Automatic detection process of solar active region based on SDO/AIA digital image

I Sukarsih¹, A Hermawanti¹, Respita wulan¹, R Priyatikanto², E Kurniati¹

¹ Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Islam Bandung, Jl. Ranggamalela 1 Bandung, Indonesia 40116
² Space Science Centre, National Institute of Aeronautics and Space, Jl.Dr. Djunjunan 133 Bandung, Indonesia 40173

E-mail: respitawulan@unisba.ac.id

Abstract. The aim of this study is to do automatic detection of the sunspot on the solar disc using MATLAB. The digital images from the satellite SDO (Solar Dynamic Observatory)/AIA (Atmospheric Imaging Assembly) with 512 x 512 pixels jpg image were being processed through several steps: segmentation threshold to separate the object from its background and obtain the area and coordinate points (centroids) of the sunspots. Then, the sunspot coordinates gained from the automatic detection are compared with the coordinates released by NASA at solarmonitor.org. The error calculation using RMSE gives 30.824022" to 1120.2831" for horizontal position and 46.509763" to 266.74988" for vertical position. The errors are still considered significant since Lembaga Penerbangan dan Antariksa Nasional (LAPAN, Indonesian National Institute of Aeronautics and Space) specify the error tolerance value of 20" maximum. The main factor for the difference is the spatial characteristics of the EUV images used for automatic detection are different with the characteristics of the photosphere image used by solarmonitor.org as the catalogue reference of the sunspots.

1. Introduction

The phenomena happened on the sun called the sun activities consist of sunspot, prominence, flare, mass hurl (CME), and corona hole. One of the activities that the development can be observed is the sun’s active area or the sunspot. Due to its link to other kinds of solar activity, sunspot occurrence can be used to help predict space weather [1], the state of the ionosphere [3],[4], and hence the conditions of short-wave radio propagation or satellite communications [5]-[7].

The forming of sunspots can be spotted by identifying the sun’s active region. The active region can be seen from the satellite images of the sun. One of the satellites for sun observation is the SDO (Solar Dynamic Observatory)/ AIA (Atmospheric Imaging Assembly) which outputs digital images data. NASA had provided the sunspot identification on the website solarmonitor.org using multiple types of image wavelength. There are also several methods for the active region detections, such as fractal-based fuzzy technique [8], thresholding and region growing technique [9], or even using the solar interior properties [10]. Here, we would like to investigate if it possible to detect the solar active region by using only the corona image of the sun.

MATLAB is a software that is often used for the purposes of computation and simulations [11]-[13]. MATLAB also provides an easy tool for image processing as a user can easily access each and every
pixel value from the image matrices and edit it [14]. Moreover, there is an ‘image processing tool box’ [15] built in MATLAB for this purpose. Because of that, automatic detection of sunspots on solar disc will be done based on the digital image from SDO/AIA using MATLAB. Furthermore, the purpose of this research is to predict the forming of sunspots on solar disc by using automatic detection based on SDO/AIA.

2. Methods

2.1. SDO/AIA satellite image
The Solar Dynamic Observatory (SDO) is a mission by NASA to observe the sun since the year 2010. SDO was launched to space on February 11th 2010. This observatory is a part of the Living with a Star (LWS) program. SDO is a stable spacecraft with a third axis that has two configurations and two antennas with high gain, and has a geosynchronous orbit that goes round the Earth. The purpose of the SDO is to understand the effect of the Sun on the Earth by studying activities that occurs on the Sun on its atmosphere on a small spatial and temporal scale and also on the wavelength simultaneously [16].

AIA (Atmospheric Imaging Assembly) can provide a clearer view on the phenomena that occurs on the outer atmosphere of the Sun that is still developing. The main purpose of AIA is for scientific research. The usage of data along with SDO instruments’ data and from other observations significantly improves understanding of the activities that was shown at the atmosphere of the Sun also occurs on the heliosphere and around the planet.

Images of solar activities can be accessed daily through the solarmonitor.org website with various formats and image sizes. This research uses an image of the surface of the Sun taken on August 30th to September 3rd, 2017.

2.2. Threshold segmentation

Image segmentation is a process to divide an image into segments or non-overlapping areas. In digital image terms, resulting segmentation areas are defined by groups of neighboring or connected pixels. The segmentation process is used in many applications, and even though the methods used can be varied, they all support the same purpose which is to obtain a simple and useful representation from an image [17].

Thresholding is one of the segmentation techniques that commonly used for images with significantly different intensity values between the background and the main object. In its application, the thresholding method requires a value that is used as a limiting value between the background and the main object, and that value is called the threshold. Thresholding is used to partitioning the image by adjusting the intensity value of all the pixels that are bigger than the value of threshold T as a foreground and pixels that are smaller than the value of threshold T as the background, with T being a floating value.

Generally, the thresholding process for a gray scale image is done to generate a binary image [18], which can be written mathematically below:

\[ g(x,y) = \begin{cases} 
1 & \text{if } f(x,y) \geq T \\
0 & \text{if } f(x,y) < T 
\end{cases} \]

With \( g(x,y) \) being a binary image from the gray scale image \( f(x,y) \), and T being the threshold value.

2.3. MATLAB regionprops function

In MATLAB, the regionprops function is an object that is represented as a region with a rectangular approach. Figure 1 shows a region that consists of white pixels represented with a rectangular approach [17].
3. Result and Discussion
Results of the automatic detection of sunspots on this research is based on the digital image from SDO/AIA taken from August 30th till September 3rd 2017. Steps that were taken can be seen on the flowchart on Figure 2:

![Flowchart](image)

**Figure 1.** Region representation with rectangle approach

**Figure 2.** Automatic Detection Process of Solar Disc Active Region

3.1. Image processing process
The data used in this research is an image in JPG format with the temporal resolution of 15 minutes and image resolution of 512 x 512 pixels that was taken from NASA SDO website [19]. The images we used are the colored image which has three color components of RGB (red, green, and blue). In MATLAB, the image will be represented in a three-dimensional matrix with the size of $512 \times 512 \times 3$. Hence, transformations of the colored image to a gray scale image was done to simplify the image to a gray level between 0 (black) to 255 (white), so the representation of the image will be a regular matrix of $512 \times 512$.

After having the result of gray scale conversion, next is the thresholding operation. The thresholding process divides gray levels into two, black and white, so that the main object and the background of the image can be identified. An image with brightness level higher than the threshold point will be given the value of 1 (with black background), while images with brightness level below the threshold point
will be given the value of 0 (object, with white color). This was done to make the bright area indicating the active region stand out from the background.

Next is the process to locate the area and coordinates of the sunspots that is derived from the centroid value using the regionprops function in MATLAB. The area has a scalar value, while the centroid has values in the form of coordinates. The values of the area and the centroid of sunspots for the August 30th, 2017 can be seen on the Table 1.

Table 1. Detected Area and Centroid (August 30th, 2017)

| Detected Area | Centroid coordinates | S (area) |
|---------------|-----------------------|----------|
| A             | 77.220                | 194.207  | 3420     |
| B             | 90.207                | 314.052  | 1647     |
| C             | 377.091               | 96.249   | 241      |
| D             | 430.169               | 231.956  | 5406     |

3.2. Comparison of the sunspots detection with NASA

To determine whether the result of the automatic detection matches the sunspots that was provided by NASA, a comparison was done between the resulting image from the automatic detection and the image provided by NASA on solarmonitor.org. comparison was done by matching the centroid coordinates from the automatic detection and coordinates from NASA.

![Automatic detection result](a) ![Result from solarmonitor.org](b)

Figure 3. Automatic detection result (a) and result from solarmonitor.org (b)

From Figure 3, the sunspot on area 12672 (solarmonitor.org) matches with area D (detection), with the coordinates for each are (835", 43") and (430.1696, 231.9558). Next, the comparison was done on the sunspot on area 12673 which was on the coordinate (−722", −207") for solarmonitor.org and area B which was on the coordinate (90.2077, 314.0516) for the detection method. Similar method was done for the remaining area and centroid. The result can be seen on table 2.

Table 2. Coordinates comparison of active region at August 30th, 2017

| Date         | NOAA (area) | solarmonitor.org | Detection Result Area | Detection |
|--------------|-------------|------------------|-----------------------|-----------|
| 2017/08/30   | 12672       | 835              | 43                    | D         |
| 2017/08/30   | 12673       | -722             | -207                  | B         |
| 2017/08/30   | 12674       | -792             | 118                   | A         |

3.3. Coordinate validation using RMSE

Because the data of the sunspots that were the result of automatic detection is in pixels while the data received from solarmonitor.org was using arc second units, transformations need to be done before the validation process.
Other than the difference in units, the scale and origin point \((0,0)\) in the cartesian coordinate of the centroid of the two images also compared. Total width of the image from solarmonitor.org is 2460” and the origin point is in the middle, while the total width of the analysed image is 512 pixels with the origin point at the top left. Thus, the coordinates’ transformation was done using these functions.

\[
x' = 4.805x - 1230 \\
y' = 1230 - 4.850x.
\]

Results of the transformation can be seen on Table 3.

### Table 3. Coordinates after Transformation

| Date      | NOAA  | NOAA Coordinates | Detection Coordinates | Transformation Coordinates |
|-----------|-------|------------------|------------------------|---------------------------|
|           | NOA   | x* y*             | x y                    | x' y'                      |
| 2017/08/30| 1267  | 835 43            | 430.17 231.956         | 856.323 105.014            |
| 2017/08/30| 12673 | -722 -207         | 90.208 314.052         | -792.493 -293.15           |
| 2017/08/30| 12674 | -792 118          | 77.220 194.207         | -855.482 288.095           |

To increase accuracy, errors were calculated between the two centroids using the RMSE. Results of RMSE of the digital image comparisons on August 30th, 2017 can be seen on table 4. The same process was done on other images from the date August 31st to September 3rd, 2017. The RMSE gained from the dates are shown on Table 4.

### Table 4. RMSE Validation

| residual x0 | residual y0 | square residual x0 | square residual y0 | RMSE x0 | RMSE y0 | Date | RMSE x0 | RMSE y0 |
|-------------|-------------|--------------------|--------------------|---------|---------|------|---------|---------|
| -21.323     | -62.014     | 454.652            | 3845.782           | 56.1365 | 115.758 | 2017/08/30 | 56.136  | 115.758 |
| 70.493      | 86.150      | 4969.214           | 7421.867           | 1120.283| 46.509  | 2017/08/31 | 1120.283 | 46.509  |
| 63.482      | -170.095    | 4029.968           | 2893.171           | 30.824  | 97.924  | 2017/09/01 | 30.824  | 97.924  |
|             |             |                    |                    |         |         | 2017/09/02 | 604.597 | 266.749 |
|             |             |                    |                    |         |         | 2017/09/03 | 181.781 | 116.333 |

The RMSE values were ranged between 30 to over 1000 arc second. This was far above the ideal research target that was determined by Lembaga Penerbangan dan Antariksa Nasional (LAPAN, Indonesian National Institute of Aeronautics and Space), which is maximum 20”. Results of this research can be affected by factors stated below:

- Spatial characteristics of the EUV image (that was detected) differs from the characteristics of the photosphere image (from solarmonitor.org) that became the reference catalogue of the sunspots.
- The image resolutions were different between the one used for detection (512 x 512 pixels) and the one taken from solarmonitor.org (1230 x 1230 pixels)
- The analysed EUV image had a different epoch. Sunspots may change in position during the rotation of the Sun (13° to 130° per day)

### 4. Conclusion

Automatic detection of sunspots was done in MATLAB with images that was taken from SDO/AIA satelites. The image that was taken was a solar ultraviolet (EUV) image in JPG format with a resolution of 512 x 512 pixels. Steps that were taken including threshold segmentation and regionprops. Images acquired from the automatic detection process were compared with the results provided by NASA in the solarmonitor.org website. Detection results were transformed so that the data would have the same unit of measurement as solarmonitor.org. Error calculation using RMSE resulting in values between 30.824” to 1120.283” for the horizontal position, and 46.509” to 266.749” for the vertical position. This value was considered significant as LAPAN had determined tolerance levels to be below 20”. Because of that, the validation for further research should be done by processing digital image that occurs on the same time period and have the same pixel size as the image provided by NASA.
**Acknowledgment**
This research was supported and funded by Research Grant 2019/2020, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Islam Bandung, Indonesia.

**References**

[1] Thomas J H and Weiss N O 2012 Sunspots and starspots *Sustainability*

[2] Storini M, Bazilevskaya G A, Fluckiger E O, Krainev M B, Makhmutov V S and Sladkova A I 2003 The Gnevyshev gap: a review for space weather *Advances in Space Research* **31**(4) pp 895-900

[3] Geisler J E and Bowhill S A 1965 Ionospheric temperatures at sunspot minimum *J. Atmospheric and Terrestrial Physics* **27**(4) pp 457-474

[4] Nishida A 1967 Average structure and storm-time change of the polar topside ionosphere at sunspot minimum *J. Geophysical Research* **72**(23) pp 6051-6061

[5] Doherty P, Coster A J and Murtagh W 2004 Space weather effects of October–November 2003 *GPS Solutions* **8**(4) pp 267-271

[6] Basu S, MacKenzie E and Basu S 1988 Ionospheric constraints on VHF/UHF communications links during solar maximum and minimum periods *Radio Science* **23**(03) pp 363-378

[7] Davies K and Smith E K 2002 Ionospheric effects on satellite land mobile systems *IEEE Antennas and Propagation Magazine* **44**(6) pp 24-31

[8] Revathy K, Lekshmi S and Nayar SP 2005 Fractal-based fuzzy technique for detection of active regions from solar images *Solar Physics* **228**(1-2) pp 43-53

[9] Benkhali A, Zharkova V, Ipson S and Zharkov S 2003 Automatic identification of active regions (Plages) in the full-disk solar images using local thresholding and region growing techniques *Proc. of the AISB* vol 3 pp 66-73

[10] Ilonidis S, Zhao J and Kosovichev A 2011 Detection of emerging sunspot regions in the solar interior *Science* **333**(6045) pp 993-996

[11] Grigoriu M and Harper E 1995 *Applied non-Gaussian processes: Examples, theory, simulation, linear random vibration, and MATLAB solutions* (Upper Saddle River, NJ: PTR Prentice Hall)

[12] Daowd M, Omar N, Van Den Bossche P and Van Mierlo J 2011 Passive and active battery balancing comparison based on MATLAB simulation *2011 IEEE Vehicle Power and Propulsion Conf.* pp 1-7

[13] Harahap E, Sukarsoh I, Gunawan G, Fajar M Y, Darmawan D and Nishi H 2016 A Model-based simulator for content delivery network using Simevents MATLAB-Simulink. *INSIST* **1**(1) pp 30-33

[14] http://www.mathworks.in/help/images/index.html

[15] Buksh R, Routh S, Mitra P, Banik S, Mallik A and Gupta S D 2014. MATLAB based image editing and color detection *International Journal of Scientific and Research Publications* **4**(1) pp 1-6

[16] Lemen J R, Akin D J, Boerner P F, Chou C, Drake J F, Duncan D W and Katz N L 2011 The atmospheric imaging assembly (AIA) on the solar dynamics observatory (SDO) *The Solar Dynamics Observatory* pp. 17-40 (New York: Springer)

[17] Ardhianto E, Hadikurniawati W and Budiarsa Z 2013 Implementasi metode image subtracting dan metode regionprops untuk mendeteksi jumlah objek berwarna RGB pada file video. *Dinamik* **18**(2)

[18] Purba D 2010 *Pengolahan Citra Digital* (Yogyakarta: Andi)

[19] Pesnell D and Addison K 2017 *SDO Data* Retrieved from URL https://sdo.gsfc.nasa.gov/data/aiahmi/