i is the mean intensity value for the region i, σ_i is the standard deviation of the intensity for the same region and Δ_i is a constant that was set to 0.4 after investigating more than 30 images.

The main stages of this technique are illustrated in Fig. 1 for Hα, and in Fig. 2 for CaIIK3 full-disk images, respectively. Subfigures (a) present the cleaned initial images; subfigures (b) show the results of their remapping into the polar coordinates. The results of the initial segmentation based on equation (1) are presented in subfigures (c). Subfigures (d) and (e) show initial and final segmentation results discussed below.

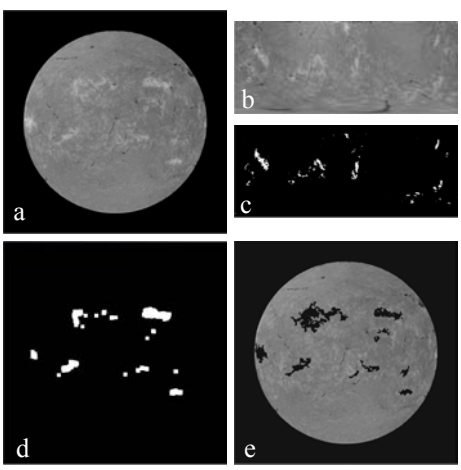

Fig. 1. Hα image detection stages

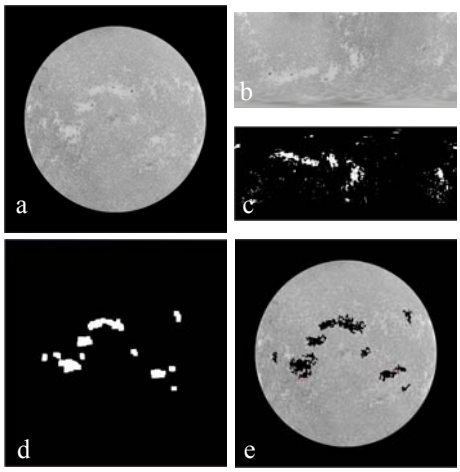

Fig. 2. CaIIK3 image detection stages

2.2 Noise reduction and Region labeling

The initial segmentation will generally include noise and unwanted small features caused by the over-segmentation. Over-segmentation is preferable to under-segmentation as the former can be remedied using Median filtering and Morphological operations whereas the latter could lose significant information. Firstly, in order to remove small features a 7×7 Median filter is used. The size was chosen through experimentation. Then Morphological opening and closing operations are applied using a structure element of size 8×8. This smoothes the features and fills in holes. Figures 1(d), 2(d) and 3(d) show the detected regions after applying Median and Morphological processing and transformation back to Cartesian coordinates. As can be seen, the noise and over-segmentation problems have been remedied.

The result of this initial segmentation is a set of the segments, each of which corresponds to an AR present on the solar disk. Every segment is labeled, and its centroid is calculated for use as a seed in a region growing procedure. Prior to this, because the shape may be complex, the location of the seed is checked in order to ensure that it is inside the region and, if not, its position is adjusted. In this case a new seed is selected by investigating pixel values in the eight nearest neighbor directions, until a new seed is found inside the region from which the region growing procedure can start.

2.3 The Region growing technique

The region growing procedure starts with a set of seed pixels and aims to grow a uniform and connected region from each seed. A pixel is added to a growing region if and only if:

- It has not been assigned to another region
- It is an 8-neighbour of the growing region
- The extended region created by the addition of a new pixel is still uniform.

The region growing algorithm takes as input a standardized image and a corresponding set of seed points obtained by the procedure described in section 2.2. The algorithm begins at each seed pixel and scans the neighboring 8 pixels in a circular fashion, to determine a membership of the region around the central pixel that complies with the rules above and the following constraints. Two forms of constraint have been considered. The first uses a fixed threshold range (with an upper and lower pixel value) and second uses a variable threshold range set to a multiple factor of the standard deviation of the pixel values in the current region. After experimentation the fixed threshold range was chosen as it was found to give more accurate control in defining the outer boundaries of regions while also reducing the occurrence of holes in the regions.

The upper and lower threshold values within initially detected ARs are determined by exploiting the statistical properties of the locally homogeneous background regions. The lower threshold value is defined as $\mu - 0.3\sigma$ (where μ is the mean and σ is the standard deviation of that region) and the upper threshold is set to the maximum intensity value of the image. As pixels are added, the process repeats with the newly added pixels as the central pixels. A list of the tested pixels is maintained, in order to avoid unnecessary duplication of the tests. In this way, the region is constructed by using the membership criteria already discussed. If in a binary image more than one seed pixel has been obtained from an AR, the region growing method will merge the detected pixels to form a single contiguous area of AR. Figs. 1(e) and 2(e) show the final results of applying the region growing procedure in the Hα and CaIIK3 images, respectively.

2.4 AR verification using magnetograms

It is not possible to confirm or deny the identification of ARs on the basis of CaIIK3 data alone and for this reason we started to look at magnetogram data also. Series of FITS file format Michelson Doppler Imager (MDI) images are obtained from the MDI instrument aboard the SOHO satellite. These images are one-minute cadence, full-disk magnetograms. The CaIIK3 images from ground observations at Meudon are acquired once a day. As The CaIIK3 and MDI images are taken at different time and from different locations (and have different sizes) the images need to be synchronized.

Firstly we select and download the MDI image that is closest in time to the Meudon CaIIK3 observation. This is generally straightforward as we have MDI image every minute. Secondly, to correct MDI-image sizes and to convert them to earth view, the MDI images are converted to Interactive Data Language (IDL) map objects using the *index2map.pro* IDL function. The Map object is a structure that contains 2D image data with accompanying pixel coordinate and spatial scale information. The map is converted from SOHO-view map to Earth-view using the *map2earth.pro* func-

tion. To make the dimensions and pixel spacing of a SOHO-MDI map object identical to the CaIIK3 images (i.e. size 1024 pixel×1024 pixel, radius 420 pixel and solar centre 511.5 pixel ×511.5 pixel) we use the *grid_map.pro* function. Finally the map objects are converted back to index images using the *map2index.pro* function. The final results are MDI images synchronized to earth view and standardized in shape and size to the CaIIK3 Meudon images.

Once an AR is detected the system will automatically crop two images, with identical sizes and locations, one from the CaIIK3 image and one from the MDI image. The latter data is then checked to identify positive and negative polarity regions which, if found, are used to confirm the detection of an AR and also to classify the ARs into newborn, highly active and decaying categories

Fig. 3 shows the results of tracking and investigating an AR (NOAA 9893) over three days 8-10/04/2002. The left column contains the CaIIK3 cropped images, the middle column contains the MDI cropped images, where the black areas indicate positive polarity and the white area indicate negative polarity, and the right column shows the magnetic field superimposed on the CaIIK3 images.

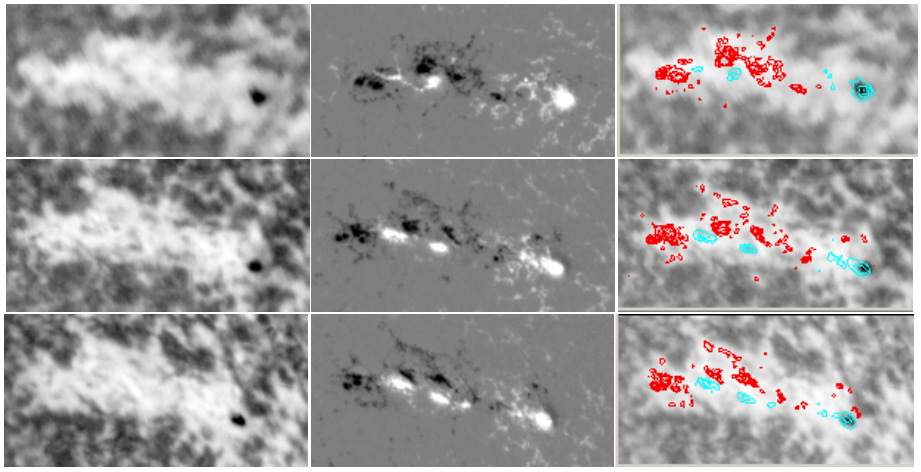

Fig. 3. Cropped images containing AR NOAA 9893. Left column CaIIK3, middle column magnetogram and right column showing magnetic field polarity superimposed on CaIIK3 image.

2.4 The accuracy of the technique

The procedures have been tested on synoptic image sequences of full-disk solar images from the Meudon observatory for the months of April and July 2002. For further testing the results obtained from the Meudon images were compared with those of the NOAA observatory as illustrated in Figs. 4.

A quantitative comparison of the results obtained using the present technique, with those done manually at the Meudon observatory and at NOAA observatory was done

for the two month (April and July 2002). In comparison with the other results, those from Meudon detect about 50% more ARs on most days. For example, on the 30/07/2002, and as shown in Fig. 4, there were 24 ARs included in the Meudon results while our procedure detected only 11 ARs and the NOAA observatory showed only 12 ARs. In order to quantify these differences the False Acceptance Rate (FAR) (where we detect an AR and they do not) and the False Rejection Rate (FRR) (where they detect an AR and we do not) were calculated for every day. In most cases there are a higher number of ARs detected by us than by NOAA with an average FAR of 1.8 per day in April and only 1 in July. The FRR was very low at about 0.2 in both months, with only 5 days in each month when we failed to detect a region detected by NOAA. In some cases we detect an AR while NOAA splits it into two regions. This does affect the quantitative comparison.

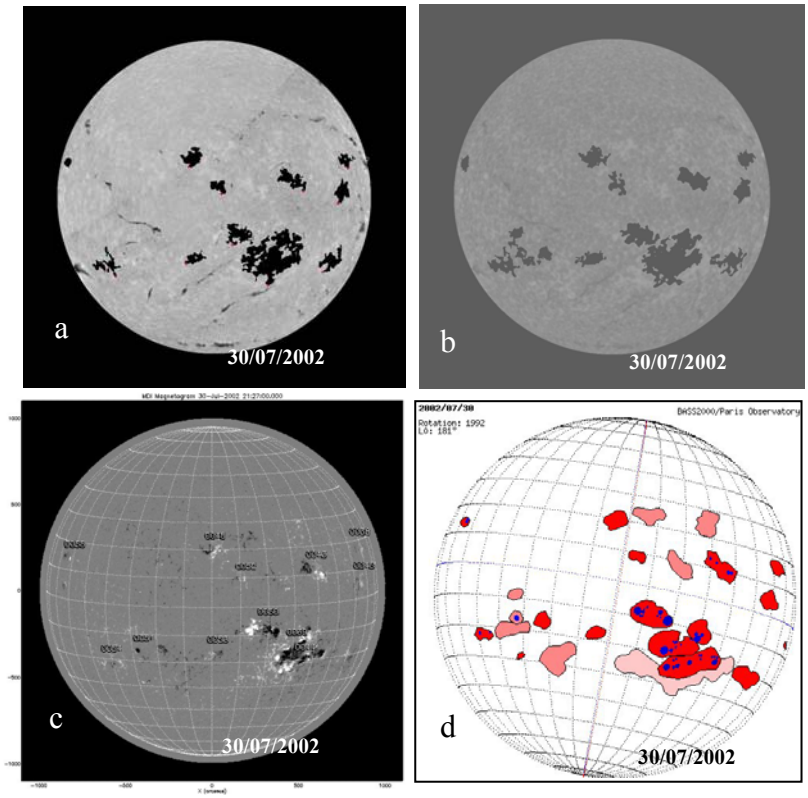

Fig. 4. A comparison of AR detection results. The present results for Hα (a), CaIIK3 solar images (b) showing 12 ARs. Results from the NOAA observatory (c) showing 12 ARs. Results from the Meudon observatory (d) showing a map of 24 ARs.

We believe the reason for these different results is due to a difference in definitions of ARs. At Meudon all bright regions (plages) are detected and these are defined as the regions in the chromosphere which are brighter than the normal "quiet" Sun back-

ground. At NOAA a detected AR is defined as a bright area on the Sun with a large concentration of magnetic field, often containing sunspots. However, not all plages contain a strong magnetic field as they might be decaying ARs with a weakening magnetic field [7]. Fig. 4 clearly illustrates this case by showing the results of ARs detection at NOAA, Meudon and using our techniques with Hα and CaIIK3 images on the same day (30/07/02). In Fig. 4(e) the Meudon map shows 24 ARs (all the white area (plages) are counted) resulting in double the number detected by us and by NOAA. In general, the agreement with NOAA is good, considering that NOAA bases its decisions on more information about magnetic field than we do at this stage.

3 Conclusions

In this paper an efficient procedure for the automated detection of solar ARs is presented. Comparisons with other recognition results derived manually show that the procedures developed have achieved a satisfactory accuracy in the automated detection and segmentation of ARs on full disk Hα and CaIIK3 solar images from Meudon. The automated solar AR detection and characterization is one component of a larger project concerned with establishing a Solar Feature Catalogue. This research is a part of the EGSO project funded by the European Commission within the IST Framework 5.

References

[1] R.D. Bentley, "EGSO - the next step in data analysis," *Proceedings of the Second Solar Cycle and Space Weather Euro-conference*, 24 - 29 September 2001, Vico Equense, Italy, Edited by Huguette Sawaya-Lacoste, ESA Publication SP-477, 2002.

[2] M. Steinegger, and P. N. Brandt, "On the determination of the quiet Sun centre-to-limb variation in Ca K spectroheliograms," *Solar Physic Journal,* vol. 177, pp.287-294, 1998.

[3] M. Hill, V. Castelli, L. Chung-Sheng, C. Yuan-Chi, L. Bergman, and B. Thompson, "Solarspire: querying temporal solar imagery by content," *International Conference on Image Processing*, Thessaloniki, Greece, 7-10 Oct. 2001, vol. 1, pp. 834–837, 2001.

[4] A. Veronig, M. Steinegger, W. Otruba, A. Hanslmeier, M. Messerotti, and M. Temmer, "Automatic Image Processing In The Frame Of A Solar Flare Alerting System," *HOBUD7*, vol. 24, no. 1, pp.195-200, 2001.

[5] M. Turmon**,** J. M. Pap, and S. Mukhtar, , "Automatically Finding Solar Active Regions Using Soho/Mdi Photograms And Magnetograms," *Proc. SoHO 6/GONG 98 Workshop, Structure and Dynamics of the Interior of the Sun and Sun-like Stars*, Boston, 1998.

[6] V. V. Zharkova, S. S Ipson, S. I Zharkov, A. K Benkhalil and J. Aboudarham, "A full disk image standardisation of the synoptic solar observations at the Meudon observatory," *Solar Phys*., vol. 214/1, pp. 89-105, 2003.

[7] L. V Driel-Gesztelyi, "Emergence and Loss of Magnetic Flux on the Solar Surface," *Proc. SOLMAG the Magnetic Coupling of the Solar Atmosphere : Euroconference and IAU Colloquium* 188, Santorini,Greece, 11-15 June 2002, pp.113-116, 2002.