# An Automatic Segmentation Algorithm for Solar Filaments in H-Alpha Images using a Context-based Sliding Window

Ibrahim A. Atoum

College of Applied Sciences
Al Maarefa Colleges for Science and Technology
Riyadh, Saudi Arabia

*Abstract*—There are many features which appear on the surface of the sun. One of these features that appear clearly are the dark threads in the Hydrogen alpha (Hα) spectrum solar images. These 'filaments' are found to have a definite correlation with Coronal Mass Ejections (CMEs). A CME is a large release of plasma into space. It can be hazardous to astronauts and the spacecraft if it is being ejected towards the Earth. Knowing the exact attributes of solar filaments may open the way towards predicting the occurrence of CMEs. In this paper, an efficient and fully automated algorithm for solar filament segmentation without compromising accuracy is proposed. The algorithm uses some statistical measures to design the thresholding equations and it is written in the C++ programming language. The square root of the range as a measure of variability of image intensity values is used to determine the size of the sliding window at run time. There are many previous studies in this area, but no single segmentation method that could precisely claim to be fully automatic exists. Samples were taken from several representative regions in low-contrast and high-contrast solar images to verify the viability and efficacy of the method.

*Keywords—Solar image processing; solar filament; segmentation; sliding window; Coronal mass ejections*

## I. INTRODUCTION

Solar filaments are huge regions of very thick relatively cool plasma compared to the surface of the sun. These solar features appear as elongated and dark filaments in Hα solar images. CMEs are enormous bubbles of hot plasma (billions of tonnes of magnetized plasma) that is propagating away from the solar corona to the interplanetary medium at a very high velocity [1]. The importance of studying solar filaments comes from considering its disappearances as a significant indicator for the possible occurrence of CMEs, which is considered as the major cause of geomagnetic storms. It is now almost certain that there is a close association between CME and filament disappearances ([1]-[12]). Most of the filament detection techniques depend on sliding windows over the image region of interest. This sliding window is the number of pixels that are shifted while scanning the image. The scanning speed of images greatly depends on the size of the sliding window and the computational power of the image processing system. A context-based sliding window is proposed in this study to fully automate the process of solar filament segmentation in Hα Images. The use of often user selectable, non-adaptable and non-automatic different sizes of

structuring elements in filament detection techniques led to them being defined as non-automated. According to this hypothesis, no single filament segmentation method thus could accurately claim to be fully automatic. Additionally, most of the previous studies ([1], [5], [7], [8], [13]-[20]) follow an image scanning process using a fixed size sliding window according to a fixed step size. Also the sliding window size affects the results of the different filtering techniques like the median filter [21]. The paper is organized as follows. The concept of the context-sensitive sliding window is presented in the next section. Section III overviews the algorithm of the context-based sliding window. Section Four is the conclusion.

Fig. 1 shows the unevenly illuminated solar images. These images were enhanced using an image enhancement technique introduced by [16].
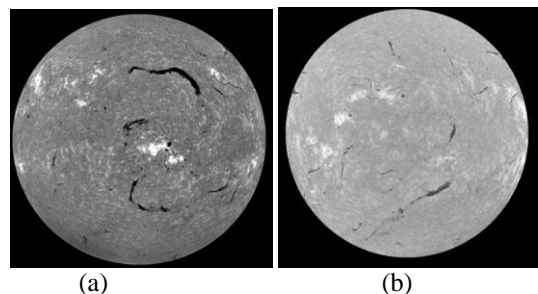


Fig. 1. (a) Solar image observed by Big Bear Solar Observatory (BBSO) on $2^{nd}$ January 2001; (b) Solar image observed by Meudon Observatory on $1^{st}$ April 2002.

## II. CONTEXT-BASED SLIDING WINDOW

The common purpose of a sliding window is to give the possibility of checking the presence of the object in a rectangular box at different positions in the image. Subsequently, it is possible to use any of the many proposed methods to check the presence of an object in the sliding window or not. All previous methods of segmenting solar filaments rely entirely on customizing the statistical measures using a fixed sized sliding window. In general, the values of the statistical parameters in a local window are affected by the size of it. This research strongly proffers to adapt the size of the sliding window according to the window contents by using the Range ($R$) as a measure of variability, that is, the

difference between the lowest and highest image intensity values contained in the sliding window. To ascertain the effectiveness of using $R$ and to optimize the image segmentation method, two samples; the first window ($W_1$) and the second window ($W_2$) were taken from two solar images as shown in the figures.

## III. HOMOGENEITY AND HETEROGENEITY OF IMAGE REGIONS

To give an adequate opportunity for segmenting solar images and hence judge between the homogeneous and heterogeneous regions, four values were calculated from different representative image samples, these being the value of $R$, the window side length; the window size ($W_{SIZE}$) as the integer square root of $R$; the mean value ($W_\mu$) and the standard deviation ($W_\sigma$). These values are shown in Fig. 2 and 3. Samples were taken to represent from non-filament regions (in order to represent them) as shown in Fig. 2(a) and Fig. 3(a). Other samples were taken from filament regions. Finally, the last samples were taken to represent the filament edges as shown in Fig. 2(c) and Fig. 3(c). Fig. 4 displays the same parameters for the low contrast image filaments.
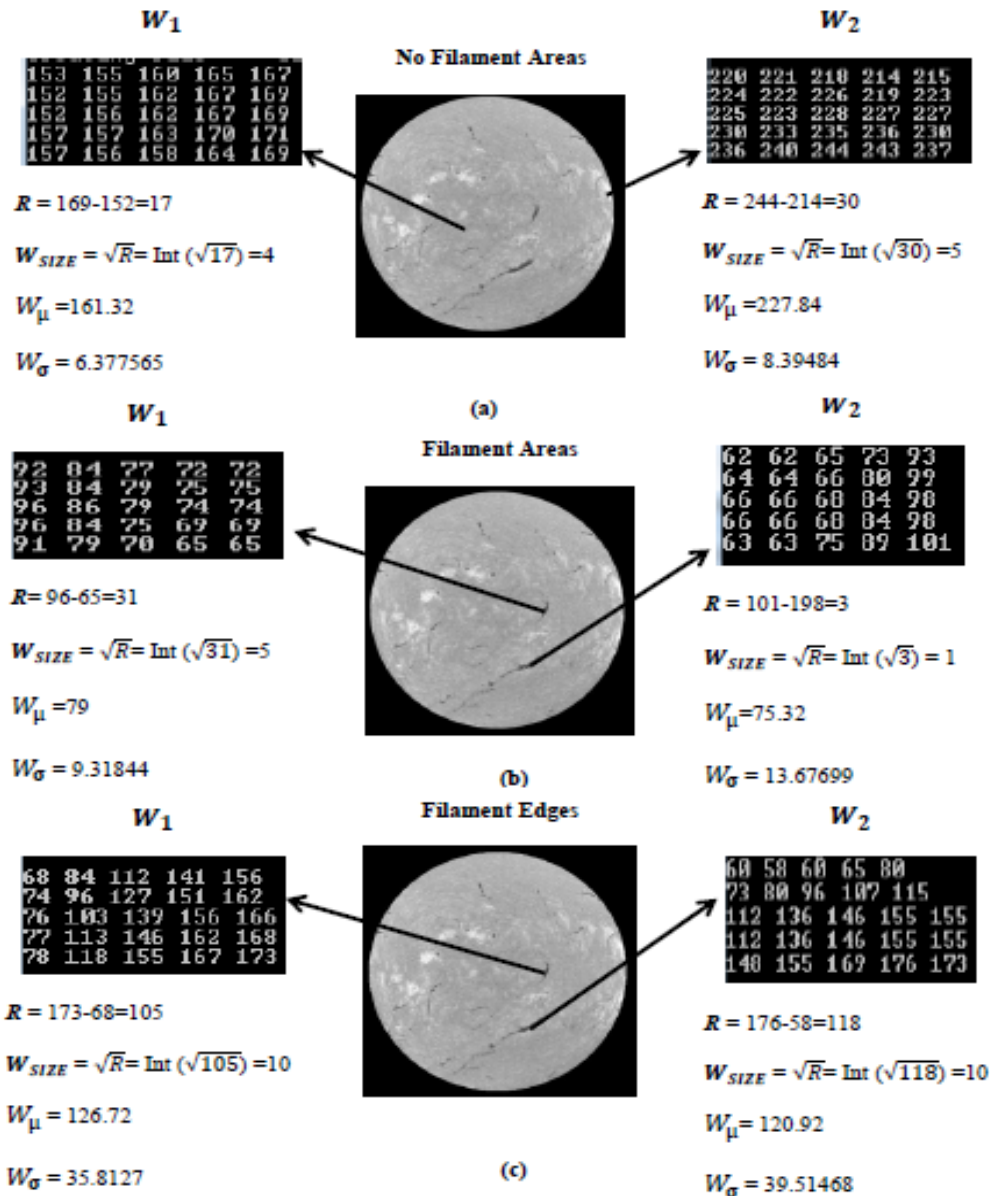


Fig. 2. Samples taken from solar image observed at Meudon Observatory on 1st April 2002: (a) Two samples from no filament pixels. (b) Two samples from filament pixels. (c) Two samples from Filament Edges.
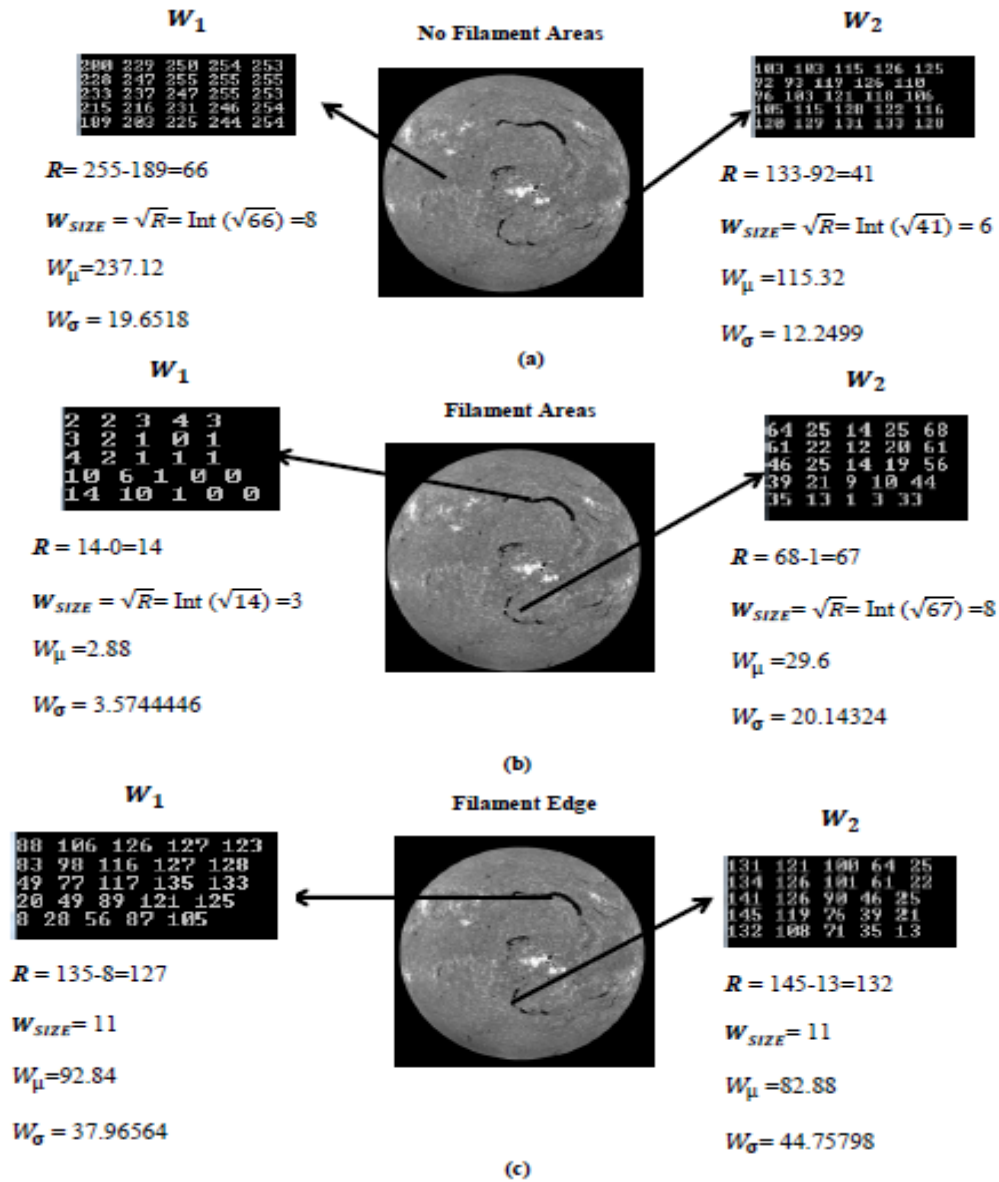
$W_1$ — No Filament Areas — $W_2$

$R = 255 - 189 = 66$

$W_{SIZE} = \sqrt{R} = \text{Int} (\sqrt{66}) = 8$

$W_\mu = 237.12$

$W_\sigma = 19.6518$

$R = 133 - 92 = 41$

$W_{SIZE} = \sqrt{R} = \text{Int} (\sqrt{41}) = 6$

$W_\mu = 115.32$

$W_\sigma = 12.2499$

(a)

$W_1$ — Filament Areas — $W_2$

$R = 14 - 0 = 14$

$W_{SIZE} = \sqrt{R} = \text{Int} (\sqrt{14}) = 3$

$W_\mu = 2.88$

$W_\sigma = 3.5744446$

$R = 68 - 1 = 67$

$W_{SIZE} = \sqrt{R} = \text{Int} (\sqrt{67}) = 8$

$W_\mu = 29.6$

$W_\sigma = 20.14324$

(b)

$W_1$ — Filament Edge — $W_2$

$R = 135 - 8 = 127$

$W_{SIZE} = 11$

$W_\mu = 92.84$

$W_\sigma = 37.96564$

$R = 145 - 13 = 132$

$W_{SIZE} = 11$

$W_\mu = 82.88$

$W_\sigma = 44.75798$

(c)

Fig. 3. Samples taken from solar image observed at Big Bear Solar Observatory 2$^{nd}$ January 2001: (a) Two samples from no-filament pixels. (b) Two samples from filament pixels. (c) Two samples from Filament Edges.
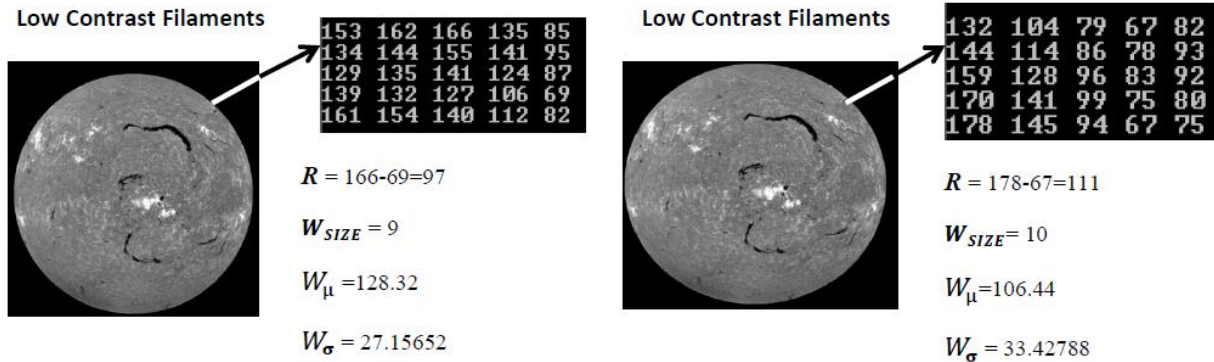
Low Contrast Filaments

$R = 166 - 69 = 97$

$W_{SIZE} = 9$

$W_\mu = 128.32$

$W_\sigma = 27.15652$

Low Contrast Filaments

$R = 178 - 67 = 111$

$W_{SIZE} = 10$

$W_\mu = 106.44$

$W_\sigma = 33.42788$

Fig. 4. Two samples from Low Contrast Filaments.

Where,

$W_\mu$: the average for the window which is computed as $W_\mu = \frac{\sum_{i=1}^{m} W_i}{m}$ , where $W_i$ is the set of pixels inside the window and $m$ being the number of pixels inside the window.

$W_\sigma$: the window standard deviation which is computed as $W_\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(W_i - W_\mu)}$, where $W_i$ are each pixel value in the window and $n$ being the number of pixels inside the window.

$R$: the window image intensity range value which is computed as $R = W_{max} - W_{min}$, where $W_{max}$ and $W_{min}$ are the maximum and minimum pixel image intensity values inside the window.

It can be observed from the previous values that the range values become high only in the regions with heterogeneous and non-convergent intensity values, while it will be low in the regions with relatively uniform (low spread) and convergent intensity values.

## IV. ADAPTIVE FILAMENT SEGMENTATION

The process of image segmentation is the first and most important phase in analyzing and interpreting these images. The success or failure of the segmentation process may be considered as the success or failure of the subsequent image classes. There are many features that appear on Hα images but our goal through this algorithm is to separate the image into white foreground filaments and the black background.

Fig. 5 shows the results after applying the segmentation code as shown in Fig. 7 on the solar image observed by Big Bear Solar Observatory (BBSO) on 2nd January 2001. Fig. 6 shows the results after applying the segmentation code on the solar image observed by Meudon Observatory (MO) on 1st April 2002.
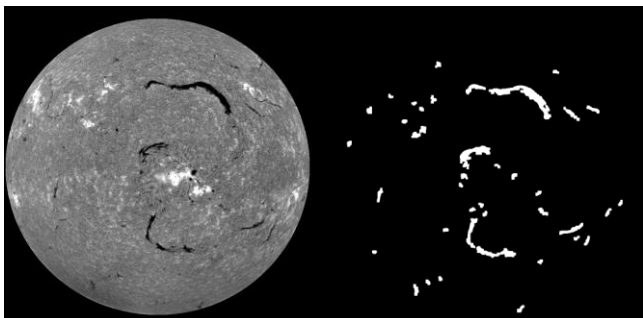


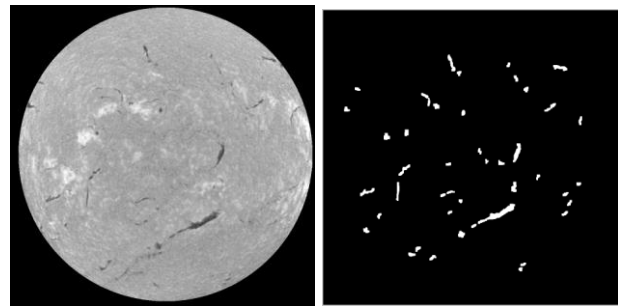Fig. 5.   The output after applying the algorithm on the BBSO image.



Fig. 6.   The output after applying the algorithm on the MO image.

Setting the window size to any initial value

> *For i=1 to m do // m is the number of image rows*
> *For j=1 to n do // n is the number of image columns*
> Calculate $I_\mu$ and $I_\sigma$
> *For n=1 to W do // W: window rows*
> *For m=1 to W do*
> Calculate $W_\mu$, $W_R$ and $W_\sigma$
> *If ((($I_\mu > (W_\mu + I_\sigma)$)) OR ($W_\mu < I_\sigma$)) then*
> *Filaments*
> *Else*
> *No Filaments*
> *End if*
> *End for m*
> *End for n*
> *End for j*
> *End for i*

Fig. 7.   The segmentation pseudo code.

## V. EVALUATION AND RESULTS

The size of the sliding window will change automatically after calculating the range in the first iteration based on the sliding window intensity values. The highest value that may exist in any sliding window is 255 (white) and the lowest one is 0 (black), which means that the maximum possible can only be, 255 - 0 = 255. A low window size indicates a homogeneous region (such as no-filament or filament region) and a high value indicates a non-homogenous region (such as the regions of filament edges). In *Hα* solar images; the filaments appear as dark features. This means if the sliding window is around the edges points of the filament pixels then the range value will be high otherwise, it will be low. The values calculated for each sample in all the regions are summarized in Table I.

TABLE I. STATISTICAL DETAILS OF THE SOLAR IMAGE AND THE SLIDING WINDOW

| | | R | | W$_{SIZE}$ | | Wμ | | W$_\sigma$ | | Iμ | I$_\sigma$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $W_1$ | $W_2$ | $W_1$ | $W_2$ | $W_1$ | $W_2$ | $W_1$ | $W_2$ | | |
| Solar Image with High Intensity values | No Filament | 17 | 30 | 4 | 5 | 161.32 | 227.84 | 6.377565 | 8.39484 | 172 | 15.7931 |
| | High contrast Filament | 3 | 31 | 1 | 5 | 75.32 | 79 | 13.67699 | 9.31844 | | |
| | Filament edge | 118 | 105 | 10 | 10 | 120.92 | 126.72 | 39.51468 | 35.8127 | | |
| Solar Image with Low Intensity values | No Filament | 41 | 66 | 6 | 8 | 115.32 | 237.12 | 12.2499 | 19.6518 | 124 | 24.6386 |
| | Low contrast Filament | 67 | 14 | 8 | 3 | 29.6 | 2.88 | 20.14324 | 3.5744446 | | |
| | Filament edge | 132 | 127 | 11 | 11 | 82.88 | 92.84 | 44.75798 | 37.96564 | | |
| | High contrast Filament | 97 | 111 | 9 | 10 | 128.32 | 106.44 | 27.15652 | 33.42788 | | |

Where,

$I_\mu$: The mean value for the whole image which is computed as $I_\mu = \frac{\sum_{i=1}^n P_i}{n}$ , where $Pi$ is the set of pixels and $n$ being the number of pixels inside the solar image.

$I_\sigma$: The standard deviation for the whole image which is computed as

$$I_\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - I_\mu)},$$

where $n$ the number pf pixels inside the whole image and $x_i$ the pixel value.

Tables II and III contain the values of the different operators of (1) and (2) for all the regions. Tables I, II and III verify that (1) is correct for harmonic (relatively uniform) regions (with filaments or without filaments).

$$I_\mu > (W_\mu + I_\sigma) \tag{1}$$

However, it was observed that (1) does not work well with the areas that have low contrast filaments; so another condition is adapted to segment these filaments, as shown in (2) and as it is clearly seen in Table III.

$$W_\mu < (R + W_\sigma) \tag{2}$$

TABLE II. FIRST WINDOW: FIRST EQUATION VERIFICATION

| | | | | | W$_1$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Iμ | R | Wμ | W$_\sigma$ | Wμ + W$_\sigma$ | | R + W$_\sigma$ | I$_\sigma$ | Wμ + I$_\sigma$ |
| 172 | 17 | 161.32 | 6.377565 | 167.697565 | No Filament | 23.377565 | 15.7931 | 177.1131 |
| | 3 | 75.32 | 13.67699 | 88.99699 | Filament | 16.67699 | | 75.32 |
| | 118 | 120.92 | 39.51468 | 160.43468 | Filament Edge | 157.51468 | | 120.92 |
| 124 | 41 | 115.32 | 12.2499 | 127.5699 | No Filament | 53.2499 | 24.6386 | 139.9586 |
| | 67 | 29.6 | 20.14324 | 49.74324 | Low contrast filament | 87.14324 | | 29.6 |
| | 132 | 82.88 | 44.75798 | 127.63798 | Filament Edge | 176.75798 | | 82.88 |
| | 97 | 128.32 | 27.15652 | 155.47652 | High Contrast Filament | 124.15652 | | 152.9586 |

TABLE III. SECOND WINDOW: FIRST EQUATION VERIFICATION

| | | | | | W$_2$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Iμ | R | Wμ | W$_\sigma$ | Wμ + W$_\sigma$ | | R+ W$_\sigma$ | I$_\sigma$ | Wμ + I$_\sigma$ |
| 172 | 30 | 227.84 | 8.39484 | 236.23484 | No Filament | 38.39484 | 15.7931 | 243.6331 |
| | 31 | 79 | 9.31844 | 88.31844 | Filament | 40.31844 | | 79 |
| | 105 | 126.72 | 35.8127 | 162.5327 | Filament Edge | 140.8127 | | 126.72 |
| 124 | 66 | 237.12 | 19.6518 | 256.7718 | No Filament | 85.6518 | 24.6386 | 261.7586 |
| | 14 | 2.88 | 3.5744446 | 6.4544446 | Low Contrast Filament | 17.5744446 | | 2.88 |
| | 127 | 92.84 | 37.96564 | 130.80564 | Filament Edge | 164.96564 | | 92.84 |
| | 111 | 106.44 | 33.42788 | 139.86788 | High Contrast Filament | 144.42788 | | |

## VI. CONCLUSION

An accurate, robust and novel algorithm for solar filament segmentation has been introduced in this paper. The algorithm adopted an adaptive, automated sliding window size according to the extent of the heterogeneity of the window pixels instead of using a fixed value. The algorithm uses extracted windows from two different solar images, one high contrast image and one low contrast image. Some statistical calculations were used to judge the presence of the solar filament or not. The range as a statistical measure played an important role in automating the segmentation process. The mean value and standard deviation of the image intensity values in addition to the range, mean value and standard deviation of the adaptive sliding window pixels have been used to make up the filament segmentation equations. This technique can be better evaluated in the subsequent solar image classes which include merging broken structures; filaments characterizations and detecting filament disappearances; additionally it can be the basis for all these subsequent automatic operations.

REFERENCES

[1] L. Alejandro, "The sourrce region of coronal mass ejections," The Astrophysical Journal, vol. 688, pp. 647-655, November 2008.

[2] N. Gopalswamy, R. Kundu, K. Manoharan, A. Raoult, N. Nitta, and P. Zarka, "X-ray and radio studies of a coronal eruption: shock wave, plasmoid, and coronal mass ejection," The Astrophysical Journal, vol. 486, pp. 1036-1044, September 1997.

[3] P. Subramanian and K. Dere, "Source regions of coronal mass ejections," The Astrophysical Journal, vol. 561, pp. 372-395, July 2001.

[4] M. Al-Omari, R. Qahwaji, T. Colak, and S. Ipson, "Machine learning-based investigation of the associations between CMEs and filaments," Solar Phys., vol. 262, pp. 511-539, April 2010.

[5] N. Fuller, N., J. Aboudarham, and D. Bentley, "Filament recognition and image cleaning on Meudon Hα spectroheliograms," Solar Phys., vol. 227, pp.61-73, March 2005.

[6] N. Gopalswamy, "Coronal mass ejections of solar cycle 23," J. Astrophys. Astr., vol. 27, pp. 243–254, June 2006.

[7] N. Gopalswamy, et. al., "The *SOHO/LASCO CME catalog*," Earth, Moon, and Planets, vol. 104, pp. 295-313, April 2009

[8] J. Jing, B. Yurchyshyn, G. Yang, Y. Xu and H. Wang, " On the relation between filament eruptions, flares and coronal mass ejections," Astrophysical Journal, vol. 614, pp. 1054-1062, October 2004.

[9] S. Pojoga and S. Huang, "On the sudden disappearances of solar filaments and their relationship with coronal mass ejections," Adv. Space Res, vol. 32, pp. 2641-2646, December 2003.

[10] E. Robbrecht, Berghmans, D. and R. Van der Linden, "Automated LASCO CME catalog for solar cycle 23: sre CMEs scale invariant," The Astrophysical Journal, vol. 691, pp. 1222-1234, February 2009.

[11] B. Schmieder, "Magnetic source regions of coronal mass ejections," J. Astrophys. Astr. Vol. 27, pp. 139-149, June 2006.

[12] P. Zhou, X. Wang, J. Zhang, "Large-scale source regions of earth-directed coronal mass ejections," Astronomy and Astrophysics, vol.445, pp.1133-1141, January 2006.

[13] J. Aboudarham, et. al., "Automatic detection and tracking of filaments for a solar feature database," Annales Geophysicae. Vol. 26, pp. 243-248. February 2008.

[14] A. Vourlidas, R. A. Howard, E. Esfandiari, S. Patsourakos, S. Yashiro, and G. Michalek ," Comprehensive analysis of coronal mass ejection mass and energy properties over a full solar cycle," The Astrophysical Journal, vol. 722, pp. 1522–1538, September 2010.

[15] P. Bernasconi, D. Rust and D. Hakim, "Advanced automated solar filament detection and characterization code: description, performance and results," Solar Phys., vol. 228, pp. 97-119, January 2005

[16] H. Fang and P.Chen, "Developing an advanced automated method for solar filament recognition and its scientific application to a solar cycle of MLSO Hα data," Sol. Phys., vol.**286**, pp. 385–404, March 2013.

[17] A. Joshi, N. Srivastava and S. Mathew, "Automated detection of filaments and their disappearance using full-disc Hα images," Sol. Phys., vol. 262, pp. 425-436, April 2010.

[18] E. Robbrecht and D. Berghmans, "A broad perspective on automated CME tracking: towards higher level space weather forecasting," Geophysical Monograph Series, vol. 165, pp. 33-42, October 2006.

[19] Y. Yuan Y., F. Shih, J. Jing, H. Wang H. and J. Chae, "Automatic solar filament segmentation and characterization," Sol. Phys., vol. 272, pp. 101-117, August 2011.

[20] F. Shih and J. Kowalski, "Automatic extraction of filaments in H-alpha solar images," Solar Phys. vol. 218, pp. 99-122, December 2003.

[21] E. Robbrecht and D. Berghmans, "Automated recognition of coronal mass ejections (CMEs) in near-real-time data," Astronomy and Astrophysics, vol. 425, pp. 1097-1106, June 2004.