Segmentation, Tracking and Characterization of Solar Features from EIT Solar Corona Images

Vincent Barra[1], Véronique Delouille[2], and Jean-Francois Hochedez[2]

[1] LIMOS, UMR 6158, Campus des Cézeaux, 63173 Aubière, France
`vincent.barra@isima.fr`
[2] Royal Observatory of Belgium, Circular Avenue 3, B-1180 Brussels, Belgium
`{verodelo,hochedez}@sidc.com`

Abstract. With the multiplication of sensors and instruments, size, amount and quality of solar image data are constantly increasing, and analyzing this data requires defining and implementing accurate and reliable algorithms. In the context of solar features analysis, it is particularly important to accurately delineate their edges and track their motion, to estimate quantitative indices and analyse their evolution through time. Herein, we introduce an image processing pipeline that segment, track and quantify solar features from a set of multispectral solar corona images, taken with eit EIT instrument. We demonstrate the method on the automatic tracking of Active Regions from EIT images, and on the analysis of the spatial distribution of coronal bright points. The method is generic enough to allow the study of any solar feature, provided it can be segmented from EIT images or other sources.

Keywords: Segmentation, tracking, EIT Images.

1 Introduction

With the multiplication of both ground-based and onboard satellites sensors and instruments, size, amount and quality of solar image data are constantly increasing, and analyzing this data requires the mandatory definition and implementation of accurate and reliable algorithms. Several applications can benefit from such an analysis, from data mining to the forecast of solar activity or space weather. More particularly, solar features, such as sunspots, filaments or solar flares partially express energy transfer processes in the Sun, and detecting, tracking and quantifying their characteristics can provide information about how these processes occur, evolve and affect total and spectral solar irradiance or photochemical processes in the terrestrial atmosphere.

The problem of solar image segmentation in general and the detection and tracking of these solar features in particular has thus been addressed in many ways in the last decade. The detection of sunspots [18,22,27], umbral dots [21] active regions [4,13,23], filaments [1,7,12,19,25], photospheric [5,17] or chromospheric structures [26], solar flares [24], bright points [8,9] or coronal holes [16] mainly use classical image processing techniques, from region-based to edge-based segmentation methods.

In this article we propose an image processing pipeline that segment, track and quantify solar features from a set of multispectral solar corona images, taken with eit EIT instrument. The EIT telescope [10] onboard the SoHO ESA-NASA solar mission takes daily several data sets composed of four images (17.1 nm, 19.5 nm, 28.4 nm and 30.4 nm), all acquired within 30 minutes. They are thus well spatially registered and provide for each pixel a collection of 4 intensities that potentially permit to recognize the standard solar atmosphere region, or more generally solar features, to which it belongs..

This paper is organized as follows : section 2 introduces the general segmentation method. It basically recalls the original $SPoCA$ algorithm, then specializes it to the automatic segmentation and tracking of solar features, and finally introduces some solar features properties suitable for the characterization of such objects. Section 3 demonstrate some results on some EIT images of a 9-year images dataset spanning solar cycle 23, and section 4 sheds lights on perspectives and conclusion.

2   Method

2.1   Segmentation

We introduced in [2] and refined in [3] $SPoCA$, an unsupervised fuzzy clustering algorithm allowing the fast and automatic segmentation of coronal holes, active regions and quiet sun from multispectral EIT images. In the following, we only recall the basic principle of this algorithm, and we more particularly focus on its application for the segmentation of solar features.

SPoCA. Let $I = (I^i)_{\{1 \leq i \leq p\}}$, $I^i = (I^i_j)_{\{1 \leq j \leq N\}}$, be the set of p images to be processed. Pixel j, $1 \leq j \leq N$ is described by a feature vector x_j. x_j can be the p-dimensional vector $(I^1_j \cdots I^p_j)^T$ or any r-dimensional vector describing local properties (textures, egdes,...) of j. In the following, the size of x_j will be denoted as r. Let $\mathcal{N}_j$ denote the neighborhood of pixel j, containing j, and $\mathrm{Card}(\mathcal{N}_j)$ be the number of elements in $\mathcal{N}_j$. In the following, we note $X = \{x_j, 1 \leq j \leq N, x_j \in \mathbb{R}^r\}$ the set of feature vectors describing pixels j of I.

SPoCA is an iterative algorithm that searches for C compact clusters in X by computing both a fuzzy partition matrix $U = (u_{ij}), 1 \leq i \leq C, 1 \leq j \leq N$, $u_{i,j} = u_i(x_j) \in [0, 1]$ being the membership degree of x_j to class i, and unknown cluster centers $B = (b_i \in \mathbb{R}^r, 1 \leq i \leq C)$. It uses iterative optimizations to find the minimum of a constrained objective function:

$$J_{SPoCA}(B, U, X) = \sum_{i=1}^{C} \left(\sum_{j=1}^{N} u_{ij}^m \sum_{k \in \mathcal{N}_j} \beta_k d(x_k, b_i) + \eta_i \sum_{j=1}^{N} (1 - u_{ij})^m \right) \quad (1)$$

subject for all $i \in \{1 \cdots C\}$ to $\sum_{j=1}^{N} u_{ij} < N$, for all $j \in \{1 \cdots N\}$ to $\max_i u_{ij} > 0$, where $m > 1$ is a fuzzification parameter [6], and

$$\beta_k = \begin{cases} 1 & \text{if} \quad k = j \\ \frac{1}{\text{Card}(\mathcal{N}_j) - 1} & \text{otherwise} \end{cases} \tag{2}$$

Parameter η_i can be interpreted as the mean distance of all feature vectors $\boldsymbol{x_j}$ to $\boldsymbol{b_i}$ such that $u_{ij} = 0.5$. η_i can be computed as the intra-class mean fuzzy distance [14]:

$$\eta_i = \frac{\displaystyle\sum_{j=1}^{N} u_{ij}^m d(\boldsymbol{x_j}, \boldsymbol{b_i})}{\displaystyle\sum_{j=1}^{N} u_{ij}^m}$$

The first term in (1) is the total fuzzy intra-cluster variance, while the second term prevents the trivial solution $U = 0$ and relaxes the probabilistic constraint $\sum_{j=1}^{N} u_{ij} = 1, 1 \leq i \leq C$, stemming from the classical Fuzzy-C-means (FCM) algorithm [6]. SPoCa is a spatially-constrained version of the possibilistic clustering algorithm proposed by Krishnapuram and Keller [14], which allows memerships to be interpreted as true degrees of belonging, and not as degrees of sharing pixels amongst all classes, which is the case in the FCM method.

We showed in [2] that U and B could be computed as

$$u_{ij} = \left[1 + \left(\frac{\sum_{k \in \mathcal{N}_j} \beta_k d(\boldsymbol{x_k}, \boldsymbol{b_i})}{\eta_i} \right)^{\frac{1}{m-1}} \right]^{-1} \quad \text{and} \quad \boldsymbol{b_i} = \frac{\displaystyle\sum_{j=1}^{N} u_{ij}^m \sum_{k \in \mathcal{N}_j} \beta_k \boldsymbol{x_k}}{2 \displaystyle\sum_{j=1}^{N} u_{ij}^m}$$

SPoCA provides thus coronal holes (CH), Active regions (AR) and Quiet Sun (QS) fuzzy maps $U_i = (u_{ij})$ for $i \in \{CH, QS, AR\}$, modeled as distributions of possibility π_i [11] and represented by fuzzy images. Figure 1 presents an example of such fuzzy maps, processed on a 19.5 nm EIT image taken on August 3, 2000.

To this original algorithm, we added [3] some pre and post processings (temporal stability, limb correction, edge smoothing, optimal clustering based on a sursegmentation), which dramatically improved the results.
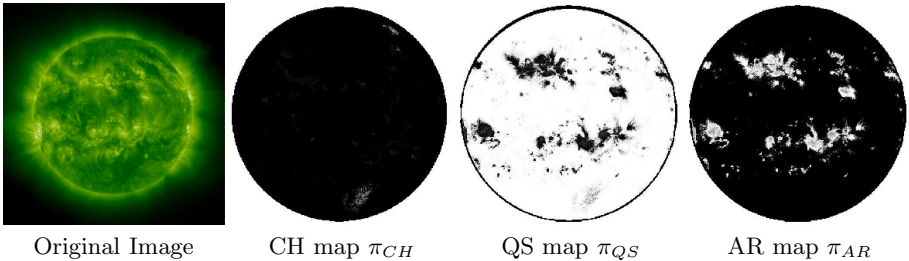

| Original Image | CH map π_{CH} | QS map π_{QS} | AR map π_{AR} |

Fig. 1. Fuzzy segmentation of a 19.5 nm EIT image taken on August 3, 2000

Segmentation of Solar Features. From coronal holes (CH), Active regions (AR) and Quiet Sun (QS) fuzzy maps, solar features can then be segmented using both memberships and expert knowledge provided by solar physicists. The basic principle is to find connected components in a fuzzy map being homogeneous with respect to some statistical criteria, related to the physical properties of the features, and/or having some predefined geometrical properties. Some region growing techniques and mathematical morphology are thus used here to achieve this segmentation process. Typical solar features that can directly be extracted from EIT images only include coronal bright points (figure 2(a)) or active regions (figure 2(b)).
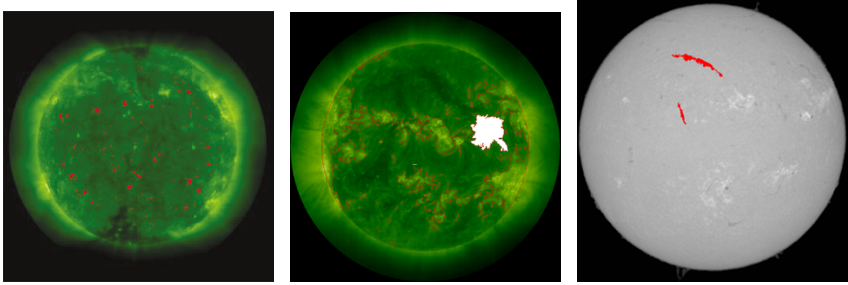
(a) Bright points from EIT image (1998-02-03)   (b) Active regions from EIT image (2000-08-04)   (c) Filaments from H-α image

Fig. 2. Several solar features

Additional information can also be added to these maps to allow the segmentation of other solar features. We for example processed in [3] the segmentation of filaments from the fusion of EIT and H-α images, from Kanzelhoehe observatory (figure 2(c)).

2.2   Tracking

In this article, we propose to illustrate the method on the automatic tracking of Active Regions. We more particularly focus on the largest active region, and algorithm 3 gives an overview of the method.

The center of mass G_{t-1} of AR_{t-1} is translated to G_t, such that the vector with start point $\boldsymbol{G_{t-1}G_t}$ equals the displacement field ν_G observed at pixel G_{t-1}. The displacement field between images I_{t-1} and I_t is estimated with the *opticalFlow* procedure, a multiresolution version of the differential Lucas and Kanade algorithm [15]. If $I(x, y, t)$ denote the gray-level of pixel (x, y) at date t, the method assumes the conservation of image intensities through time:

$$I(x, y, t) = I(x - u, y - v, 0)$$

where $\nu = (u, v)$ is the velocity vector. Under the hypothesis of small displacements, a Taylor expansion of this expression gives the gradient constraint equation:

Data: $(I_1 \cdots I_N)$ N EIT images
Result: Timeseries of parameters of the tracked AR

```
// Find the Largest connected component on the AR fuzzy map of I_1
```
AR_1=FindLargestCC(I_1^{AR})
```
// Compute the Center of mass of AR_1
```
G_1=ComputeCenterMass(AR_1)
for $t=2$ **to** N **do**
    ```// Compute the Optical flow between I_{t-1} and I_t```
    F_{t-1}=opticalFlow(I_{t-1}, I_t)
    ```// Compute the New center of mass, given the velocity field```
    $G_t = Forecast(G_{t-1}, F_{t-1})$
    ```// Find the Connected component in AR fuzzy map of I_t, centered```
    ```   on G_t```
    $AR_t = \text{FindCC}(G_t)$
```
// Timeseries analysis of regions AR_1 ··· AR_t
```
return Timeseries($AR_1 \cdots AR_N$)

Fig. 3. Active region tracking

$$\nabla I(x,y,t)^T \nu + \frac{\partial I}{\partial t}(x,y,t) = 0 \tag{3}$$

Equation (3) allows to compute the projection of ν in the direction of ∇I, and the other component of ν is found by regularizing the estimation of the vector field, through a weighted least squares fit of (3) to a constant model for ν in each of small spatial neighborhood Ω:

$$Min \sum_{(x,y)\in\Omega} W^2(x,y) \left[\nabla I(x,y,t)^T \nu + \frac{\partial I}{\partial t}(x,y,t) \right]^2 \tag{4}$$

where $W(x,y)$ denotes a window function that gives more influence to constraints at the center of the neighborhood than those at the surroundings. The solution of (4) is given by solving

$$A^T W^2 A \nu = A^T W^2 b$$

where for n points $(x_i, y_i) \in \Omega$ at time t

$$A = (\nabla I(x_1, y_1, t) \cdots \nabla I(x_n, y_n, t))^T$$
$$W = diag(W(x_1, y_1) \cdots W(x_n, y_n))$$
$$b = \left(-\frac{\partial I}{\partial t}(x_1, y_1, t) \cdots -\frac{\partial I}{\partial t}(x_n, y_n, t) \right)^T$$

A classical calculus of linear algebra directly gives $\nu = (A^T W^2 A)^{-1} A^T W^2 b$.

In this work, we applied a multiresolution version of this algorithm : the images were downsampled to a given lowest resolution, then the optical flow algorithm was computed for this resolution, and serves as an initialization for the computation of optical flow at the next resolution. This process was iteratively applied

until the initial resolution was reached. This allows a coarse-to-fine estimation of velocities. This procedure is simple and fast, and hence allows for a real-time tracking of AR.

Although we can suppose here that because of the slow motion between I_{t-1} and I_t, G_t will lie in the trace of AR_{t-1} in I_t (and thus a region growing technique may be sufficient, directly starting from G_t in I_t), we use the optical flow for handling non successive images I_t and $I_{t+j}, j >> 1$, but also for computing some velocity parameters of the active regions such as the magnitude, the phase, etc, and to allow the tracking of any solar feature, whatever its size (*cf.* section 3.3).

2.3   Quantifying Solar Features

Several quantitative indices can finally be computed on a given solar feature, given the previous segmentation. We investigate here both geometric and photometric (irradiance) indices for a solar feature S_t segmented from image I_t at time t:

- location L_t, given as as function of the latitude on the solar disc
- area $a_t = \int \int_{S_t} dx dy$,
- Integrated and mean intentities: $i_t = \int \int_{S_t} I(x, y, t) dx dy$ and $m(t) = i_t/a_t$
- fractal dimension, estimated using a box counting method

All of these numerical indices give relevant information on S_t, and more important, the analysis of the timeseries of these indices can reveal important facts on the birth, the evolution and the dead of solar features.

3   Results

3.1   Data

We apply our segmentation procedure on subsets of 1024×1024 EIT images taken from 14 February 1997 up till 30 April 2005, thus spanning more than 8 years of the 11-year solar cycle. During the 8 years period, there were two extended periods without data: from 25 June up to 12 October 1998, and during the whole month of January 1999. Almost each day during this period, EIT images taken with less than 30 min apart were considered. These images did not contain telemetry missing blocks, and were preprocessed using the standard *eit_prep* procedure of the *solar software (ssw)* library. Image intensities were moreover normalized by their median value.

3.2   First Example: Automatic Tracking of the Biggest Active Region

Active regions (AR) are areas on the Sun where magnetic fields emerge through the photosphere into the chromosphere and corona. Active regions are the source of intense solar flares and coronal mass ejections. Studying their birth, their

evolution and their impact on total solar irradiance is of great importance for several applications, such as space weather.

We illustrate our method with the tracking and the quantification of the largest AR of the solar disc, during the first 15 days of August, 2000. Figure 4 presents an example on a sequence of images, taken from 2000-08-01 to 2000-08-10. Active Regions segmented from SPoCA are highlighted with red edges, the biggest one being labeled in white. From this segmentation, we computed and plotted several quantitative indices, and we illustrate the timeseries of area, maximum intensity and fractal dimension over the period showed in figure 4.
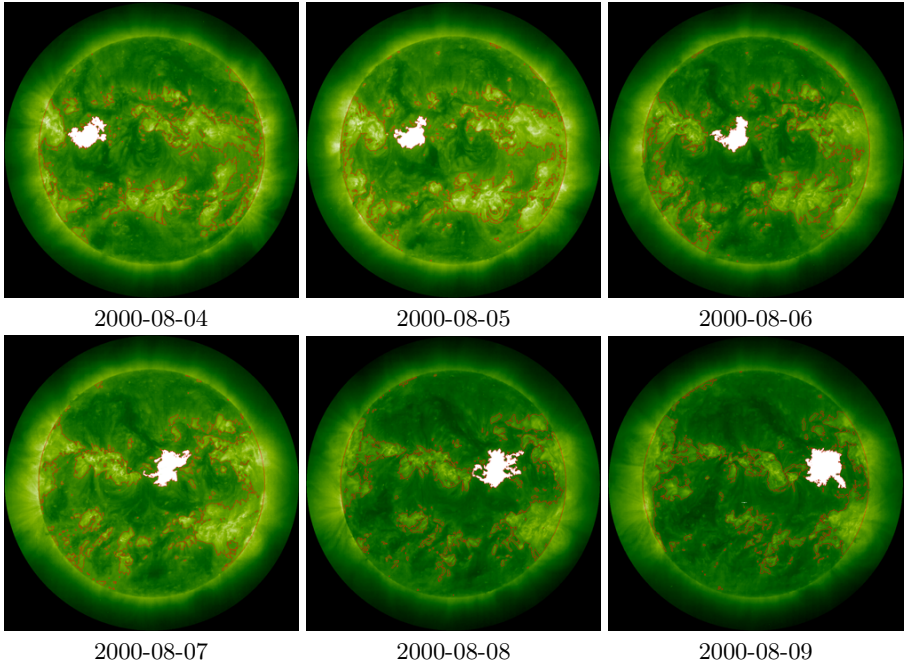

| 2000-08-04 | 2000-08-05 | 2000-08-06 |
| 2000-08-07 | 2000-08-08 | 2000-08-09 |

Fig. 4. Example of an AR tracking process. The tracking was performed on an active region detected on 2000-08-04, up to 2000-08-09.

area   maximum intensity   fractal dimension
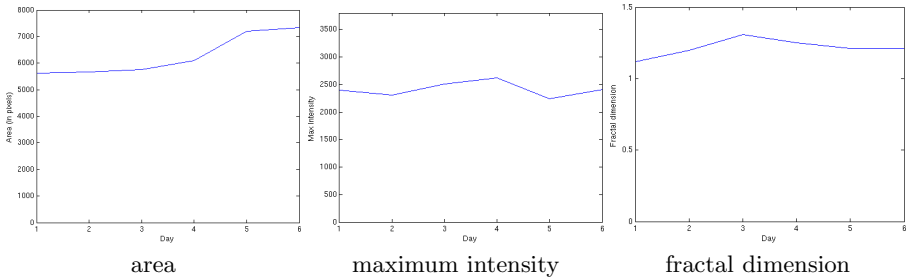
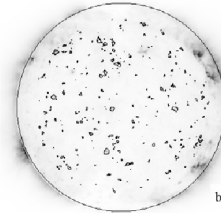**Fig. 5.** Example of AR quantification indices for the period 2000-08-04 - 2000-08-09

Such results demonstrate the ability of the method to track and quantify active regions. It is now important not only to track such a solar feature over a solar rotation period, but also to record its birth and capture its evolution through several solar rotations. For this, we now plan to characterized solar features with their vector of quantification indices, and to recognize new features appearing on the limb, among the set of solar feature already been registered, using an unsupervised pattern recognition algorithm.

3.3   Second Example: Distribution of Coronal Bright Points

Coronal Bright Points (CBP) are of great importance for the analysis of the structure and dynamics of solar corona. They are identified as small and short-lived ($<$ 2 days) coronal features with enhanced emission, mostly located in quiet-Sun regions and coronal holes. Figure 6 presents a segmentation of CBP of an image taken on February, 2nd, 1998. This image was chosen so as to compare the results with the one provided by [20] Several other indices can be computed from this analysis, such as N/S assymetry, timeseries of the number of CBP, intensity analysis of CBP...
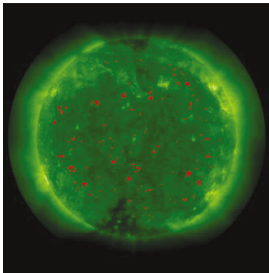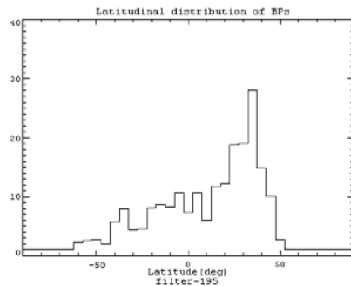

Sgmentation of CBP using 19.5 nm EIT image   CBP [20]
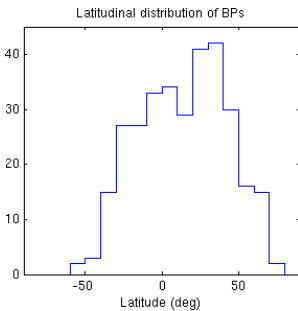
Number of CBP as a function of latitude   same from [20]

Fig. 6. Number of CBP as a function of latitude: comparison with [20]

4   Conclusion

We proposed in this article an image processing pipeline that segment, track and quantify solar features from a set of multispectral solar corona images, taken with eit EIT instrument. Based on a validated segmentation scheme, the method is fully described and illustrated on two preliminary studies: the automatic tracking of Active Regions from EIT images taken during solar cycle 23, and the analysis of spatial distribution of coronal bright points on the sular surface. The method is generic enough to allow the study of any solar feature, provided it can be segmented from EIT images or other sources. As stated above, our main perspective is to follow solar feature and to track their reappearance after a solar rotation S. We plan to use the quantification indices computed on a given solar feature F to characterize it and to find, over new solar features appearing on the solar limb at time $t + S/2$, the one closest to F. We also intend to implement a multiple activity region tracking, using a natural extension of our method.

References

1. Aboudarham, J., Scholl, I., Fuller, N.: Automatic detection and tracking of filaments for a solar feature database. Annales Geophysicae 26, 243–248 (2008)
2. Barra, V., Delouille, V., Hochedez, J.F.: Segmentation of extreme ultraviolet solar images via multichannel Fuzzy Clustering Algorithm. Adv. Space Res. 42, 917–925 (2008)
3. Barra, V., Delouille, V., Hochedez, J.F.: Fast and robust segmentation of solar EUV images: algorithm and results for solar cycle 23. A&A (submitted)
4. Benkhalil, A., Zharkova, V., Zharkov, S., Ipson, S.: Proceedings of the AISB 2003 Symposium on Biologically-inspired Machine Vision, Theory and Application, ed. S. L. N. in Computer Science, pp. 66–73 (2003)
5. Berrili, F., Moro, D.D., Russo, S.: Spatial clustering of photospheric structures. The Astrophysical Journal 632, 677–683 (2005)
6. Bezdek, J.C., Hall, L.O., Clark, M., Goldof, D., Clarke, L.P.: Medical image analysis with fuzzy models. Stat. Methods Med. Res. 6, 191–214 (1997)
7. Bornmann, P., Winkelman, D., Kohl, T.: Automated solar image processing for flare forecasting. In: Proc. of the solar terrestrial predictions workshop, Hitachi, Japan, pp. 23–27 (1996)
8. Brajsa, R., Whöl, H., Vrsnak, B., Ruzdjak, V., Clette, F., Hochedez, J.F.: Solar differential rotation determined by tracing coronal bright points in SOHO-EIT images. Astronomy and Astrophysics 374, 309–315 (2001)
9. Brajsa, R., Wöhl, H., Vrsnak, B., Ruzdjak, V., Clette, F., Hochedez, J.F., Verbanac, G., Temmer, M.: Spatial Distribution and North South Asymmetry of Coronal Bright Points from Mid-1998 to Mid-1999. Solar Physics 231, 29–44 (2005)
10. Delaboudinière, J.P., Artzner, G.E., Brunaud, J., et al.: EIT: Extreme-Ultraviolet Imaging Telescope for the SOHO Mission. Solar Physics 162, 291–312 (1995)
11. Dubois, D., Prade, H.: Possibility theory, an approach to the computerized processing of uncertainty. Plenum Press (1985)
12. Fuller, N., Aboudarham, J., Bentley, R.D.: Filament Recognition and Image Cleaning on Meudon Hα Spectroheliograms. Solar Physics 227, 61–75 (2005)

13. Hill, M., Castelli, V., Chu-Sheng, L.: Solarspire: querying temporal solar imagery by content. In: Proc. of the IEEE International Conference on Image Processing, pp. 834–837 (2001)
14. Krishnapuram, R., Keller, J.M.: A possibilistic approach to clustering. IEEE Trans. Fuzzy Sys. 1, 98–110 (1993)
15. Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereovision. In: Proc. Imaging Understanding Workshop, pp. 121–130 (1981)
16. Nieniewski, M.: Segmentation of extreme ultraviolet (SOHO) sun images by means of watershed and region growing. In: Wilson, A. (ed.) Proc. of the SOHO 11 Symposium on From Solar Min to Max: Half a Solar Cycle with SOHO, Noordwijk, pp. 323–326 (2002)
17. Ortiz, A.: Solar cycle evolution of the contrast of small photospheric magnetic elements. Advances in Space Research 35, 350–360 (2005)
18. Pettauer, T., Brandt, P.: On novel methods to determine areas of sunspots from photoheliograms. Solar Physics 175, 197–203 (1997)
19. Qahwaji, R.: The Detection of Filaments in Solar Images. In: Proc. of the Solar Image Recognition Workshop, ed. Brussels, Belgium (2003)
20. Sattarov, I., Pevtsov, A., Karachek, N.: Proc. of the International Astronomical Union, pp. 665–666. Cambridge University Press, Cambridge (2004)
21. Sobotka, M., Brandt, P.N., Simon, G.W.: Fine structures in sunspots. I. Sizes and lifetimes of umbral dots. Astronomy and astrophysics 2, 682–688 (1997)
22. Steinegger, M., Bonet, J., Vazquez, M.: Simulation of seeing influences on the photometric determination of sunspot areas. Solar Physics 171, 303–330 (1997)
23. Steinegger, M., Bonet, J., Vazquez, M., Jimenez, A.: On the intensity thresholds of the network and plage regions. Solar Physics 177, 279–286 (1998)
24. Veronig, A., Steinegger, M., Otruba, W.: Automatic Image Segmentation and Feature Detection in solar Full-Disk Images. In: Wilson, N.E.P.D.A. (ed.) Proc. of the 1st Solar and Space Weather Euroconference, Noordwijk, p. 455 (2000)
25. Wagstaff, K., Rust, D.M., LaBonte, B.J., Bernasconi, P.N.: Automated Detection and Characterization of Solar Filaments and Sigmoids. In: Proc. of the Solar image recognition workshop, ed. Brussels, Belgium (2003)
26. Worden, J., Woods, T., Neupert, W., Delaboudiniere, J.: Evolution of Chromospheric Structures: How Chromospheric Structures Contribute to the Solar He ii 30.4 Nanometer Irradiance and Variability. The Astrophysical Journal, 965–975 (1999)
27. Zharkov, S., Zharkova, V., Ipson, S., Benkhalil, A.: Automated Recognition of Sunspots on the SOHO/MDI White Light Solar Images. In: Negoita, M.G., Howlett, R.J., Jain, L.C. (eds.) KES 2004. LNCS, vol. 3215, pp. 446–452. Springer, Heidelberg (2004)