Computational Engineering in Systems Applications, CESA'2003

An Automated Detection of the 3D-Structure of Active Regions from Full Disk Solar Images

Valentina Zharkova, Ali Benkhalil, Sergei Zharkov, and Stanley Ipson
Cybernetics Department, University of Bradford, BD7 1DP, UK
Phone: +44-1274-234269, Fax: +44-1274-236600, E-mail: a.k.benkhalil@bradford.ac.uk

Abstract - The techniques are developed for an automated detection of Active Regions, bright areas on a darker background, using Hα, Ca II K3 solar images from the Meudon observatory and EUV solar images from the EIT instrument aboard of the Solar and Heliospherical Observatory (SOHO). An initial segmentation of active regions is achieved using intensity thresholds determined using statistical information from each quarter of full disk solar images. Median filtering and morphological operations are applied to the resulting binary image to remove noise and to merge broken regions. The centre of gravity of each labelled region is used as a seed for a region growing procedure. Statistically based local thresholding is applied to calculate upper and lower threshold values defining the spatial extents of the regions. This procedure has been tested on full-disk solar images from the Meudon observatory for the whole month of April 2002 and compared with their manual synoptic maps. The active regions detected using the presented technique were also compared with those of the National Oceanic and Atmospheric Administration observatory (NOAA) revealing a very good correspondence. The vertical (pseudo-3D) structures of active regions observed from the corona, chromosphere and the photosphere are presented for the few biggest features present during this month.

I. INTRODUCTION

There are a growing number of archives of digitised images of the Sun taken from ground-based and space instruments in various wavelengths. These archives are available from different locations and are to be unified by the EGSO project [BEN, 01]. The digitized solar images have different sizes, resolutions, instrumental and weather distortions. All are to be subjected to an automated recognition process in order to provide a reliable location of features and their evolution with time relative to solar rotation. With the growing demand for forecasts of solar activity by the space weather Living with the Star (LWS) project and by many industrial organisations, there is a great need for the development of reliable and fast techniques for automated feature recognition on solar disks and their unification into feature catalogues.

The solar atmosphere consists of many features with different characteristics and distinct physical origins. The most clearly distinguished in the Hα-line images are active regions, which are closely associated with sunspots in the white light images and plages in Ca II K3 spectroheliograms observed from the photosphere. The plages can be large in extent and normally are associated with a strong magnetic field and increased brightness. Bipolar active regions are widely accepted to be the manifestations of emerging buoyant magnetic flux tubes [ZWA, 87]; [DRI, 02] and the references therein). These flux tubes are seen on the photosphere in white light images as sunspots or in the corona as very diffusive bright areas in extreme ultraviolet (EUV) emission. Filaments on the photosphere or filament channels in the corona are often found to be located alongside active regions separating the magnetic fields of opposite polarity [SCH, 02a]; [AUL, 02]. The same active regions are often found to produce numerous flares during their life time, in a process that is related to the changes in their magnetic configurations [DRI, 02], [SCH, 02b]. These magnetic field changes can occur on even smaller scales related to the emergence of mini and nano-flux tubes causing smaller flaring events [SCH, 02b]. This produces different responses at different levels of the solar atmosphere, which allows the testing of theoretical models describing these responses.

Therefore, the structure of active regions at various levels of the solar atmosphere can provide a key to the understanding and proper forecast of solar activity manifestations such as: solar flares, coronal mass ejections (CMEs), eruptive filaments etc. Their reliable automated detection will facilitate the building of a major database of solar active features and enable the analysis of solar activity on a comprehensive database of active regions taken in various wavelengths.

There are three different approaches identified in the literature for automatic identification of bright active regions (plages). The first one 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 [STE, 98]. The second approach is based on region growing techniques applied in a number of solar images in various wavelengths, including Hα, segmenting the image into bright and dark regions [HIL, 01]; [VER, 01]. Finally, the third approach uses the Bayesian inference method for automatically identifying various surface structures on the Sun [TUR, 98]. All these approaches can give a reasonable accuracy of detection, but the Bayesian based methods were concluded to be relatively computationally expensive.

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, in the current paper techniques combining the first two approaches above are developed for an automated detection of Active Regions (plages). The following three types of solar images are used: Hα and Ca II K3 line spectroheliograms from the Meudon observatory and EUV images (Fe XIV 171Å and Fe XII, 195 Å) from SOHO/ Extreme ultraviolet Imaging Telescope (EIT) with the aim of building a pseudo-3D reconstruction of the detected active regions. The region growing technique, with some modifications, is used in each wavelength in conjunction with intensity thresholds which are applied twice: in the initial processing stage, in order to find seed locations, and in the subsequent active region segmentation. The methods used for active regions detection and the resulting identified regions are discussed in Section II and conclusions are given in Section III.

II. THE TECHNIQUES FOR ACTIVE REGIONS RECOGNITION

There are two basic assumptions about the solar images used by the developed techniques. The first is that the input images are of standardised size (1024 pixel×1024 pixel) and solar disk radius (420 pixel) with 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 on full-disk high-resolution solar images which have been standardized using procedures [ZHA, 03] for limb fitting, shape correction and limb darkening removal that also convert the images into the required format. The second assumption is related to the active region properties, namely, that active regions 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.

A. 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, statistically derived localised intensity threshold values 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 particular problems in some Meudon images. To overcome these problems optimised local intensity threshold values (T) are calculated for quarter-sized regions of an image as follows:

$$T_i = \mu_i + (1+\Delta_i)\times\sigma_i \qquad (1)$$

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, in Fig. 2 for Ca II K3 and in Fig. 3 for Fe XII 195 Å 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.

B. The Median filtering, Morphological operations and Region labelling

The initial segmentation will include noise and unwanted small features caused by the over-segmentation. The over-segmentation is preferable to under-segmentation as the former can be remedied using Median filtering and Morphological operations while 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. Figs. 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 completely remedied.

The result of this initial segmentation is a set of the segments, each of which corresponds to an active region present on the solar disk. Every segment is labelled, and its centroid is calculated for use as a seed in a region growing procedure. Prior to this, 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 neighbour directions, as illustrated in Fig. 4, until a new seed is found inside the region from which the region growing procedure can start.

C. The Region growing technique

Region growing is a basic procedure that starts with a set of seed pixels. The aim is 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 new region created by the addition of a new pixel is still uniform.

The algorithm used for the region growing takes as input a clean image and a set of seed points from which the region-growing process starts. The algorithm begins at each seed pixel and scans the neighbouring 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. The two forms of the constraints 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 a standard deviation of the current region pixel values. 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 active regions are determined by exploiting the statistical properties of the locally homogeneous background regions. As illustrated in the pseudo code below, 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 centre 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 pixels have been obtained from an active region, the region growing method merges the detected pixels and forms a single active region area. Figs. 1(e), 2(e) and 3(e) show the final results of applying the region growing procedure in the Hα, Ca K II and Fe XII 195Å images, respectively.

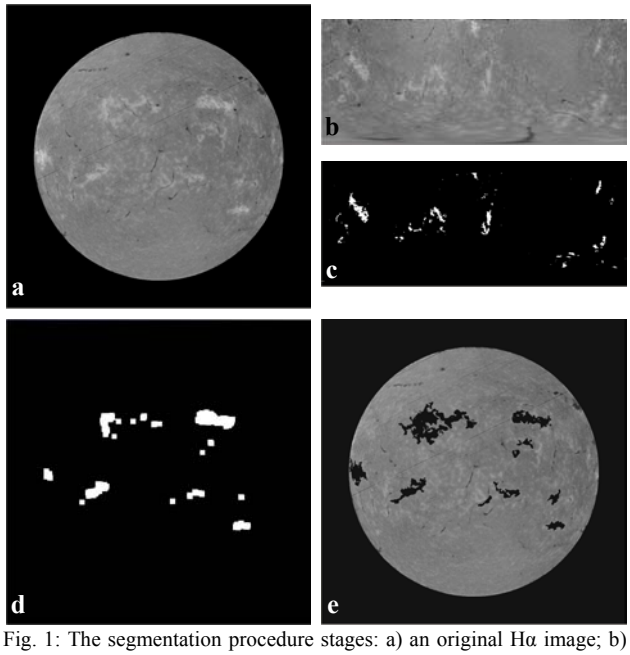

Fig. 1: The segmentation procedure stages: a) an original Hα image; b) after a transformation to Polar coordinates; c) after an initial thresholding; d) after a transformation back to the Cartesian coordinates, cleaning and morphological processing and e) a final result of the region growing.

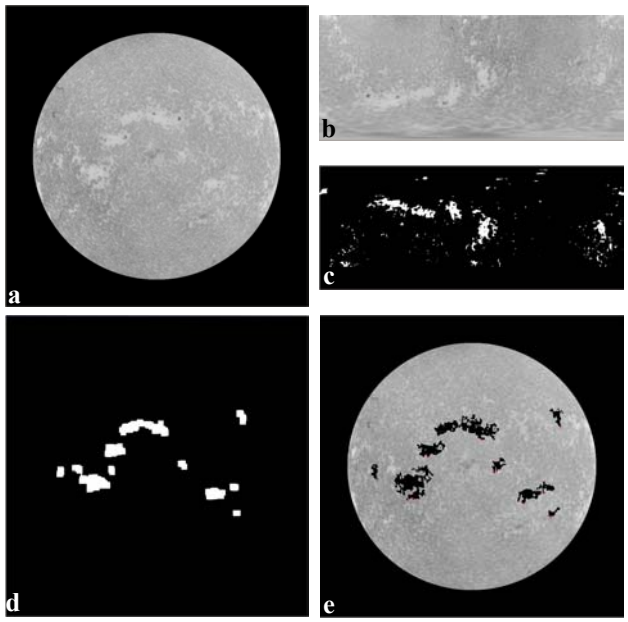

Fig. 2: The segmentation procedure stages: a) an original Ca II K3 image; b) after a transformation to Polar coordinates; c) after an initial thresholding; d) after a transformation back to the Cartesian coordinates, cleaning and morphological processing and e) a final result of the region growing.

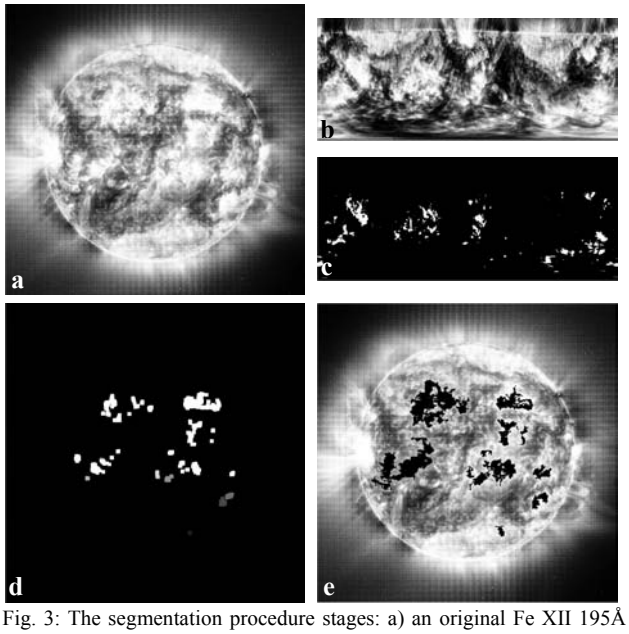

Fig. 3: The segmentation procedure stages: a) an original Fe XII 195Å image; b) after a transformation to Polar coordinates; c) after an initial thresholding; d) after a transformation back to the Cartesian coordinates, cleaning and morphological processing and e) a final result of the region growing.

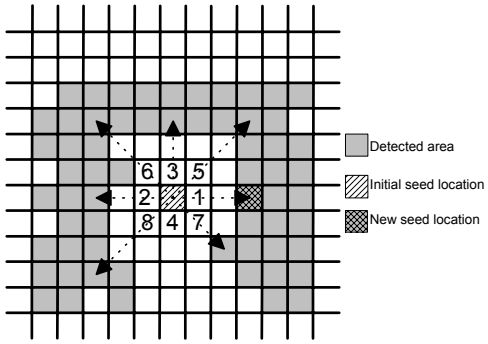

Fig. 4: The search scheme for a seed pixel location.

The following pseudo-code defines the region-growing algorithm:

Let f be an image, and $R_1, R_2,...R_n$ a set of regions each consisting of a single seed pixel.

```
for i=1...n
  Repeat
    for each pixel p at the border of R_i
      for all neighbours of p
        Let x,y be the neighbour's coordinates
        Let lo_i be the lowest grey level of pixels in R_i
        lo_i = μ_i – 0.3 ×σ_i
        Let hi_i be the highest grey level of pixels in R_i
        hi_i = max(f)
        if neighbour unassigned and
          (hi_i ≥ f(x,y) ≥ (lo_i + Δ))
          Add neighbour to R_i, update hi_i and lo_i
  Until no more pixels are being assigned to region
```

D. The images synchronisation and 3D reconstruction

In order to perform a morphological analysis of the active regions detected in various wavelengths the images have to be synchronised in time, matched in size and overlaid. The images in the Hα and Ca II K3 wavelengths are obtained from the ground-based observatory in Meudon and have a size of 927x923 pixel while the Fe XII, 195Å images are obtained from the SOHO space satellite and have a size of 1024x1024 pixel. These images are also captured at slightly different time. This requires the images to be transformed from the heliographic coordinates into heliocentric ones and the pixels from these images to be mapped to a new synchronised image using a back-projection technique. Then the images from all three or more wavelengths can be overlaid and boundaries and pixel values of active regions can be compared. Fig. 5 shows overlaid images from the three wavelengths illustrating the synchronisation results.

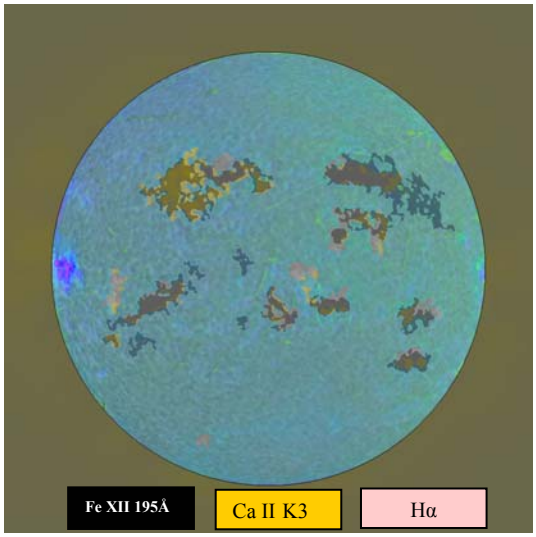

Fig. 5: The Active Regions overlaid on the images taken on the same day (12/04/2002)in the three wavelengths Hα, Ca II K3 and Fe XII 195Å.

Now we can estimate the 3D structure of an active region detected at the three wavelengths: Hα, Ca K II and Fe XII 195Å. First the detected regions are cropped from the original images to rectangular fragments of size 380 × 300 pixel. Fig. 6 shows the cropped fragments of the same active region detected from images captured on 12/04/2002.

Bearing in mind that the images in different wavelengths show features at different heights in the solar atmosphere, it is possible to place these 2D-fragments into a 3-D plot at different heights to indicate their 3D-structure. Fig. 7 presents a pseudo-3D structure of these regions showing the effective height of line formation in a given wavelength together with a 3D wire structure of the active region in these wavelengths. A comparison of the active region fragments detected on different heights in solar atmosphere and different times and dates allows us to investigate an active region morphology, intensity and evolution during its lifetime. This will be the scope of a forthcoming paper.

E. 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 month of April 2002. For the further testing the results obtained from the Meudon images were compared with those of the NOAA observatory and the Big Bear Solar Observatory (BBSO) as illustrated in Table I and Figs. 8-9.

A quantitative comparison of the results obtained using the present technique, with those done manually at the Meudon observatory [MOU, 98] and at BBSO/NOAA observatory is shown in Table I. In comparison with other results, those from Meudon detect about 50% more active regions on most days. For example, on the 07/04/2002, there were 22 active regions in the Meudon results while our procedure detected only 13 active regions and the NOAA observatory showed only 12 active regions. In order to quantify these differences the False Acceptance Rate (FAR) (where we detect an active region and they do not) and the False Rejection Rate (FRR) (where they detect an active region and we do not) were calculated for every day and presented in the last two columns of Table I. As can be seen, in most cases there is a higher number of active regions detected by us than by NOAA with an average FAR of 1.7 per day. The FRR was very low at about 0.2, with only 5 days when we failed to detect a region detected by NOAA. In some cases, as indicated in Fig. 8, we detect an active region while NOAA splits it into two regions. This does affect the quantitative comparison.

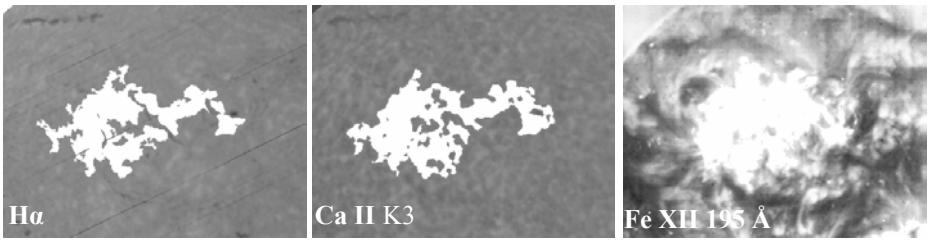

Fig. 6: The fragments of the same active region detected from the solar images on 12/04/2002 in the three wavelengths.

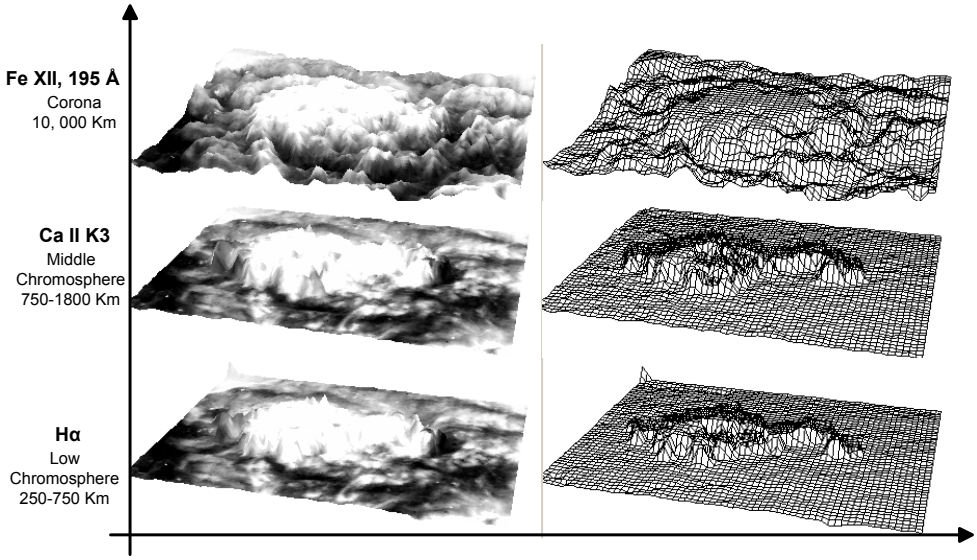

Fig. 7: The active region fragments detected from the three wavelengths: Hα, Ca K II, and Fe XII 195Å.

TABLE I
A COMPARISON OF ACTIVE REGIONS DETECTED WITH THE PRESENT TECHNIQUE IN APRIL 2002 WITH THOSE FROM MEUDON AND BBSO/NOAA (SEE EXPLANATION IN THE TEXT)

Date	**Meudon**	**BBSO/ NOAA**	**Ours**	**FAR**	**FRR**	**NOTE**
01/04/02	19	6	8	3	0	We have merged 2 regions to 1
02/04/02	14	6	8	2	0	
03/04/02	18	9	12	3	0	
04/04/02	22	8	11	3	0	
05/04/02	**	7	11	6	0	We have merged 2 regions to 1
06/04/02	24	9	9	1	0	We have merged 2 regions to 1
07/04/02	22	12	13	1	0	
08/04/02	24	8	11	4	0	We have merged 2 regions to 1
09/04/02	26	11	11	3	0	We have merged 4 regions to 2
10/04/02	**	10	15	5	0	
11/04/02	**	10	10	2	0	We have merged 4 regions to 2
12/04/02	20	11	8	0	0	We have merged 6 regions to 2
13/04/02	21	12	**	**	**	
14/04/02	21	9	10	1	0	
15/04/02	22	6	7	1	0	
16/04/02	20	7	7	0	0	
17/04/02	19	6	6	0	0	
18/04/02	18	5	7	2	0	
19/04/02	21	6	7	1	0	
20/04/02	17	6	6	0	0	
21/04/02	15	7	9	2	0	
22/04/02	14	7	8	1	0	
23/04/02	15	9	9	3	2	We have merged 2 regions to 1
24/04/02	17	8	9	3	2	
25/04/02	14	10	10	0	0	
26/04/02	11	8	**	**	**	
27/04/02	11	6	8	3	1	
28/04/20	7	8	7	0	1	
29/04/02	6	7	7	2	1	We have merged 2 regions to 1
30/04/02	**	7	8	1		
Total	**458**	**241**	**251**	**53**	**7**	****no data on that day**

The reason for these differences is owing to a different definition of active regions. At Meudon all bright regions (plages) are detected, and these are defined as the regions in the chromosphere that are brighter than the normal "quiet" Sun background. At NOAA a detected active region 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 active regions with a weakening magnetic field [DRI, 02]. Fig. 9 clearly illustrates this case by showing the results of active regions detection at NOAA, Meudon and using our techniques with Hα, Ca II K3 (Meudon) and Fe XII 195Å (SoHO/EIT) solar images on the same day (01/04/2002). In Fig. 9(e) the Meudon map shows 19 active regions (all the white area (plages) are counted) resulting in double the number detected by us and by NOAA. In general, the agreement with both Meudon and NOAA is good, considering that NOAA bases its decisions on more information about magnetic field than we do at this stage.

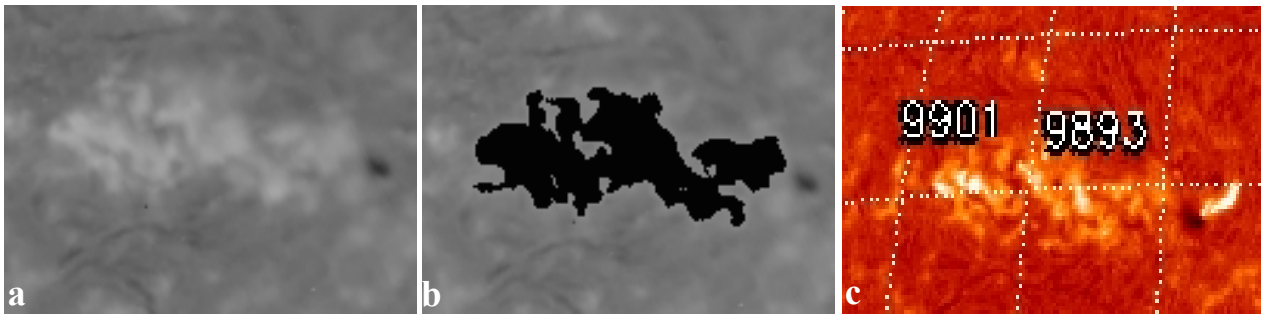

Fig. 8: The example of two active regions merge: a) the Meudon Hα input image, b) the same image after applying the region growing procedure and c) the BBSO Hα image shown the same region divided into the two active regions.

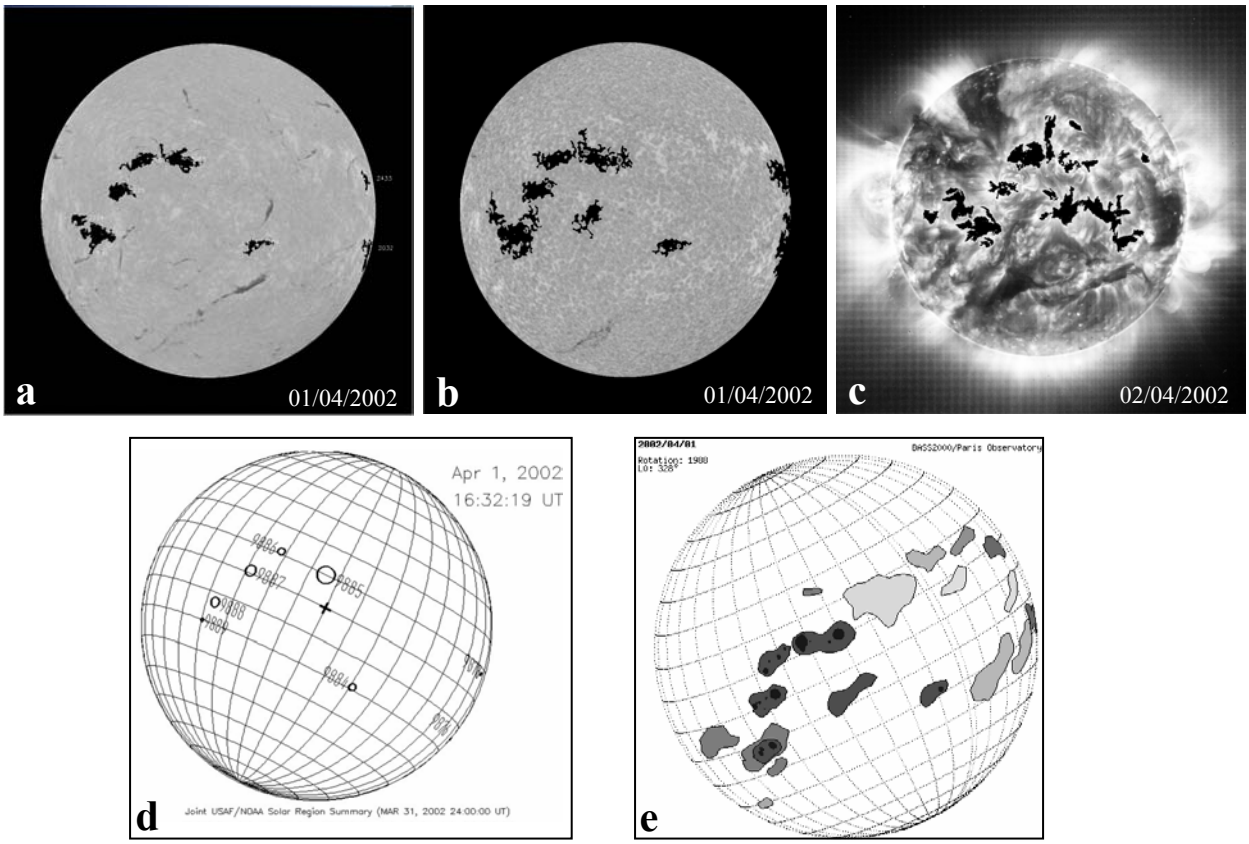

Fig. 9: A comparison of Active Region detection; the present results on Hα (a), Ca II K3 (b) and Fe XII 195Å (c) solar image showing 8 active regions with the NOAA observatory map (d) showing 8 active region and the Meudon observatory map (e) showing 19 active regions The images are taken on 01/04/2002) except the Fe XII 195Å image on the 02/04/2003.

III. CONCLUSIONS

In this paper an efficient procedure for the automated detection of solar active regions is presented. Statistically derived local intensity threshold values and region growing methods are used in the procedure. A local intensity threshold is applied to detect bright pixels that are then used as seed candidates for growing active region. The subsequent region-growing algorithm exploits statistical properties of the locally homogeneous background regions in order to determine the local upper and lower threshold values which control the spatial extent of a region growth. The procedures developed have achieved a satisfactory accuracy in the automated detection and segmentation of active regions on the full disk solar images from the Meudon Hα, Ca II K3 and from SOHO/Fe XII 195Å. A pseudo-3D structure for the active regions, detected in these three wavelengths, was estimated to demonstrate the active region variations in different wavelengths which correspond to different heights of the solar atmosphere. A comparison of the active region fragments detected on different heights will allow us to investigate the active region morphology and evolution during its lifetime. This will be the subject of a forthcoming paper.

This research is a part of the European Grid of Solar Observations (EGSO) project funded by the European Commission within the IST Framework 5.

REFERENCES

[AUL, 02] Aulanier G. [1] and Schmieder B., "The magnetic nature of wide EUV filament channels and their role in the mass loading of CMEs," *Astronomy and Astrophysics*, vol. 386, pp 1106-1122, 2002.

[BEN, 02] Bentley, R.D., "EGSO - the next step in data analysis," in 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.

[DRI, 02] Driel-Gesztelyi, L. V., "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.

[HIL, 02] Hill M, Castelli V, Chung-Sheng Li, Yuan-Chi Chang, Bergman L, Smith J.R. and Thompson B., "Solarspire: querying temporal solar imagery by content," International Conference on Image Processing, Thessaloniki, Greece, 7-10 Oct. 2001, vol. 1, pp. 834–837, 2001.

[MOU, 98] Mouradian Z., "Synoptic Data Findings," Synoptic Solar Physics ASP Conferences Series, vol. 140, pp. 181-204, 1998.

[SCH, 02a] Schmieder, B., Engvold, O., Lin, Y., Deng, Y.Y., Mein, N., "Magnetic Cancellation and Small-Scale Activity in An AR Filament,": proc. of SOLMAG 2002 the Magnetic Coupling of the Solar Atmosphere : Euroconference and IAU Colloquium 188, 11 - 15 June 2002, Santorini, Greece, pp.223-226, 2002.

[SCH, 02b] Schmieder, B., Aulanier G., Pariat E., Georgoulis M.K., Rust D. M. and Bernasconi P. N."Vector Magnetic Field Observations of Flux Tube Emergence,": proc. of SOLMAG 2002 the Magnetic Coupling of the Solar Atmosphere : Euroconference and IAU Colloquium 188, 11 - 15 June 2002, Santorini, Greece, pp.223-226, 2002.

[STE, 98] Steinegger, M and Brandt, P.N., "On the determination of the quiet Sun centre-to-limb variation in Ca K spectroheliograms," *Solar Physic Journal,* vol. 177, pp.287-294, 1998.

[TUR, 98] Turmon**,** M., Pap J. M., and Mukhtar, S., "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.

[VER, 01] Veronig A., Steinegger M., Otruba W., Hanslmeier A., Messerotti M., and Temmer M., "Automatic Image Processing In The Frame Of A Solar Flare Alerting System," *HOBUD7*, vol. 24, no. 1, pp.195-200, 2001.

[ZHA, 03] Zharkova V.V., Ipson S. S, Zharkov S. I, Benkhalil A. K and Aboudarham J., "A full disk image standardisation of the synoptic solar observations at the Meudon observatory," *Solar Phys*., vol. 214/1, pp. 89-105, 2003.

[ZWA, 87] Zwaan, C., "Elements And Patterns In The Solar Magnetic Field," *Ann. Rev. Astron. Astrophys,* vol. 25 , pp.83-111, 1987.