The Development of Automatic and Solar Imaging Techniques for the Accurate Detection, Merging, Verification and Tracking of Solar Filaments

Ibrahim Ali Ahmad Atoum

PhD

UNIVERSITY OF BRADFORD

2012

The Development of Automatic and Solar Imaging Techniques for the Accurate Detection, Merging, Verification and Tracking of Solar Filaments

Ibrahim Ali Ahmad Atoum

A thesis submitted for the degree of
Doctor of Philosophy

School of Computing, Informatics & Media
University of Bradford

2012

October 2012

University of Bradford

To My Beloved Wife Faten, My Daughters Sarah,

Salma, Sadeen, and My Son Osama,

For their endless love and continuous support.

Abstract

Based on a study of existing solar filament and tracking methods, a fully automated solar filament detection and tracking method is presented. An adaptive thresholding technique is used in a segmentation phase to identify candidate filament pixels. This phase is followed by retrieving the actual filament area from a region grown filament by using statistical parameters and morphological operations. This detection technique gives the opportunity to develop an accurate spine extraction algorithm. Features including separation distance, orientation and average intensities are extracted and fed to a Neural Network (NN) classifier to merge broken filament components. Finally, the results for two consecutive images are compared to detect filament disappearance events, taking advantage of the maps resulting from converting solar images to Heliographic Carrington co-ordinates.

The study has demonstrated the novelty of the algorithms developed in terms of them now all being fully automated; significantly the algorithms do not require any empirical values to be used whatsoever unlike previous techniques. This combination of features gives the opportunity for these methods to work in real-time. Comparisons with other researchers shows that the present algorithms represent the filaments more accurately and evaluate computationally faster - which could lead to a more precise tracking practice in real-time.

An additional development phase developed in this dissertation in the process of detecting solar filaments is the detection of filament disappearances. Some filaments and prominences end their life with eruptions**.** When this occurs, they disappear from the surface of the Sun within a few hours. Such events are known as disappearing filaments and it is thought that they are associated with coronal mass ejections (CMEs).

Filament disappearances are generally monitored by observing and analysing successive solar H-alpha images. After filament regions are obtained from individual H-alpha images, a NN classifier is used to categorize the detected filaments as Disappeared Filaments (DFs) or Miss-Detected Filaments (MDFs). Features such as Area, Length, Mean, Standard Deviation, Skewness and Kurtosis are extracted and fed to this neural network which achieves a confidence level of at least 80%. Comparing the results with other researchers shows high divergence between the results. The NN method shows better convergence with the results of the National Geophysical Data Centre (NGDC) than the results of the others researchers.

Acknowledgments

First of all, my immense gratitude to ALLAH The Most Merciful, The Most Beneficent, without His help and blessings, this thesis would not have progressed nor have seen the light.

My most sincere thanks go to my dissertation principal supervisor, Prof. Rami Qahwaji for introducing me to the wonders of scientific research. I thank him for his guidance, encouragement and support during the entire period of my doctoral study years.

I am grateful to Dr. Tufan Çolak, who encouraged and helped me through long discussions. He has given me a lot of comments, support and help to take the first step of this valuable research. I thank him for all the many constructive ideas, helpful discussions, guidance and advice, which have all been much appreciated.

I am deeply thankful for my external supervisor, Dr. Maaruf Ali, who encouraged and helped me to publish most of my research. I thank him for all his useful ideas, cooperative discussions and meetings, and I appreciate his supportive guidance and meticulous checking of all my chapters.

Last but not the least; I would like to thank my family for their endless love, encouragement and support. My gratitude to them can never be overstated.

Publications

1) **Atoum, I. A.**, Qahwaji, R. S., Colak, T. and Ahmed, Z. H., (2009) Adaptive Thresholding Technique for Solar Filament Segmentation, *Ubiquitous Computing and Communication Journal*, **4** (4), pp. 91-95.

Papers under Review

2) **Atoum, I. A.**, Qahwaji, R. S., Colak, T. and Ali M., (2012) Automated Algorithms for the detection and merging of broken filaments using neural network, *Solar Physics*, (SUBMITTED).
3) **Atoum, I. A.**, Qahwaji, R. S., Colak, T. and Ali M., (2012) Automatic technique for detecting the disappearance of solar filaments, *Journal of Space Weather and Space Climate*, (SUBMITTED).

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ALT	Adaptive Local Thresholding
ALT&V	Adaptive Local Thresholding and Verification
ANN	Artificial Neural Network
BBSO	Big Bear Solar Observatory
BF	Broken Filament
CAO	Catania Astrophysical Observatory
CME	Coronal Mass Ejection
CMP	Central Meridian Passage
DF	Disappeared Filaments
EI	Enhanced Image
FAR	False Acceptance Rate
FM	False Merged parts
GHN	Global High-resolution Hydrogen-alpha Network
Hα	Hydrogen Alpha
HCM	Heliographic Carrington Maps
HSOS	Huairou Solar Observing Station
KSO	Kanzelhöhe Solar Observatory
MDF	Miss-Detected Filaments
MP	Merged Parts
MSE	Mean Squared Error
NBP	Number of Broken Parts
NASA	National Aeronautics and Space Administration
NGDC	National Geophysical Data Centre

NJIT	New Jersey Institute of Technology
NMP	Not Merged Parts
NN	Neural Network
NOAA	National Oceanic and Atmospheric Administration
SDO	Solar Dynamic Observatory
SE	Structuring Element
SF	Stable Filament
SOHO	SOlar and Heliospheric Observatory
SOON	Solar Observing Optical Network
SSN	Station Serial Number
SVM	Support Vector Machine
SWRL	Space Weather Research Lab
USET	Uccle Solar Equatorial Table
YAO	Yunnan Astronomical Observatory

List of Mathematical Symbols

$\bar{X}$	Mean Value of the sliding window
$\bar{X}_{SW}$	The average intensity value of the small window.
μ	The average intensity value.
μ_f	The average intensity value of the whole filament.
σ	Standard Deviation
σ_f	The standard deviation intensity value of the whole filament.
σ_w	The standard deviation intensity value of the sliding window.
Θ	Erosion Operation
img_{max}	The maximum intensity value in the solar image.
img_{min}	The minimum intensity value in the solar image.
img_{mean}	The average intensity value in the solar image.
LW	Large Window
Max	Maximum
Min	Minimum
MR	Mid – Range
R	Range
R_{LW}	The range intensity value in the large window.
SW	Small Window
TH	Threshold
W	Window
W_{max}	The maximum intensity value in the sliding window.
W_{min}	The minimum intensity value in the sliding window.

CHAPTER ONE

1 INTRODUCTION

1.1 Background

As the importance of forecasting space weather increases, the need for detecting solar features affecting space weather also increases. The increasing powers of computers, image processing and machine learning techniques offer opportunities to develop automatic detection methods for solar features and activities that may affect life on Earth.

Solar flares and CMEs are the most important solar events that lie behind space weather; these solar eruptions release vast quantities of electromagnetic radiation and charged particles (Al-Omari *et al*., 2010). Solar flares are sudden, short lived, burst of energy on the Sun's surface, lasting from minutes to hours as described in (Colak and Qahwaji, 2007) and shown in Figure 1.1.

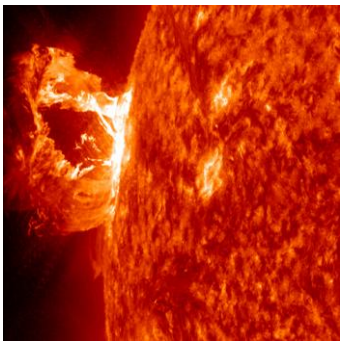


Figure 1.1. Solar flare recorded by the NASA Solar Dynamics Observatory on April 16th 2012.

The most important method for space weather prediction is the accurate detection and monitoring of the evolution of solar features affecting space weather. Detecting filaments can indicate the possible occurrence of CMEs. Hence, solar filaments are features that play a vital role in the study of space weather.

In Hα images as shown in Figure 1.2, filaments appear as dark ribbons against brighter and hotter background. At the limb, they become bright features against the sky and are then called prominences.

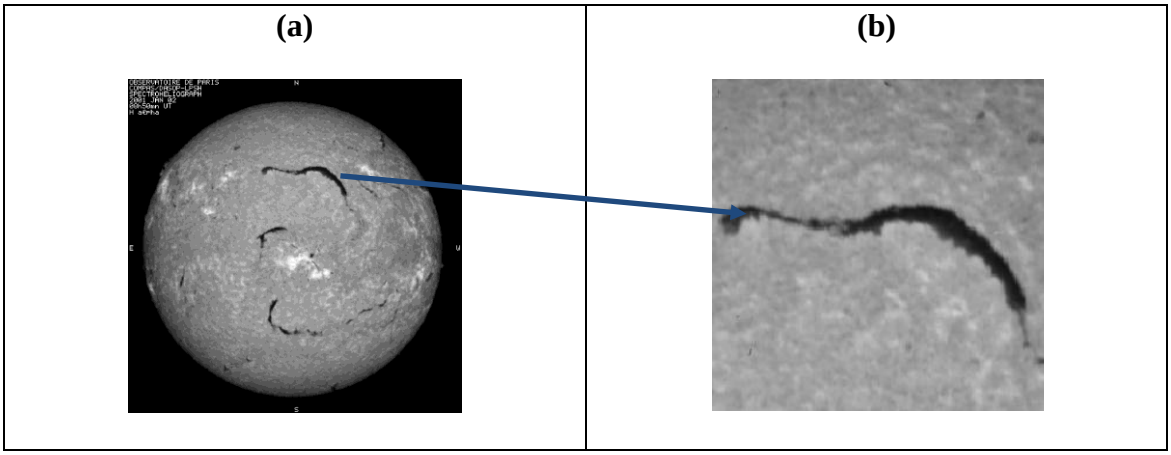


Figure 1.2. Solar filament (a) Hα solar image observed at the Meudon observatory on January 2nd 2001. (b) Solar filament as shown in (a). (http://bass2000.obspm.fr/home.php)

Some filaments erupt and disappear within hours and are known as disappeared filaments as shown in Figure 1.3. Artificial neural networks (ANNs), usually called Neural Networks (NNs) are fairly simple automated models that simulate the neural structure of the brain. This new approach to computing provides a more attractive automatic processing compared to traditional methods, because computers can perform complicated problems but they still have problems in recognizing simple patterns.

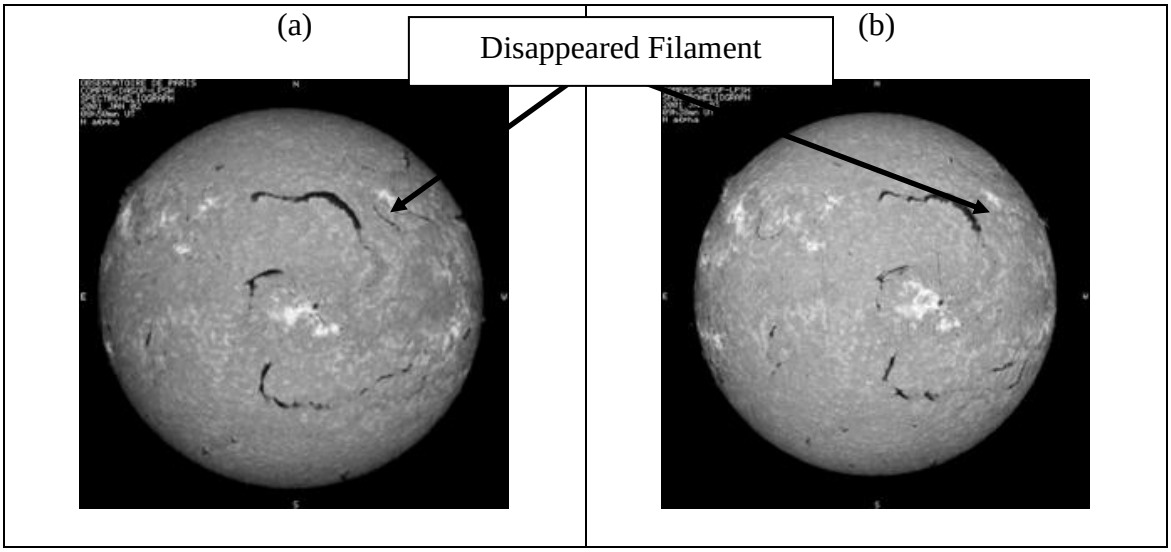


Figure 1.3. Solar images illustrate a disappeared filament (a) Hα solar image observed at the Meudon observatory on January 2nd 2001. (b) Hα solar image observed at the Meudon observatory on January 3rd 2001.

In this research, the aim is to design a fully automated real-time system that can detect the disappearance of solar filaments by analysing solar images; the fundamental steps are shown in Figure 1.4. This system uses solar images for the segmentation phase, detecting solar filaments, merging broken filaments and finally tracking their disappearance.

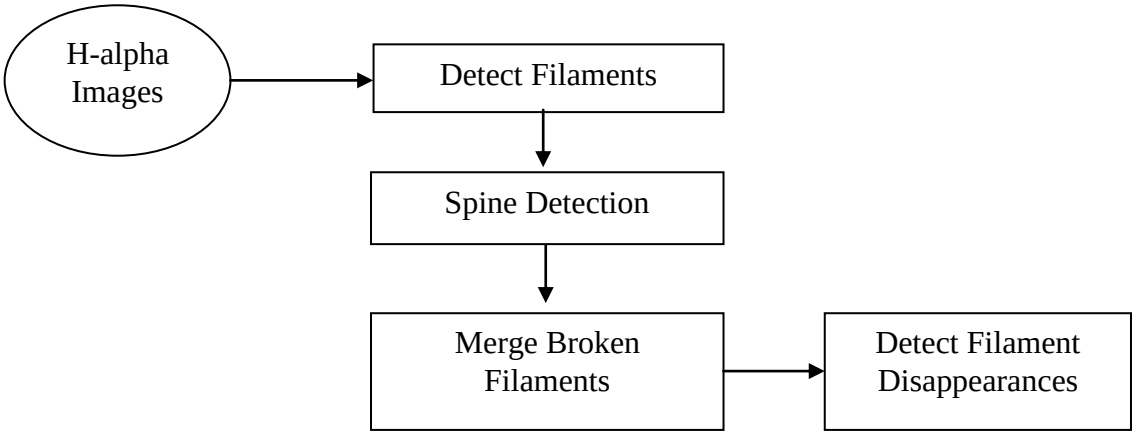


Figure 1.4. Suggested system for automatic detection of filament disappearances.

1.2 Solar filament segmentation, detection, merging and tracking

Solar images are obtained by ground and space observatories. There are plenty of public solar images that are obtained using different wavelengths. Solar features seen in those images could be used by scientists to observe and study solar events. However, there are many challenges facing solar imaging. The automated detection and tracking of solar features is one of these challenges. Solar image processing and feature extraction techniques are usually used to extract numerical features that can provide efficient representation for solar features, solar activities and/or general regions of interest in the solar images. Knowledge extraction and prediction of forthcoming activities can be achieved using machine learning.

A successful integration between image processing and machine learning could provide automatic detection, tracking and even classification of filaments. In this case, the automated system could be used to represent the detected filaments using statistical and geometrical features (i.e., size, location, shape, orientation, etc.).

Solar filaments are large regions of very dense, relatively cool ionized gases, held in place by magnetic fields. They are elongated structures and dark features appearing in H-alpha (Hα) solar images as shown in Figure 1.5.a.

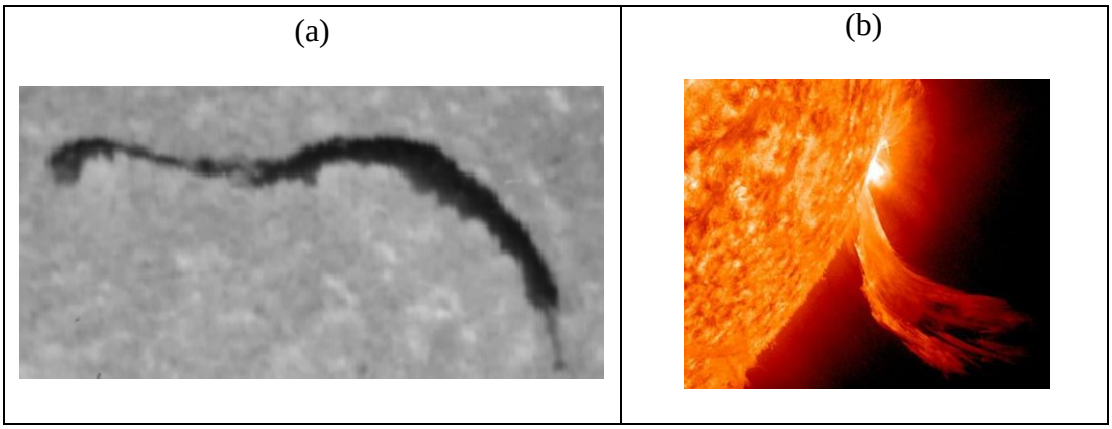


Figure 1.5. (a) Solar filament seen in the Hα solar image observed at the Meudon observatory on January 2nd 2001. (b) Solar prominence seen from the NASA/Solar Dynamics Observatory on September 8th, 2010.

Filaments appear dark because they are cooler than their surroundings, while they appear bright when they become visible on the edge of the solar disk and are then called prominences as shown in Figure 1.5.b.

Despite the advances expressed in the previous works, machine learning has not been used for large-scale analysis of filaments during the detection or classification phases. Still one of the challenges that face this paradigm is the need to develop and implement a detection technique that could avoid the use of any empirical values to produce a fully automated technique.

Additional requirements include producing a fully automated filament detection algorithm that is both fast and accurate. Time requirement is definitely a factor which would affect the ability to synthesize a filament detection and tracking system as part of a proposed real-time system. The successful real-time implementation of such a system would give the opportunity to produce timely space weather alerts and quick look-up of results.

Solar filaments are characterized by its low intensity values in Hα images because they are darker in colour. This nature, being well separated from the background, gives the opportunity to use a thresholding segmentation technique. This is achieved by defining a range of brightness values in the original image, choosing the image pixels that fall in this range to be foreground and ignoring other pixels as they represent the background. The same technique has been used in this study without using any range of brightness value in segmenting unambiguous filaments whilst introducing minimal noise compared to the other techniques.

This segmentation method is followed by a region growing technique developed by (Qahwaji and Colak, 2005). This method uses the foreground pixels as seeds; then it combines the adjacent pixels and the unwanted seeds are eliminated. A fully automated detection technique is developed in this study to retrieve the actual filament area from

these region grown filaments. This technique involves statistical parameters and morphological operations and avoids using any empirical values whatsoever. The shape of the detected features is represented by determining its spine geometry which gives the opportunity to extract all the filament morphological features.

In some solar images, some filaments may be broken or segmented into several small filaments due to differences in intensity values or fail in the pre-processing parts of this phase. These filaments should be merged to get the correct filament numbers and to get the actual filament parameters like the spine length. An NN-based merging technique is used by extracting some of the filament characteristics and feeding it to an NN to classify filaments into two groups, merged or not merged.

After obtaining the areas of individual filaments, the resultant filaments in two consecutive images are compared to detect the filament disappearances. Again a NN-based tracking technique is followed by extracting a set of characteristics of individual filaments and feeding it to an NN to distinguish it from a miss-detected one.

1.3 Motivation

The importance of studying solar filaments comes from considering its disappearances as a significant indicator for possible occurrence of CMEs, which is considered as the major cause of geomagnetic storms.

CMEs could be initiated from closed magnetic field regions such as filament regions and it is now almost certain that there is a close association between CMEs and filament disappearances (Gopalswamy *et al*., 2003, Moon *et al*., 2002, Pojoga and Huang, 2003, Jing *et al*., 2004, Gopalswamy, 2006, Schmieder, 2006, Alejandro 2008, Robbrecht *et al*., 2009, Al-Omari *et al*., 2010).

CMEs are enormous bubbles of hot plasma (billions of tons of magnetized plasma) that propagate away from the solar corona into the interplanetary medium at a very high velocity (Alejandro, 2008). Figure 1.6 shows one CME. These bubbles which are

carrying a huge amount of energy can – if directed towards our planet – cause massive disruption in our communication networks, failures in navigational systems and satellites, and power outages and disruptions.

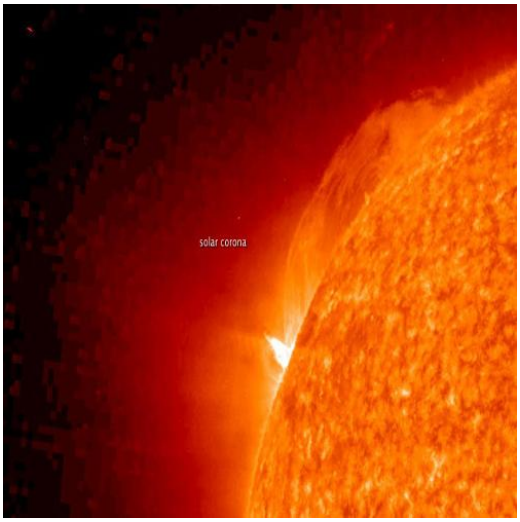


Figure 1.6. Coronal Mass Ejections observed at SDO, the NASA Solar Dynamics Observatory, on January 27th, 2010.

These CMEs have been observed in the images of the solar corona obtained by Solar Heliospheric Observatory (SOHO) mission's Large Spectrometric Coronagraph (LASCO) since 1996. A coronagraph is a telescope that uses a disk to block the Sun's direct light, while permitting light from surrounding sources; this light reveals the solar corona and a coronagraph could be regarded as a producer of artificial solar eclipses.

SOHO has two coronagraphs on board; "C2" coronagraph and C3 coronagraph. C2 coronagraph images are usually coloured red; C3 coronagraph images are blue as shown in Figure 1.7.

There are three possible indicators of CME onsets, which are: filament disappearances, coronal dimming and solar flares. The coronal dimming phenomenon occurs when the intensity of the Sun corona decreases (Attril *et al*., 2006).

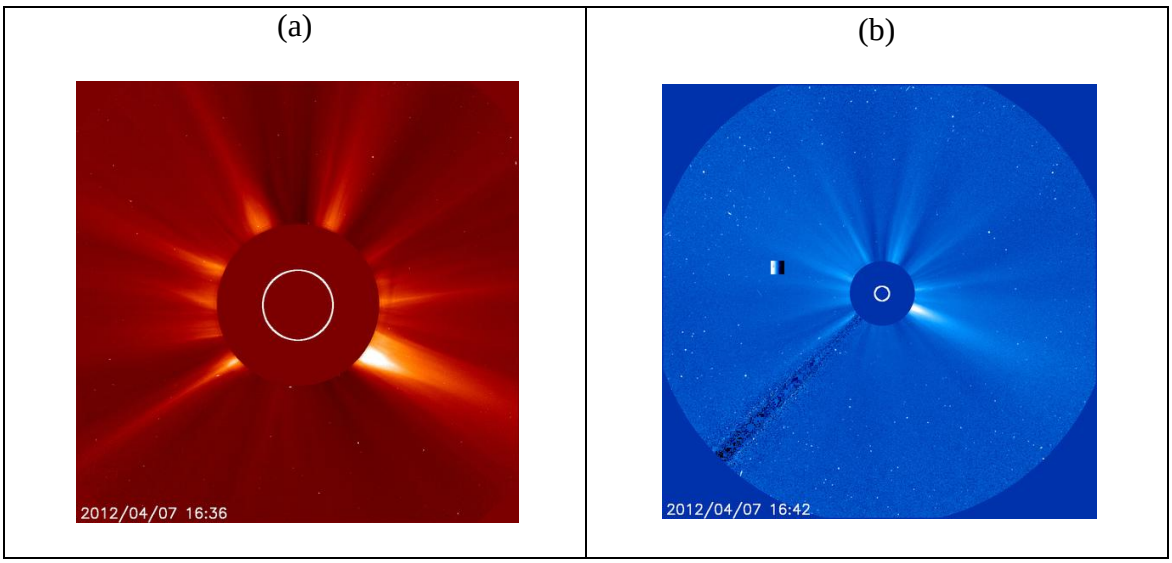


Figure 1.7. Two examples of coronagraph images. (a) SOHO LASCO C2 image taken on 07th April, 2002 at 16:36. (b) SOHO LASCO C3 image taken on 07th April, 2002, 16:43.

With the current huge development in Space instruments, computers and communications, the need becomes greater for quick alerts and warnings of the risks that threaten our planetary atmosphere and for forecasting space weather. Thus the need becomes more critical for developing a real-time CME prediction system. The work in this underlying study will stop at detecting filament disappearances and in future may be used for developing this CME prediction system.

1.4 Research Aims and Objectives

The main aims of this research were the accurate detection of solar filaments by the process of merging broken filaments to create an automated detection system for solar filament disappearances. The technique will receive the real-time solar images that are observed by different ground-based observatories and are available online. The development of this real-time system using machine learning algorithms will help in modelling a reliable tracking technique. The different objectives of this research are summarized as:

- Designing an automated computer platform for automatic detection of solar filaments.
- Designing an automated tool for extracting different features of solar filaments. A geometrical-based approach is used for extracting the filament spine and determining its properties. These properties give the opportunity to accurately accomplish the subsequent processing tasks.
- Designing an automated tool for merging the broken filaments structures. This method has exploited the findings of the initial extraction of the filament spine.
- Designing novel techniques to effectively track the disappearances of solar filaments.

1.5 Original Contributions

The main original contributions presented in this thesis can be summarised as follows:

- The development of an adaptive thresholding technique for segmenting Hα solar images in order to detect filaments as Regions of Interest (RoI) and discard everything else as background. These well-defined and visible filaments could be considered for further analysis by characterizing their features which may give the ability to provide suitable inputs for machine vision techniques. (Atoum *et al.*, 2009).
- Presentation and development of a fully automated technique for the detection of solar filaments by manipulating statistical parameters and morphological operations. The present detection process avoids using any empirical values whatsoever. This technique is described in Chapter four and in a submitted paper.

- Representation of the shape of the detected filaments by determination of its spine geometry. The algorithm gives the opportunity to extract the filament morphological features, such as: the filament length, filament centre, filament head-end points, filament tail-end points and the filament boundary. The algorithm is valuable as part of a real-time system for detecting and tracking solar filaments. This algorithm is presented in Chapter four and in a submitted paper.
- The use of the NN classifier algorithm to achieve the merging of the broken filaments with a higher merging percentage than before. The algorithm avoids the use of empirical constant values for different filament merging attributes. This algorithm is described in Chapter five and also in a submitted paper.
- Finally a solar filament tracking technique was implemented to detect disappearing filaments. The technique exploits the relatively small movements of the filaments over HCMs. It uses an NN classifier to distinguish between the actual disappeared filaments and the phenomenon of disappearing miss-detected filaments. This technique is presented in Chapter six and in a submitted paper.

1.6 Outline of the Thesis

This thesis is organized as follows:

- Chapter Two provides a literature review of recent research on the automatic detection techniques for solar filaments. The methods used in each phase of the detection process are detailed.
- Chapter Three explores the available sources of solar data that can be used in the research presented in this thesis. Some of these data are used for

solar filament segmentation, other solar data is used for automatic detection, spine representation, merging algorithms and tracking techniques.

- Chapter Four introduces an improved image segmentation algorithm; an automated detection algorithm which includes retrieving the approximate actual filament from the region grown one and the detection of the filament boundary. Additionally, it describes the implementation of the filament spine representation.
- Chapter Five describes and implements a merging technique that utilizes the findings of the spine extraction to merge the broken structures using a neural network.
- Chapter Six describes the implementation of a tracking technique. It includes the filament detection stage and creating the Heliographic Carrington Map. A neural network is used then to categorize the detected disappearances as true or false.
- Finally the concluding remarks and recommendations for future work are presented in Chapter Seven.

CHAPTER TWO

2 LITERATURE REVIEW

2.1 Solar Filament Detection

In image processing, the term feature detection refers to the methods that aim to make a decision at every image pixel on whether it is an image feature or not. The outcomes of these methods will be subsets of the image domain, often in the form of isolated points.

In solar imaging, given a solar image, filament detection means to determine whether or not the specified filament is present, and, if present, determine its characteristics. These characteristics include properties such as: length, area, centre, head-start points, tail-end points and filament boundary points. Once these characteristics are found, they are utilized in trying to merge the broken filaments in order to approximately restore the actual size of the detected filaments. The image processing task also involves the very important task of detecting the filament disappearance - where studies have shown that filament disappearances are usually associated with the occurrence of CMEs. The filament detection process should also create algorithms which execute speedily.

Filaments usually appear above the chromosphere as thin elongated structures in Hα solar images. Detecting and characterizing solar filaments is important for several aspects of solar activities because of their association with geomagnetic storms (Al-Omari *et al.*, 2010). An efficient detection system which is fully automated and further, works in real-time is highly desirable, which makes its implementation in software rather challenging.

There have been a number of methods developed for detecting solar filaments, merging broken filaments and detecting the filament disappearances during the last

decade. Each of these methods is addressed in detail in the following sections. The detection methods are presented in Section 2.2, the detection techniques are presented in Section 2.3, the merging methods are described in Section 2.4, the techniques used for detection of filament disappearances are described in Section 2.5 and finally conclusions are given in Section 2.6.

2.2 Detection methods

The papers on the various stages of filament processing found in the literature are listed below. Methods used for detecting solar filaments include:

- Threshold and Region-based Technique (Gao *et al*., 2002)
- Artificial Neural Networks (Zharkova and Schetinin, 2003)
- Morphological Operations (Shih and Kowalski, 2003)
- Adaptive Thresholding Method and Support Vector Machine (SVM) (Qu *et al*., 2005)
- Thresholding, Region Growing and Spine detection (Fuller *et al.*, 2005)
- Thresholding, Region Growing and Spine Tracing (Bernasconi *et al*., 2005)
- Thresholding, Region Growing and Feature Verifications (Qahwaji and Colak, 2005)
- Intensity and Size Threshold (Joshi *et al.*, 2010)
- Adaptive Thresholding Method (Yuan *et al.*, 2011)

Methods used for merging broken structures include:

- Distance Criterion (Gao *et al.*, 2002)
- Implementation of Distance and Angle (Bernasconi *et al*., 2005)
- Closing, Thinning, Pruning and Adaptive Edge Linking (Qu *et al*., 2005)
- Morphological Closing Operation (Fuller *et al.*, 2005)
- Grouping and Distance Criteria (Joshi *et al.*, 2009)

- Closing Operation (Yuan *et al.*, 2011)

Methods used for tr*acking* solar filaments include:

- Comparison of Two Consecutive Images (Gao *et al.*, 2002)
- Component Matching (Qu *et al.*, 2005)
- Comparison of Three Consecutive Images (Bernasconi *et al.*, 2005)
- Applying a region growing techniques over Carrington Maps (Aboudarham *et al.*, 2008)
- Labelling Criterion (Joshi *et al.*, 2009)

2.3 Detection Process

In this section, all of the algorithms that were implemented and discussed in the literature are presented. There have been many successful attempts at designing algorithms for detecting solar filaments. Gao (Gao *et al.*, 2002), combined thresholding and region growing techniques to achieve filament detection. Filament candidates were obtained using a thresholding technique. The region growing technique was then used for grouping the candidate pixels to form filament areas. The detection method includes three particular operations; deletion of any adjacent pixels that touch the solar limb, because these pixels could be part of filaments, so this avoids the detection of prominences; remove any filament area with less than half the median intensity value of the whole image; because it is considered noise; an eighty-adjacency connection method is used to connect the adjacent pixels (eighty pixels around the central pixel are checked for connectivity) to avoid small errors in solar images. During the image processing, some large filaments could become broken into small fragments which could become excluded because they are now less than a 220 pixel area threshold. This method could not be considered a fully automatic detection technique because it uses

many different constant values in the detection process. Also it cannot handle low contrast filaments and thus produces unstable results.

A feed-forward ANN technique is used by Zharkova (Zharkova and Schetinin, 2003); this network is composed of two hidden and one output neurons to extract solar filaments automatically from Hα solar images. In total 55 filament fragments were selected depicting filaments with different backgrounds; one is used for training the network and the other 54 are used for testing. The main aim of the proposed ANN is to remove a contribution of the variable background elements which is defined as a background function. The technique is based on a standard sliding window technique as follows: The given image is transformed into columns; with each column representing a set of pixels taken from a sliding window width of size 3×3. The output neuron makes a decision on whether the central pixel is a filament or non-filament pixel. The results of this technique were not validated nor optimized.

A superposition of morphological closing operations was applied by Shih (Shih and Kowalski, 2003) to separate filaments from the granular background. The process used eight directional linear 11×11 structuring elements with 90°, 0°, 45°, 135°, 67°, 112.5°, 22.5°, 157.5° slopes, respectively as shown in Figure 2.1. An additional closing operation with a 3×3 structuring element was applied to eliminate spurious features. The dark features obtained are used as seeds for a region growing process. This checks the neighbourhoods of the detected features and compares them against the original pre-processed image. All connected black points in the pre-processed image neighbouring the detected filament are marked as belonging to the filament. Using 8-neighbor connectivity, all connected black points in the pre-processed image neighbouring the detected filament are marked as belonging to the filament. This detection method uses an extensive set of morphological operations in detecting solar filaments which could thus consume significant computational time.

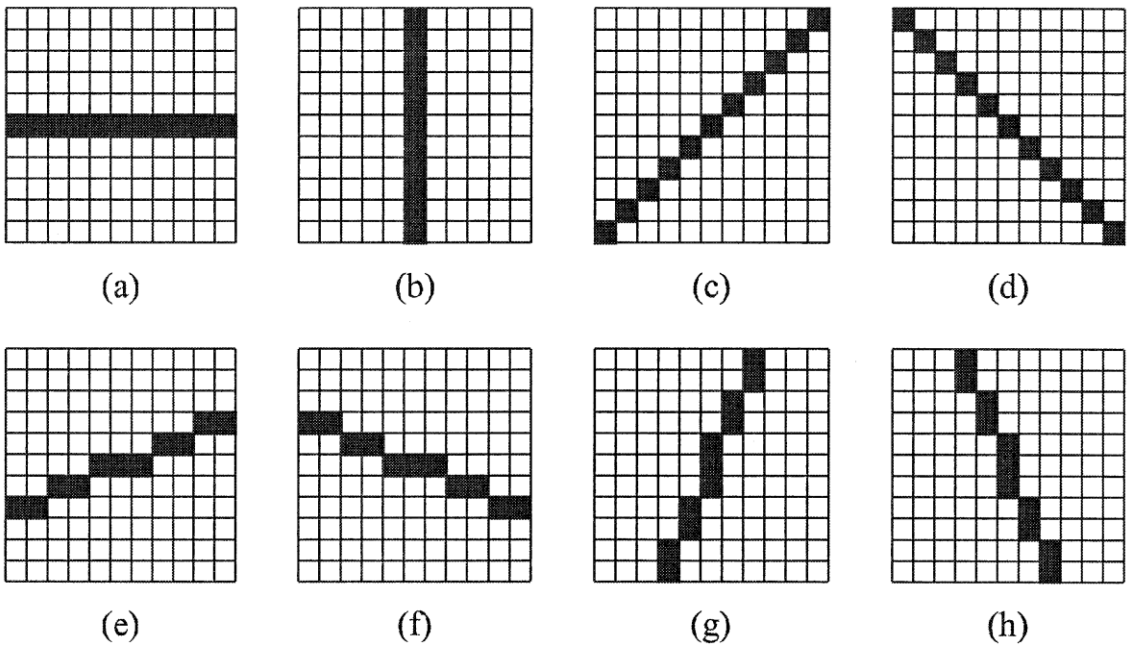


Figure 2.1. Eight directional linear 11×11 structuring element with 90°, 0°, 45°, 135°, 67°, 112.5°, 22.5°, 157.5° slopes, respectively.

The Sobel operator was applied by Qu (Qu *et al*., 2005) to detect the edges of the filaments. This operator emphasizes the regions of high spatial gradient that correspond to edges by performing a two dimensional spatial gradient measurement on an image. Then, two thresholding operations are performed; a global one for the whole image to select filaments that have high contrasts relative to the background. A second local one is performed for sub-images of size 100×100 pixels to select filaments that have locally high contrasts relative to the background. Initially, a set of adaptive thresholds ranging from zero to the median of pixel intensities of the image are chosen for segmenting dark regions. The global threshold is computed by dividing the result of applying the Sobel operator over the segmented region by the number of the pixels in the same region. When the new expanded region meets the edges of filaments then the threshold value reaches the maximum which then represents the best global threshold for segmenting solar filaments. The local threshold is computed according to two criteria; firstly, the local threshold equals the optimal global threshold plus or minus 30 and secondly; the size of the region obtained by the local threshold is less than three times that of the

region obtained by the global threshold. Additionally, any small regions with area less than 10 pixels are considered noise and removed. Finally, a SVM classifier is used to discriminate sunspots from filaments, where the sunspots are represented by nine features representing the input of the SVM, which are extracted from a window of size 100 ×100 pixels. This method cannot be considered a fully automatic detection technique because several constant values have been utilized for computing the local threshold and for removing the noisy pixels.

The detection process by Fuller (Fuller *et al*., 2005) is started by detecting the seed regions that need to be grown. Primarily, the image is divided into local square windows with size 1/6 of the image size and then any pixel value less than the mean value in this window is discarded from the calculation. The threshold for each sub-image is then computed by using a local quiet Sun intensity minus a constant value multiplied by the standard deviation as shown in equation 2-1. The constant value was chosen to be 3.7.

$$T_{win} = QS_{win} - \alpha \times \sigma_{win} \quad (2\text{-}1)$$

T_{win} is the threshold for each sub image, QS_{win} is the local quiet Sun intensity, α is the constant value and σ_{win} is the standard deviation for each sub image. Any remaining spurious seeds were removed by keeping only seeds whose size is more than 20 pixels. A region growing technique is then applied to group pixels together; the upper and lower limits of the growing process are identified by threshold values; the lower threshold is set to be zero while the upper threshold is computed by multiplying the standard deviation of the bounding rectangle of this seed region by 1.5. Then this value is subtracted from the quiet Sun intensity of the bounding rectangle. To fill in small holes and to merge nearby regions a closing operator is applied. The spine detection process is also achieved using multiple morphological operators. They have applied an iterative thinning method to produce the skeleton of the region followed by a pruning process to remove the branches of the skeleton tree. This method cannot be considered a

fully automatic detection technique because several constant values are used for computing the threshold for each sub-image, for removing the spurious pixels and in calculating the threshold that is used in the region growing technique.

The detection process by Bernasconi (Bernasconi *et al*., 2005) starts by creating a binary filament mask in which pixels labeled one are considered part of a filament and the background is filled by zeros. Then the process searching for filaments is limited to within a latitude circle of 60° heliocentric from the Sun center. Secondly, sunspots are removed from the mask according to the thresholding method that was introduced by Shih (Shih and Kowalski, 2003). The resulting pixels are then used as seeds for a region-growing operation using upper and lower threshold values to avoid detecting too large spots and too small spots. After that a new threshold value is assumed to extract the filament mask; this threshold extracts not only the required filaments but also spurious pixels which then should be removed; this is done by applying a morphological filtering operation that was introduced by Shih (Shih and Kowalski, 2003). The resulting pixels are now used as seeds for a region growing operation that extends the filament size to an adopted threshold. After all filaments are detected to their full shape, a threshold is again used to exclude the too small filaments or the false detection results. Thereafter the filament boundary is extracted as an array of Cartesian co-ordinates of each pixel along the outline of the filament. Finally a principal curve algorithm uses a multi-step iterative technique to determine the filament spine. This detection process has limited the searching process within a latitude circle of 60° which means it will not be able to detect filaments outside this. Several constant values are used to implement the detection process which prevents it being considered a fully automatic detection technique.

The detection process by Qahwaji (Qahwaji and Colak, 2005) starts by detecting seed pixels using an intensity filtering technique. This technique uses the standard

deviation, the mean value and a constant value in a thresholding equation to detect the seeds pixels for the desired solar filaments. A region growing algorithm is then applied to detect the RoIs (Regions of Interest) which have not been detected by the previous filtering algorithm; this process involved two main stages: firstly combining the adjacent seed pixels and secondly eliminating the unwanted seeds. The results of these calculations are then fed to a neural network to verify the detected regions and minimize the false acceptance rate. This method uses constant values in the statistical equations for detecting the solar filament, thus it cannot be considered as a fully automated detection technique.

Joshi (Joshi *et al.*, 2010), applied an intensity threshold based on a variable local thresholding method introduced by Shih (Shih and Kowalski, 2003) to identify when a pixel value is related to a filament and not to the background. The method starts by calculating the median at every pixel in the image using a 19×19 pixel neighborhood centred on the pixel. Two cut-off values are used, a lower cut-off value which is 10% of the intensity range of the image and a higher cut-off value which is 90% of the intensity range of the image. The threshold is equal to the lower cut-off value if the median value at a pixel is less than or equal to this value. If the threshold is equal, the higher cut-off value is chosen if the median filter at the pixel is less than or equal to this value. This threshold value, thereafter, is used to extract the solar filaments. Then a fixed threshold value is then applied to eliminate some non-filament features, this is set to 12 pixels in size. Finally, dark sunspots are removed by selecting a threshold to be 30% of the disc-centre intensity of the normalized images (Shih and Kowalski, 2003). A constant value was used to remove the non-filament features; this prevents the algorithm from being considered as a fully automated detection process.

An adaptive segmentation procedure is presented by Yuan (Yuan *et al.*, 2011). It consists of five different stages; convolving the solar image by a high-pass Laplacian

filter; using the median and standard deviation to create a series of thresholds by generating an arithmetic progression method; segmenting the solar image using these threshold values to obtain the segmented regions and the difference regions; the segmented regions and the difference regions are then used to calculate a series of thresholds; and the final threshold value is then chosen to be the maximum one. After this segmentation method, the filaments are removed if they are too small (3.5% of the radius of the solar disk), then a closing operation is applied to connect the broken filaments using a disk structuring element with size of 0.5% of the radius of the solar disk. A shape-based threshold is then adopted to remove the sunspots. The area and perimeter were used to create a shape measurement value, these values will be within the [0, 1] interval. The larger value is more likely if it is a sunspot; an experimental value of 0.7 is chosen; so if the shape measurement of an object is more than 0.7 then it is removed. A morphological reconstruction operation is then applied to fill in the holes inside each filament. An iterative morphological thinning is applied after that to obtain the skeleton of the filament. The image is initially thinned by the left SE as shown in Figure 2.2 and then by the right SE as shown in Figure 2.2 and followed by the remaining six 90° rotations of the two mentioned structure elements.

0	0	0
x	1	x
1	1	1

x	0	0
1	1	0
x	1	x

Figure 2.2. The two basic structural elements used in the morphological thinning process, (x indicates do not care).

The process is repeated until none of the thinning processes produces any further changes. A graph theory approach is then implemented to remove the skeleton branches.

The adjacent matrix representation for each filament is created; for each pixel on the skeleton a vertex that represents it is created; all the vertices and connectivities of the vertices are then created. If the pixels of the skeleton are 8-connected, an edge connecting the two vertices corresponding to the two pixels is then created. After that, two graph algorithms are implemented to find the main skeleton; the first one; finding the shortest path between all pairs of vertices and the path with the maximum length is the main skeleton. The second is finding all the end vertices by searching over the shortest paths between each pair of the end vertices. The part with the maximum length is the main skeleton. A constant value was used for creating a shape measurement value; this prevents it from being a fully automated detection process.

2.4 Merging Process

In some solar images, some filaments may be broken or segmented into several smaller filaments. In order to detect the correct numbers of filaments and to extract the correct characteristics of these merged filaments, such as length, starting end-points and ending end-points, a merging algorithm is needed to combine these broken filaments. Most of these broken structures are large ones; large dark filaments that disappear generally produce larger interplanetary atmospheric disturbances that are capable of having an impact on Earth. So, merging and detecting these large filaments is a significant measure in detecting filament disappearances. There are two criteria specified in the literature to merge broken elongated objects (Cheng *et al*., 2007 and Ingrid *et al*., 2002): the broken features should be close enough and they should satisfy some relative orientation requirements.

In one of the early studies in this area by Gao (Gao *et al.*, 2002), a distance constant value to be 40 pixels is adopted as a merging threshold; if the distance between each two filaments is less than this threshold then they are considered as one filament. A

constant value is used for the distance between the filaments to achieve the merging process, again it cannot be considered as a fully automatic merging technique.

The two mentioned merging rules were used in (Bernasconi *et al*., 2005) by estimating two constant values for the two criteria; one for the distance (d) of each spine end-point of each filament to all other filaments and one for the angle (θ) between the last two segments of the two spines relative to the line connecting the two end-points as shown in Figure 2.3. If d<25 pixels then the two filaments are merged. If d>25 pixels and d<100 pixels then the θ is calculated, if for both segments θ_1 and θ_2 are both <22.5° then they are merged. This merging technique is considered imperfect as stated by the author.

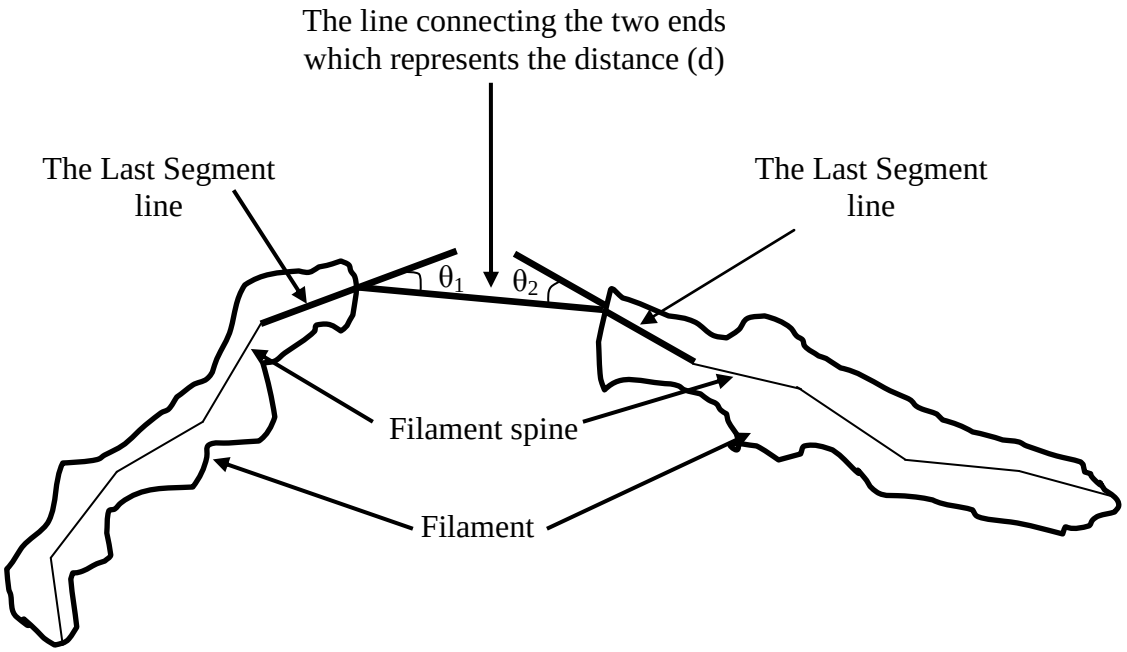


Figure 2.3. The angles between the last two segments and the line connect the two end-points of the two filaments.

It must be noted that this merging process is implemented after characterizing the solar filament by Bernasconi. Two constant values have been used for the distance and the angle in order to implement the merging process. This prevents the algorithm from being an automatic merging process.

A set of mathematical morphological operations that includes: closing, thinning and pruning and adaptive edge linking methods were undertaken by Qu (Qu *et al*., 2005). A closing operation is used to eliminate the small gaps. A thinning operation is then used to determine the skeleton of the filaments; this process is repeated using a sequence of eight structuring elements until no further changes can occur. A pruning operation is then used to determine the spines of the filament; this operation also uses eight structural elements. Even after using the previous set of morphological operations there could be big gaps in broken filaments which could not be merged. Therefore, an adaptive edge linking method was used to connect edges based on the orientation of the filament spines (Shih and Cheng, 2004). In this method, an adaptive dilation morphological operation is applied at each endpoint with an adaptive elliptical structuring element (SE). The size and orientation of the SE are adjusted according to local properties, such as curvature and slope. Some post-processing operations are applied, such as, thinning and pruning to remove noisy edge segments. This extensive use of morphological operations in extracting the filaments spine could require significant computational time.

A morphological closing operation was implemented by Fuller (Fuller *et al*., 2005) to merge closed regions. A simple morphological closing operation to achieve the merging process is useful for the filaments that are close to each other. To get a more efficient and accurate result for this operation and to fill in the big gaps in broken filaments, a larger structuring element or multiple dilations followed by the same numbers of erosions is utilized. This will, however, destroy the structural integrity of the underlying filament. Thus this algorithm is not suitable for widely disjointed broken filaments.

Joshi (Joshi *et al.*, 2009) have employed a grouping criterion to identify the fragments that belonged to the same single filament. The largest filament is labeled ‘1’ and is compared with all the other fragments of the same image that lie within a 40 pixel

distance (Gao *et al.*, 2002). According to this criterion, any filament fragment found close to the one labeled ‘1’ is also labeled ‘1’. Again this fragment is compared with all other remaining fragments to check their distances. Once this fragment ‘1’ is compared with all other fragments, the next largest fragment is labeled ‘2’. The process is repeated until all the fragments are finished. The spatial relationship between the regions that need to be merged is important but the orientation relationship between these regions is also important and it is not considered here. Additionally the algorithm has used a constant value for the distance between the fragments; thus it is not considered as fully automated.

Mathematical morphology closing is used in (Yuan *et al.*, 2011) to connect broken filaments using a disk-shaped SE; the size of the SE used was 0.5 percent of the radius of the solar disk according to their experiments. Again using a simple morphological closing operation to achieve the merging process could be useful for the filaments that are close to each other.

These previous merging methods could be classified into two groups. The first group has adopted the spatial (distance) and orientation (angle) criteria and uses a few constant values. The shortcoming of the second group is the extensive use of morphological operations which could consume additional computational time. None of these current state of the art techniques can be said to be suitable for fully automated use.

2.5 Tracking Process

Object tracking is a general imaging term which represents an important task in the field of computer vision. Tracking can be defined as the problem of estimating the path of a moving or moved object in the image plane as needed for further analysis (Yilmaz *et al.*, 2006). The lack of automatic tracking catalogues and the need for automatic object analysis and representation; the increasing power of computers and the availability of high quality instruments, and the need to know the reasons behind the

occurrences of CMEs - these offer a reason or an opportunity to develop automatic tracking methods for solar features and activities that may affect life on Earth. Thus the accurate tracking of solar filaments is imperative for the reliable prediction of geomagnetic storms and the near-earth space weather as they affect the communication networks and power grid systems of our increasingly networked planet.

The majority of current tracking methods that I have looked at appear to have some deficiencies in that they do not offer a complete study which indicate and explain the stages of tracking and the evaluation of the tracking results accurately. Some of the filaments that are classified as being disappeared have actually not disappeared, but have simply not been detected because the applied detection technique is not sensitive enough. These classes of miss-detected filaments have not been considered in the majority of previous research (Gao *et al*., 2002,Shih and Kowalski, 2003,Qu *et al*., 2005,Fuller *et al*., 2005,Bernasconi *et al*., 2005,Qahwaji and Colak, 2005, Aboudarham *et al*., 2008, Joshi *et al.*, 2009 and Yuan *et al.*, 2011).

One of the early studies in the area of detection of filament disappearances was made by Gao (Gao *et al.*, 2002). In this study, the results of two consecutive images are compared after the filament areas are obtained in individual images. The method takes into consideration the rotation of the Sun during the time elapsed between the two solar images when they were taken. Solar rotation can take the disappeared filaments behind the limb, so in this method any filament that was out of the view is excluded from the detection process. Also an extra threshold is set to control the vertically movement of the solar filaments. A "size threshold" for disappeared filament area is set to 750 $arcsec^2$. Additionally a "size threshold" for filament area is set to 250 $arcsec^2$ to exclude small filaments from detection. Many constant values are used to control the vertically movement of solar filaments, for restricting the size of the disappeared filaments and the size of the small filaments that should be excluded.

A component matching method was used by Qu (Qu *et al.*, 2005) to detect filament disappearances. The results of two consecutive images were matched after considering a differential solar rotation correction (Freeland and Handy, 1998). After dilating the spine with a 20×20 structuring element, a matching process is implemented according to the size of the spines and the intensity of the filaments. If the spine size of the current day filament is greater than 40% of the previous day and the mean intensity value of the current day filament is less than 1.5 times of the previous day, then the filaments are considered as matched. Unmatched filaments are reported as disappeared. Constant values were used in the method to control the size of the filament spine.

Another filament tracking scheme was developed by Bernasconi **(**Bernasconi, *et al.*, 2005). This method is based on tracking the day-by-day evolution of the filaments and then composing a filament tracking table that includes all the relative attributes of a tracked filament. The method uses two main thresholds for achieving the comparison process while comparing the tracking tables for three consecutive days. The first limits the search area and the second bounds the possible final search region. The process starts by reading the location (latitude and longitude) of each filament in the first table and compares that with the entries in the second table within a 5° circle from the predicted location. If no match is found then the search process continues in the third table. The search process is extended up to three days or until the predicted locations falls beyond the 60° filament detection limit. If after three days the filament is not found, it is considered as disappeared. Two latitude circles have been adopted to restrict the search results between the tables of the different days within 5° circle and to restrict the search process to a latitude circle of 60° which will prevent the detecting of the disappearance of prominences.

Carrington maps are used to achieve the tracking process by Aboudarham (Aboudarham *et al.*, 2008) where all the detected filaments during each solar rotation

are plotted by applying a region growing technique on those plots. This study needs to be validated as expressed clearly by the authors themselves.

An automatic tracking method based on a labelling criterion has been described by Joshi (Joshi *et al.*, 2009). Filaments of the new images are compared with filaments of the previous images; if they are found to lie within a distance of 15 pixels from each other, the filament label of the original image is assigned to the corresponding filament of the new image. This 15 pixel search area was the major limitation of Joshi's work.

2.6 Conclusions

Although there have been tremendous efforts over the past decade in the area of automatic detection of solar filaments, it can be shown from the previous description of the three fundamental classes of digital image processing operations that there still exist key challenges in all these classes. These include: achieving fully automatic system implementation for filament detection, merging broken structures and detecting filament disappearances; consideration of the existence of miss-detected filaments that still need to be identified and detected correctly; the implementation of a real-time system; the use of spatial and orientation considerations for the merging process instead of simple morphological closing operations; a faster implementation of the filament spine process than previous methods; finally, the extensive use of morphological operations for achieving a more optimal merging process. To reach the ultimate goal of this research, which is implementing an automated filament tracking system, the process should run through several related phases: image pre-processing; image segmentation; solar filament detection; characterizing these filaments; merging broken structures; and detecting filament disappearances. Many filament studies during the last decade have covered these phases by using different image processing techniques, including (Gao *et al*., 2002, Shih and Kowalski, 2003, Qu *et al*., 2005, Fuller *et al*., 2005, Bernasconi *et al*., 2005, Qahwaji and Colak, 2005, Aboudarham *et al*., 2008, Joshi *et al.*, 2009 and

Yuan *et al.*, 2011). However, for the last phase few investigators have actually implemented a fully automatic detection of filament disappearances: (Gao *et al*., 2002, Qu *et al.*, 2005, Bernasconi *et al.*, 2005, Aboudarham *et al*., 2008 and Joshi *et al.*, 2009).

Regardless of the number of studies undertaken; all of them look to solve the two main challenges of implementing an automated system and achieving this in real-time. This makes their implementation and development more difficult.

The following chapters discuss the stages involved in implementing a fully automated filament detection system in real-time. Chapter Four introduces an improved image segmentation algorithm and describe the implementation of the filament spine representation. Chapter Five describes the merging technique that utilizes the findings of the spine extraction to merge the broken structures using a NN. Chapter Six describes the implementation of a tracking technique which includes the filament detection stage and the creation of an HCM. A NN is used to categorize the detected filaments as DFs or MDFs.

CHAPTER THREE

3 SOLAR DATA

3.1 Introduction

This chapter reviews the two data types that were used in the practical work of this thesis. These types include solar images that are downloaded online from observatories such as the Meudon Observatory[1], included in the Global High Resolution H-alpha Network (GHN)[2] that is operated by the Space Weather Research Lab (SWRL)[3]. These images are used throughout the different phases of the present algorithms. The last phase, filament tracking, makes use of another type of solar data. These are taken from the data catalogues: the NGDC[4] and filament disappearance catalogue of SWRL.

This chapter is organized as follows: Section 3.2 identifies the solar images used in the different methods of this thesis. The catalogues of filament disappearances are described in Section 3.3. Conclusions on the contents of the different observatories and catalogues are discussed in Section 3.4.

3.2 Solar Images

Generally, solar images are obtained from many ground-based and space-based observatories. Although the space-based observatories being above the Earth's atmosphere avoid the effects of atmospheric turbulence and instability, they can more specifically be used to observe at particular wavelengths of the electromagnetic spectrum (like infrared, ultraviolet, x-ray, and gamma rays) that are impossible or difficult to observe using ground-based observatories. This is because they are either

[1] http://www.obspm.fr/
[2] http://swrl.njit.edu/ghn_web/
[3] http://swrl.njit.edu/
[4] http://www.ngdc.noaa.gov

absorbed by the Earth's atmosphere or are blocked by some other mechanism. However, ground-based observatories are cheaper to build and easier to maintain than space-based observatories. They have also been improved to get clearer solar images using adaptive optical compensation techniques to combat the turbulence of the atmosphere. These ground based observatories are used to monitor the Sun in specific segments or windows of the electromagnetic spectrum (like radio and visible light).

All the solar images that were used in this research are full-disk solar images (spectroheliograms) which were observed at the Meudon observatory and downloaded from the Global High Resolution H-Alpha Network observatories**.** The Spectroheliogram is an instrument designed to produce monochromatic images of the Sun in various wavelengths. The Meudon spectroheliograms include Hα images, Ca II K1 images, Ca II K3v images and Ca II K3p images. The Spectroheliograph provides images of the solar photosphere via the K1v images and of the solar chromospheres via the K3 and H-alpha solar images. These data are acquired once a day. The four types are shown in Figure 3.1. The Ca II K observations are very sensitive to the presence of regions with very strong magnetic fields like active regions and sunspots (Qahwaji, *et al.*, 2005), (Ermolli *et al.*, 2009).

The chromosphere can be observed using the very narrow spectral band of light known as the hydrogen-alpha line. The Hα line is a strong spectral line (high absorption) which has a wavelength of 656.3 nm (red light)**.** In this line we can see up to 1700 km above the visible layer (National Solar Observatory, 1996).

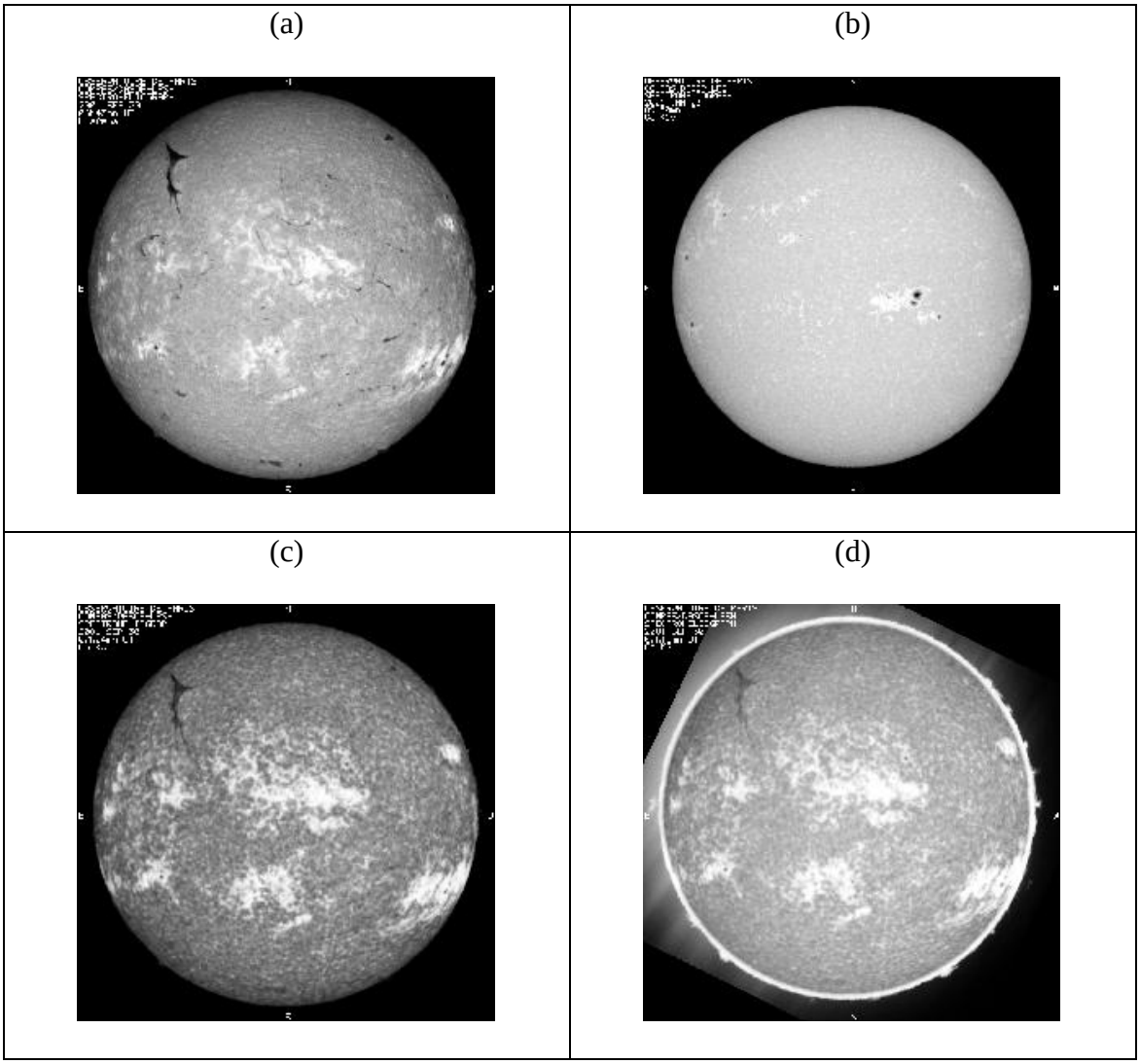


Figure 3.1. The four types of Meudon Spectoheliohraphs observed on 30th September 2001. (a) *Hα* spectroheliogram observed at 06:43:00. (b) *K1v spectroheliogram* observed at 09:45:00 (c) *K3 spectroheliogram* observed at 07:04:00. (d) *K3 (prominence) spectroheliogram* observed at 07:15:00.

The spectroheliograms in which filaments are best seen are Hα images, as shown in Figure 3.2.a, so that these are used in this study. Figure 3.2.b shows a synoptic map manually constructed at the Meudon Observatory to record the locations of the observed filaments.

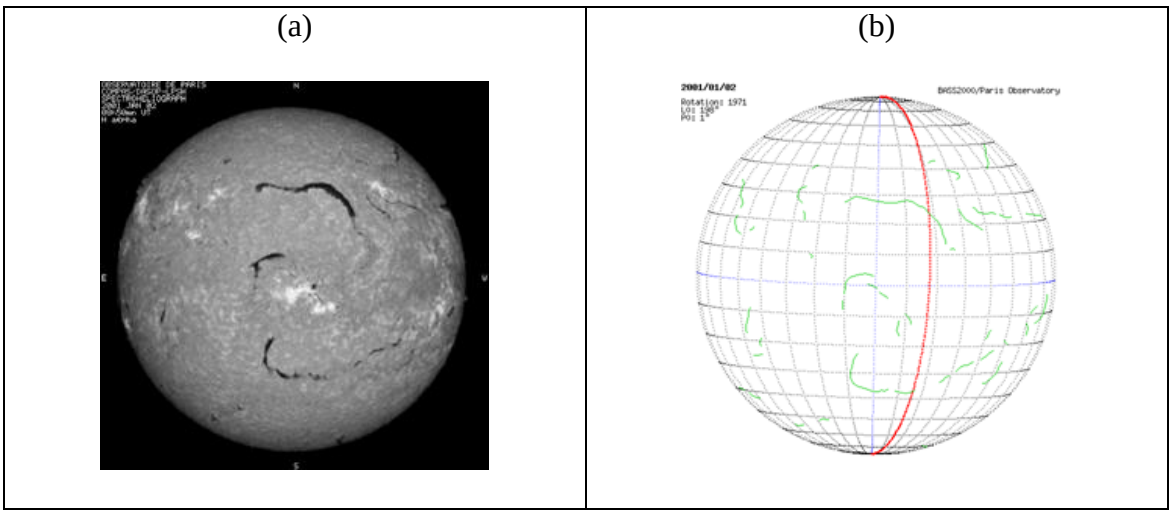


Figure 3.2. (a) Hα image observed at Meudon Observatory on 02-Jan-2001 08:50:00. (b) The synoptic map for the solar image of (a).

In the segmentation phase of this system, a False Acceptance Rate (FAR) (Hong and Jain, 1997) is estimated by comparing the detected filaments with those detected manually and recorded in the synoptic maps. Figure 3.3 illustrates the different layers of the Sun: the corona, the chromosphere and the photosphere.In addition to filaments, Hα images show other solar features such as: plages, short-lived solar flares, sunspots and elongated filaments. Filaments in these images appear obviously as dark ribbons against their brighter background.

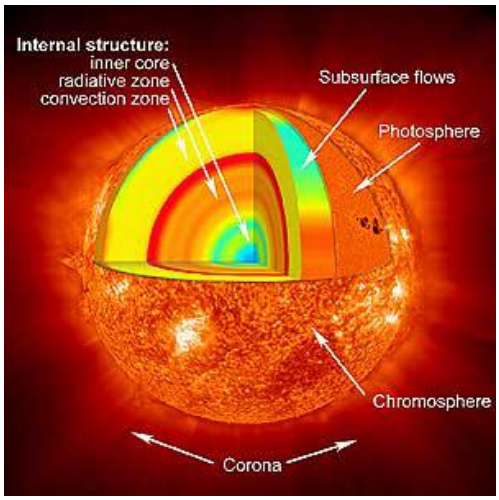


Figure 3.3. Layers of the Sun. NASA.

The general format of solar image filenames from the Meudon Observatory[5] is composed of three parts, image type, date and time. An image filename is shown in the example below.

Example: mh010102.085000.gif

Starting from the left, the first two characters, mh, represent the type of the image (mh for Hα images (chromosphere), mk for Ca II K1 images (wing of the line, photosphere) and for Ca II K3 images (centre of line, chromosphere) and mp for Ca II K3 prominences). The second part, 010215.082336, corresponds to the image date and time (the first part of it is the date in the format (YYMMDD); note that the first two numbers are the last two digits of the year, whilst the second part is for the time, with the format (HHMMSS)). The last part of the image filename, gif, is the image file format.

The global high resolution H-alpha network includes the set of H-alpha stations which are managed and maintained by the SWRL. These stations include: the Big Bear Solar Observatory (BBSO) in California, the Kanzelhöhe Solar Observatory (KSO) in Austria, the Catania Astrophysical Observatory (CAO) in Italy, Meudon and Pic du Midi Observatories in France, the Huairou Solar Observing Station (HSOS) and the Yunnan Astronomical Observatory (YAO) in China, the Mauna Loa Solar Observatory in Hawaii, and the Uccle Solar Equatorial Table (USET) in Belgium. These stations monitor the solar activity of the chromosphere twenty four hours a day. The data set used here is mainly collected from the: BBSO, Kanzelhöhe solar observatory (KNZ) and Observatoire De Paris, Section De Meudon (MEUDON) observatories.

The SWRL was established in September, 2008 by the New Jersey Institute of Technology (NJIT)[6]. SWRL focuses scientific research in the area of space weather. Its mission is to understand the magnetic activities of the Sun and their effects on the near-

[5] http://bass2000.obspm.fr/data_guide.php?what=all#spectro_general
[6] http://www.njit.edu/

Earth environment. Its tools are being developed to monitor, understand and forecast solar activity and its geomagnetic effects. The reason behind the use of this network in the present study is the availability of successive images from its different observatories that gives benefit in the process of tracking.

The SWRL network follows the SOHO convention for file naming[7], except that the observation time will have six characters with the format HHMMSS instead of four characters and the miscellaneous info code will be modified to represent different spatial resolution and calibration data. The file name will consist of 33 characters (including four underscore characters and a dot) as shown in the example below.

Example: bbso_halph_fr_20010216_164558.jpg

Reading from left to right: the first four characters, bbso, are the institution code, the next five characters, halph, are the image type code, the next two characters, fr, are the miscellaneous info code (all such codes are shown in Table 3.1), the next eight characters, 20010216, are the observation date in yyyymmdd format (16/02/2001), the next six characters, 164558, are the observation time in hhmmss format (16:45:58) and the last three characters, jpg, represent the file type.

Table 3.1. BBSO Miscellaneous Info Code.

Code	Description
FI	Full disk raw image (Singer or photometric telescope).
FL	Full disk image dark subtracted and flat field corrected.
FR	Full disk image limb darkening subtracted.
FF	Full disk flat field frame.
FD	(Full disk dark frame).

[7] http://www.mssl.ucl.ac.uk/grid/iau/extra/local_copy/BBSO_SOHO_filenames.html

The Singer telescope mentioned in Table 3.1 was built by Boller and Chivens under contract to the Link Division of Singer-General Precision in New York; it is from here the name Singer originates (Denker, *et al.* 1998). Corrections to the raw sensor data are required because the camera pixels have non-zero responses to no light (dark) and different sensitivities to light (flat field). Separately, the brightness of the solar disk decreases towards the limb due to increased absorption of light in the solar atmosphere (limb darkening). In the work presented here, BBSO images with the FR miscellaneous info code, which are corrected for dark, flat-field and limb-darkening have been used.

3.3 Data Catalogues

Solar catalogues are lists or tabulation displays which have been produced for several aspects of solar activities over many years. Most of these catalogues are available online in electronic format and can be freely downloaded. Solar filament catalogues are intended to provide values for the features of these filaments. A related type is catalogues of solar filament disappearances, which are the subject of the discussion in the next subsection. They are available in two existing online catalogues, which are dependent on the NGDC and SWRL ground-based observatories.

3.3.1 Filament Disappearances Catalogues

Two different catalogues were used in this study for filament tracking, the last phase of this research. These are the NGDC[8] and the SWRL[9] catalogues. The former catalogue provides near complete detail about different solar features and activities; in particular, it tabulates the Solar Filament Disappearances as text files, like that shown in Table 3.2. This table lists all DF events detected in January 2001. The results of filament disappearances detected by the algorithm were evaluated against the contents

[8] ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SOLAR_FILAMENTS/, last access: 2012.
[9] http://swrl.njit.edu/ghn_web/

of the NGDC manual catalogue. The contents of this catalogue are classified and arranged based on human observation of the filaments.

Table 3.2. The solar filaments and prominences as tabulated by NGDC for January, 2001.

Date	Start Time	End Time	Position	Event Type	Extent	CMP Date	End Time		NOAA	SSN	Quality
77010117	1404	1424	S14E90	EPL1	E 99	10124.4				1	2SVTO
77010119	2137U	1141U	S18E09	DSF	E0800	10120.6	S22E16			1	3RAMY
77010121	1912	1917	S07E40	DSF3	E0700	10124.8	S10E37	S10E36	9313	1	3RAMY
77010124	1616U	1205U	S14W47	DSF	E0800	10121.1	S18W40			1	3RAMY
77010124	1616U	1205U	N13W13	DSF	E1600	10123.7	N20W15	N20W25		2	3RAMY
77010129	2322U	1534U	N29E12	DSF3	E2400	10130.9	N19W10			1	3HOLL
77010131	1536	1544	S22W80	BSD	E 00	10125.5	S23W86		9316	1	3HOLL
77010131	1536	1544	S77W21	BSD	E1100	10129.7	S89W23		9316	1	3RAMY
77010204	1326U	0634U	N24E55	DSF2	E0800	0102 8.8	N20E61			1	3SVTO
77010209	2137U	1212U	N15W29	DSF	E1400	0102 7.7	N01W31			1	3RAMY
77010209	2137U	1212U	N11E27	DSF	E0900	10211.9	N20E24			2	3RAMY
77010209	1519U	0939U	N29W11	DSF	E0900	0102 8.8	N27W03			1	3SVTO
77010209	1519U	0939U	N27E13	DSF	E0600	10210.6	N24E18			2	3SVTO
77010215	0018U	1407U	N20E20	DSF	E1500	10216.5				1	3HOLL
77010215	0944U	2317U	S25E17	DSF	E0700	10216.7	S18E20			1	3LEAR
77010215	0949U	2346U	S45E21	DSF2	E2500	10217.1	S35W02			1	3LEAR
77010215	1954U	1444U	S27E02	DSF	E1700	10216	N18E13			1	3RAMY
77010216	0944U	2317U	S25E17	DSF	E0700	10217.7	S18E20			1	3LEAR
77010216	2113U	1138U	S50E36	DSF	E0300	10219.9	S48E39			1	3RAMY
77010216	1539U	0627U	S27E11	DSF	E0800	10217.5	S21E16			1	3SVTO
77010217	1928U	1444U	S25E08	DSF3	E0800	10218.4	S19E16			1	3HOLL
77010217	0947U	2311U	N38W38	DSF	E0500	10214.3	N39W33			1	3LEAR
77010219	2027U	1700U	S35W23	DSF	E1000	10218	S39W13			1	3RAMY
77010226	1912U	1214U	S16W10	DSF	E1200	10226	S17W05	S22E00		1	3RAMY
77010226	1545U	1213U	S16W09	DSF	E1100	10226	S21E01			1	3SVTO
77010228	1146	1323U	S17W05	DSF1	E1100	10228.1	S25W13			1	3RAMY
77010228	1734U	1113U	S30E40	DSF	E0900	0103 3.9	S22E42			2	3RAMY
77010301	0946U	0039U	N09W21	DSF2	E250000	10227.9				1	4LEAR
77010301	1023U	2302U	S33E18	DSF	E0800	0103 2.9	S30E26			1	3LEAR
77010305	1025E	0	N22W44	LPS	E 00	0103 2.0			9371	1	3SVTO
77010305	1025E	1307D	N21W44	LPS	E 00	0103 2.1			9371	1	3SVTO
77010306	1011U	2301U	S33E26	DSF	E0700	0103 8.5	S27E22			1	2LEAR
77010307	1015U	2304U	N40E63	DSF2	E2100	10312.5	N29E45			1	3LEAR
77010309	0958U	2316U	S05E18	DSF2	E 0000	10310.8				1	3LEAR
77010311	2333	6	S12E80	EPL3	E 00	10318			9376	1	3HOLL
77010311	2006U	1119U	S55E04	DSF	E2200	10312.2	S56E27			1	3RAMY
77010311	1430U	1356U	S56E03	DSF	E2400	10311.9	S56E27			1	3SVTO
77010312	0030U	1338U	S60E05	DSF	E2500	10312.5	S60E30			1	3HOLL
77010313	2329U	1411U	S41W29	DSF	E0900	10311.6	S34W34			1	3HOLL
77010313	2111U	1141U	S27W40	DSF	E1700	10310.8	S40W31			1	3RAMY
77010314	625	1158	S37W38	DSF3	E1000	10311.2	S41W45	S43W40		1	3SVTO

The source images used in this catalogue are from the Solar Observing Optical Network (SOON) in Boulder, Colorado, USA. This network is an American Air Force Global Network for ground-based solar observatories (U.S. Air Force Fact Sheet, 2010). The first column of Table 3.2 shows the data code which always starts with 77 and is

followed by the estimated date (YYMMDD) of filament disappearances. The second column shows the estimated time of disappearances in hours and minutes (or start of a data gap). The third column indicates the uncertainty in time of disappearance in hours and minutes (or end of data gap). The start and end time in this catalogue are sometimes accompanied by the three letters, D (After), E (Before) and U (Uncertain). Column four indicates the location of the active filament centroid with N or S for North and South latitudes, E or W for East or West central meridian distance. The type parameter in column five classifies filaments into one of the fifteen different types shown in Table 3.3. Column Six contains for limb events; the radial extent above the limb while for disk events it contains the heliographic extent in whole degrees. The CMP date is the Central Meridian Passage Date which is shown in column seven. When a filament is longer than 10 degrees, the locations of the ends are also given in column eight. Column Nine shows the NOAA/USAF region number, if it is known. SSN is the Station serial number for this region and is shown in column ten. The final column contains the Quality (1 = poor to 5 = excellent), in addition to the station name.

Table 3.3. Filament Types.

Type	Description
SSB	Solar Sector Boundary
MDP	Mound Prominence
CRN	Coronal Rain
CAP	CAP Prominence
LPS	Loops Prominence System
SPY	Spray
BSD	Bright Surge on Disk
APR	Active Prominence
DSD	Dark Surge on Disk
ADF	Active Dark Filament
ASR	Active Surge Region
AFS	Arch Filament System
BSL	Bright Surge on Limb
EPL	Eruptive Prominence on Limb
DSF	Disappearing filament

Another tabulated form for filament disappearances, like that shown in Table 3.4, was created by SWRL. This catalogue lists features of the disappeared filaments, which include: in the first column, the last appearance time (the last date and time the filament appeared); in the second column, the first disappearance time (the first date and time the filament disappeared); in the third column the position (the location of the disappeared filament in latitude and longitude); and finally in column four, the filament size (area).

Table 3.4. Part for the filament disappearance, tabulated by SWRL for Year 2005.

Last Appearance Time	First Disappearance Time	Position(latitude,longitude)	Size(square arcsec)
05/08/2005 15:26	05/08/2005 22:06	W2N33	3397
05/08/2005 22:06	08/08/2005 17:49	W17N18	3947
10/08/2005 16:50	10/08/2005 18:50	W40N19	470
10/08/2005 20:50	10/08/2005 22:50	W51N15	780
17/08/2005 15:24	17/08/2005 17:23	E51N20	719
18/08/2005 23:24	19/08/2005 15:29	E24N52	1407
19/08/2005 19:29	19/08/2005 21:29	W9N1	1103
29/07/2005 17:24	29/07/2005 23:16	W12N29	807
28/07/2005 22:07	29/07/2005 15:23	W26S47	771
28/07/2005 22:07	29/07/2005 15:23	W5N16	619
28/07/2005 22:07	29/07/2005 15:23	E24S20	1845
27/07/2005 23:17	28/07/2005 15:22	W7S20	464
26/07/2005 20:36	27/07/2005 18:37	E19N19	1303
26/07/2005 20:36	27/07/2005 18:37	E5N12	1782
25/07/2005 17:41	25/07/2005 22:23	E29S17	841
25/07/2005 17:41	25/07/2005 22:23	E44N17	1040
21/07/2005 22:21	22/07/2005 15:29	W15N13	860
19/07/2005 17:29	19/07/2005 22:15	E1S15	831

3.4 Conclusions

Some of the solar images used in this work were downloaded from the Meudon observatory, but images were found to be missing from the solar survey archive over certain time periods. Other solar images were downloaded from the Global high resolution H-alpha network of SWRL. This network includes observations for around 30 years obtained from several observatories spread round the globe. Hα images were downloaded from some of these stations in different formats with different miscellaneous information codes. The one most suited for the various phases of this study is the FR miscellaneous information code - which is dark and flat-field

corrected and with limb-darkening correction. Although this network collected solar images from different solar stations; there are still some images which are missing for certain time periods. The first of two catalogues that record filament disappearances is the SWRL filament disappearances catalogue (Yuan Yuan. yy46@njit.edu. Solar Filament Disappearances. 17 November 2011). This represents results for the time period from 1991 to 2005. These results are displayed within nine pages: each containing fifty disappeared filaments, to give a total of four hundred and fifty disappeared filaments over the fifteen years. The second catalogue is the NGDC manual catalogue which is used in the current study as a reference for validating the number of disappeared filaments. Its tables list all the DF events detected between 1991 and 2012 (C D Slisser. Justin.Mabie@noaa.gov. Catalogue of Solar Filament Disappearances. 15 July 2011).

CHAPTER FOUR

4 AUTOMATED ALGORITHM FOR DETECTING SOLAR FILAMENTS

4.1 Introduction

Detecting and characterizing solar filaments are important for several aspects of solar activities because of their association with the occurrences of CMEs – these are major solar eruptions that could cause geomagnetic storms on Earth. Efficient detection systems should be fully automated and work in real-time, which makes their implementation and development harder.

This chapter introduces an automated detection method for solar filaments that has avoided using any empirical values to produce a fully automated technique. Definitely, the time-factor will affect synthesizing a filament detecting and tracking system to be part of a proposed real-time system. This factor could prevent the system from producing real-time space weather alerts and quick look-up results. Significantly, the accuracy of detecting the spines of filaments plays an important role in identifying its main attributes under consideration and as well as better achieving the subsequent image post-processing tasks.

This chapter is organized as follows: Section 4.2 introduces an improved image segmentation algorithm. The detection algorithm filament boundary extraction and spine detection techniques are described in Section 4.3. Section 4.4 presents the computational demands of the proposed technique. Conclusions on the associated findings and comparisons with previous work are discussed in Section 4.5.

4.2 Solar Image Segmentation

The solar image segmentation stage is still a challenge which can be addressed by the adaptive detection of solar features. A more accurate implementation would play a key role in recognizing and detecting the features with higher probability. In this section, a modified and enhanced segmentation phase, based on the pre-processing and cleaning technique proposed in (Qahwaji and Colak, 2005), to improve their segmentation results is described. This adaptive local thresholding (ALT) technique depends on sliding two windows over the whole image. In this technique, a selection is made to classify the pixel contents of the enhanced image (EI) into potential filament pixel or background pixel based on the instructions in 4-1 (Automated detection of masses in mammograms Kom *et al*., 2007):

For i=1 to m do
For j=1 to n do
If(EI(i,j)>TH(i,j) and $R_{LW} > \bar{x}_{SW}$*)then*
EI(i,j)∈ Candidate filament area
Else *(4-1)*
EI(i,j)∈ Normal area
Endif
Endfor j
Endfor i

Where m × n is the size of image, subscripts *LW* and *SW*, indicate large and small windows centred on pixel at *i, j,* $TH(i, j)$ is an adaptive threshold value that is calculated by the formula shown in Equation 4-2, R_{LW} is the range of values in the large window also shown in Equation 4-2, $\bar{x}_{SW}$ represents the average pixel intensity in the small

window. $EI_{\max}(i,j)$ and $EI_{\min}(i,j)$ are the maximum and minimum intensity values respectively within the large window.

$$TH(i,j) = \bar{x}_{SW} - R_{LW}$$

and *(4-2)*

$$R_{LW} = EI_{max}(i,j) - EI_{min}(i,j)$$

From experiment, the dimensions of the large window were chosen to be 17×17 pixels and those of the small window were chosen to be 3×3 pixels. These sizes gave the best trade-off between accuracy and speed of execution. The whole algorithm for segmenting the solar image can be summarised as shown below:

Segment_Algorithm( )

{

For each pixel:

Max:= maximum intensity value in the large window

Min:= minimum intensity value in the large window

R_{LW}: = Max-Min

$\bar{x}_{SW}$: = average of intensity value in the small window

TH: = $\bar{x}_{SW} - R_{LW}$

If ((Pixel Intensity value > *TH*) AND (R_{LW}>$\bar{x}_{SW}$)) then

The Pixel belongs to the region of interest

Else

The Pixel belongs to a non-region of interest

}

The algorithm was applied to Hα solar images like the one shown in Fig 4.1.a. The segmented filaments were tested by comparing the resultant image with the manually constructed synoptic maps like that shown in Figure 4.1.b. The maps contain all the

solar filaments seen on a given day. Figure 4.1.a represents the enhanced image that is used as input to the segmentation method. Some results from applying the technique are shown in Figure 4.1.c. The primary goal of all solar filament segmentation techniques is to obtain well defined filaments, and a low FAR.

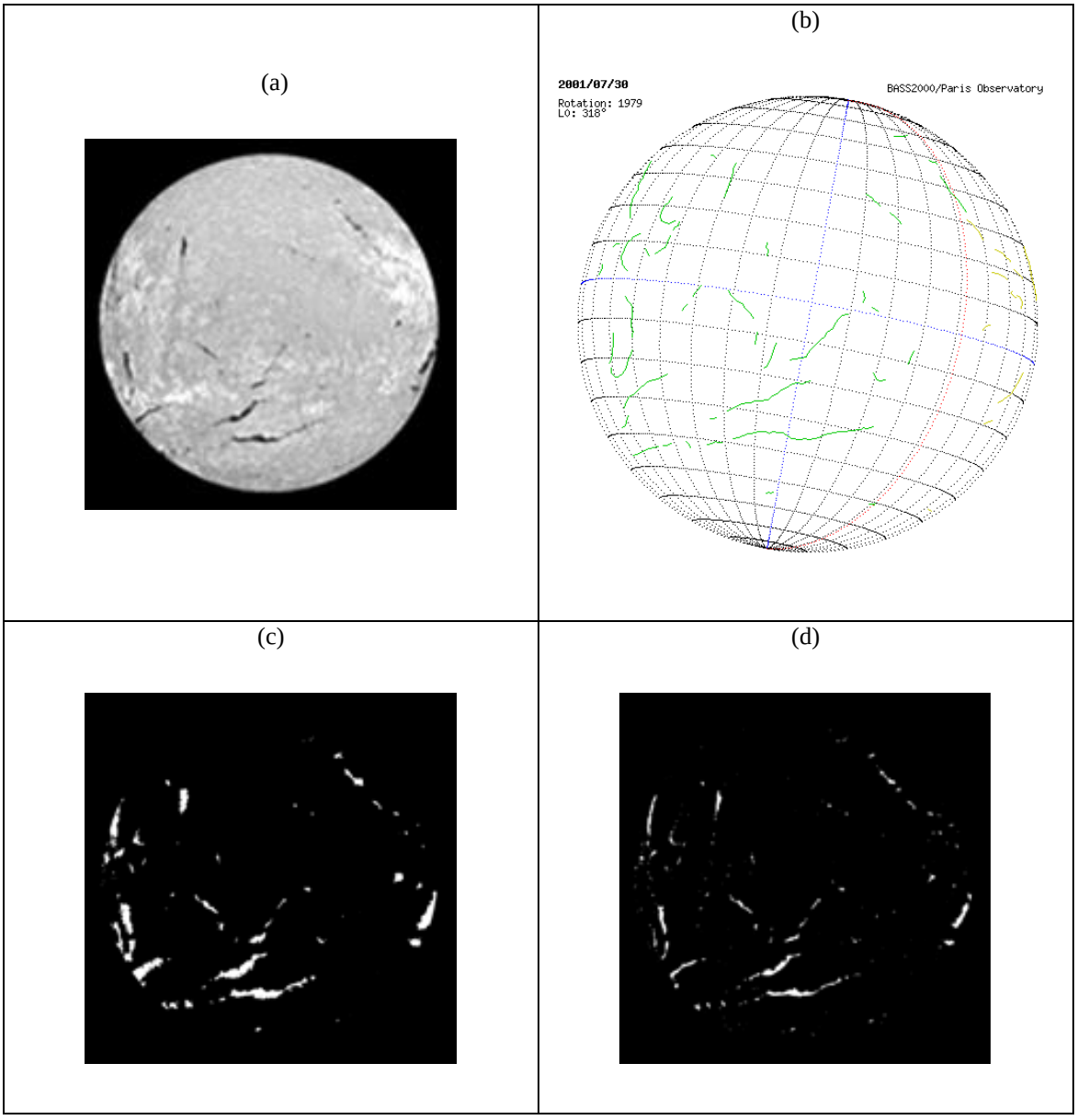


Figure 4.1. Results of applying the ALT technique. (a) Input Image (Enhanced Image), original Hα image observed at Meudon observatory on July 30, 2001. (b) Synoptic map for the Hα image shown in (a) downloaded from Meudon Observatory. (c) The output image after Applying ALT. (d) The output image after applying ALT&V on the same image.

The performance of the detection algorithms are evaluated using the FAR measure (Hong and Jain, 1998) which is the probability of a non-region of interest (non-RoI) being detected as a RoI. The algorithm is compared with the results of the Adaptive Local Thresholding and Verification (ALT&V) which is presented by Qahwaji (Qahwaji and Colak, 2005) as shown in Figure 4.1.c and Table 4.1. It seems, by comparing figures 4.1.c and 4.1.d, that the ALT has the advantage of detecting more unambiguous filaments whilst introducing less noise than the ALT&V result shown in Figure 4.1.d.

Table 4.1. FAR Values for Synoptic Maps and ALT.

Problem	**Synoptic maps**	**ALT**		**ALT&V**	
	Filaments	**Filaments**	**FAR (%)**	**Filaments**	**FAR (%)**
02/07/2001	44	28	0	44	15
03/07/2001	45	33	0	45	5
04/07/2001	38	18	0	38	3
06/07/2001	50	43	10	50	18
09/07/2001	41	28	2	41	12
10/07/2001	39	21	0	39	15
11/07/2001	32	29	9	32	7
15/07/2001	32	32	19	32	31
16/07/2001	26	28	42	26	36
17/07/2001	34	20	0	34	31
19/07/2001	41	43	22	41	27
20/07/2001	36	29	8	36	24
21/07/2001	36	37	6	36	4
22/07/2001	40	46	23	40	21
25/07/2001	34	22	0	34	18
26/07/2001	37	33	11	37	9
29/07/2001	38	23	0	38	11
30/07/2001	52	36	2	30	18
31/07/2001	43	38	5	31	2
03/08/2001	46	44	22	3	29
04/08/2001	37	26	0	4	33
Average	**39.1**	**31.3**	**9**	**33.9**	**19**

Generally, the system performance requirement is specified in terms of FAR, where FAR of zero means that no non-RoI is detected as a RoI. The findings from Figure 4.1 are confirmed according to this criterion by observing the number of detected filaments and FAR values for the two algorithms, which are shown in Table 4.1. The first column shows the date of every Hα image, while the total number of filaments that are detected

manually by synoptic maps is shown in the second column. The remaining columns show the number of filaments and FAR values obtained using the ALT and ALT&V techniques. The table shows the improvement that is achieved over ALT&V by reducing the average FAR error rate from 19% to 9%, while the difference in detections is less than this. Both methods detect fewer filaments on average than shown in the synoptic maps.

4.3 Filament Detection and Boundary Extraction

In this section, algorithms used to detect the shape and size of filaments and extract their boundaries are described.

4.3.1 Filament Detection

Filament detection means extracting and identifying the actual filament area as an elongated object in Hα solar images. An important achievement in this area is the work carried out by (Qahwaji and Colak, 2005) - where processes for recognizing and verifying solar filaments and active regions from Hα images were introduced. The solar images used in this work were pre-processed (cleaned and enhanced) as described in (Qahwaji and Colak, 2005). The filling algorithm defined in this paper is also used to distinguish between the background region that lies outside the solar disk and the region that lies inside the solar disk. This was followed by applying an intensity filtering technique to detect the candidate pixels for the filaments. The process ends by applying a region growing technique to detect regions of interest that were not detected by intensity filtering. Additionally, the intensity filtering stage in (Qahwaji and Colak, 2005) is further improved by applying the adaptive thresholding technique presented by Atoum (Atoum *et al*., 2009). This method is further enhanced by developing a fully automated detection technique that includes retrieving the actual filament area from the region grown one.

The detection method starts by computing the mean μ_f and standard deviation σ_f for every filament area. Thereafter, the standard deviation σ_w, maximum value W_{max}, and minimum value W_{min} are calculated over a 5×5 pixel neighbourhood centred on every pixel of the region grown image. Then the threshold *TH*, calculated using Equation 4-3.

$$TH = MR + \alpha * \sigma_f \qquad (4-3)$$

Where *MR* is the mid-range value $(W_{max} - W_{min})/2.0$. The value of $\boldsymbol{\alpha}$ is calculated using Equation 4-4 for solar images from the Meudon observatory or using Equation 4-5 for solar images from BBSO and KANZ observatories**.**

$$\alpha = \sigma_w/MR \qquad (4-4)$$

$$\alpha = \sqrt{\sigma_f}/\sigma_w \qquad (4-5)$$

The thresholding rule shown in Equation 4-6, is then applied:

$$f(x,y) = \begin{cases} 1, & I(x,y) < TH \quad and \quad TH < \mu_f \\ 0, & Otherwise \end{cases} \qquad (4-6)$$

Where $f(x,y)$ represents the detected filament and $I(x,y)$ is the region grown filament. Figure 4.2 illustrates the main steps of the filament detection technique. The result of the whole detection process is shown in Figure 4.3.d. This technique has the advantage of being fully automated.

For *i*=1 to number of images

For *j*=1 to number of filaments in each image

Step1: Compute Filament μ_f.

Step2: Compute Filament σ_f .

Step3: Over a 5×5 window:

(i) Compute the window min and max values.

(ii) Compute the σ_w.

(iii) Compute the threshold value (*TH*) according to Equation (4-3)

(iv) If (Pixel Value <*TH* AND *TH* < μ_f) Then

Pixel Value =1;

Otherwise

Pixel Value=0;

Figure 4.2. Pseudo-code for the Detection Process.

The output image of this detection method still suffers from the presence of small unwanted pixels and holes. These should be removed without distorting the image, noting that it is impossible to remove such noise totally without distorting the image. Nevertheless it is a vital step in reducing the image noise to an acceptable level which is a fine line between removing the noise completely which will distort the image and affect the important filaments or retaining sufficient noise so as not to have too many falsely detected filaments. Morphological operators are one tool which could be used to remove these pixels and fill the small holes at the same time whilst relatively preserving the corner details.

These operators could achieve the previous task by using a combination of dilation and erosion operators. Note that, they have they the disadvantage of merging close

elements but using a small structuring element will, due to the existing small size of these ‘spots’ and ‘holes’, avoid such shortcomings.

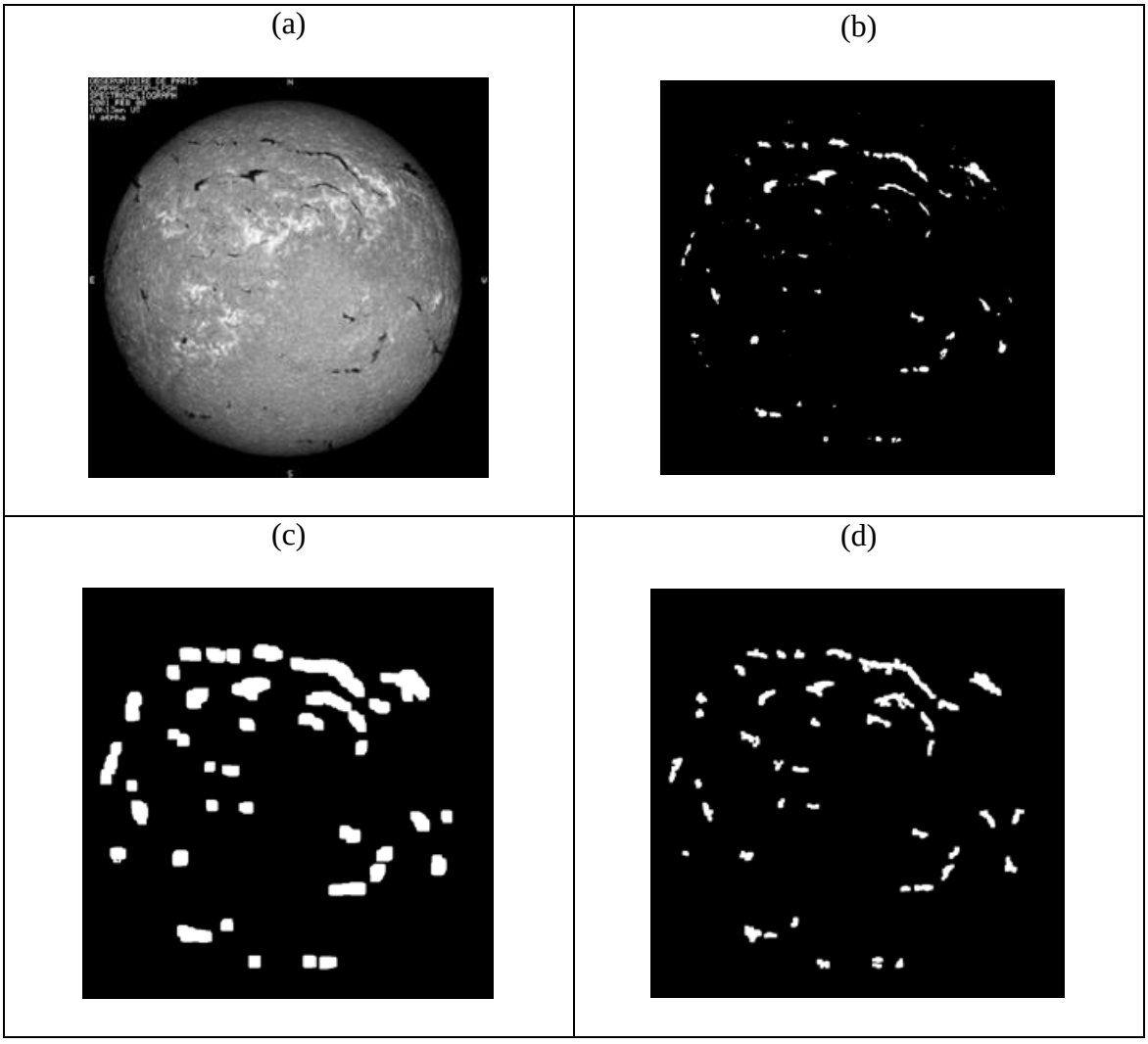


Figure 4.3. The results from the whole detection process applied to Hα image observed at Meudon observatory on February 8, 2001. (a) Original Image. (b) Segmented Image. (c) Region Grown Image. (d) Detection Results.

The opening morphological operator, consisting of erosion followed by dilation, has been implemented using a 3×3 square structuring element to eliminate noise (small unwanted pixels in the detected image) while preserving the shape and size of the larger objects. To remove other unwanted isolated spurious pixels an area thresholding operation has been set as given in Equation 4-7.

$$T = {img_{mean}}/{img_{MR}} \quad (4\text{-}7)$$

Where img_{mean} is the mean value of the whole image and img_{MR} is the mid range of the whole image and it is equal ${(img_{max} - img_{min})}/{2}$, img_{max} and img_{min} are the maximum and minimum respectively for the whole image. If the object area is less than this threshold value *T* then it will be removed. Another closing operation was implemented afterwards using a 3×3 square element to fill in the empty pixels left after the previous operations. The closing operation is a dilation followed by an erosion operation.

4.3.2 Boundary Extraction

Determining the filament's boundary is a necessity for the subsequent activities; spine extraction and filament merging. Morphological image processing is a collection of image processing techniques that deal with the structure of features in an image (Efford, 2000). There are three primary morphological techniques: erosion, dilation and hit-or-miss operators. One application of erosion operation is boundary finding. The boundary is extracted by subtracting the eroded image from the original image *r*, as shown in Equation 4-8 below:

$$g = r - (r \ominus s) \quad (4-8)$$

s is the SE, Θ is the erosion operation, and *g* is the image of the region boundaries. The size of the SE is chosen to be of 3×3 pixels in order to identify a one-pixel-wide boundary. Figure 4.4.a shows a detected filament and Figure 4.4.b shows its boundary.

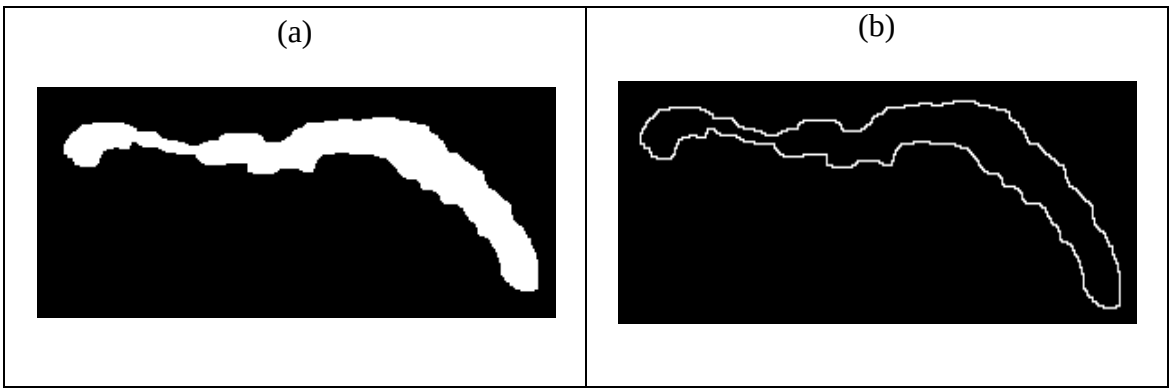


Figure 4.4. Boundary detection results by using a horizontally aligned filament from an Hα image observed at Meudon observatory on January 2, 2001. (a) Detected Filament. (b) Filament Boundary.

4.3.3 Spine Description and Extraction

The accurate detection of the filament's spine could provide accurate description for the shape, size and orientation of the filament. An accurate spine detection algorithm should satisfy the following requirements **(**Fuller *et al*., 2005; Ingrid *et al*., 2002):

1. Non-linearity: The spine should pass through the middle of local regions, following the body of the filament as a curve not as a line.
2. Morphology: The spine should retain the shape of the original feature.
3. Adaptation: The shape of the spine should match any changes in the filament shape.
4. Connectivity: The spine points should be connected.

This work does not describe the spine as a set of line segments, where a threshold can be used to control their size and number as done by Bernasconi (Bernasconi *et al*., 2005). Additionally it does not make extensive use of the morphological operations (standard thinning and skeletonization) to extract the filament spine as done by Fuller (Fuller *et al*., 2005) because this could require significant computational time.

Instead a new automated technique for detecting and extracting these spines is proposed. This new technique represents the spine as a smooth curve, which passes

through the middle of the boundary data. This produces a curve, instead of a set of line segments, represented by a set of connected points. To determine these curves, six consecutive steps are implemented for each filament as shown below and illustrated in Figure 4.5.

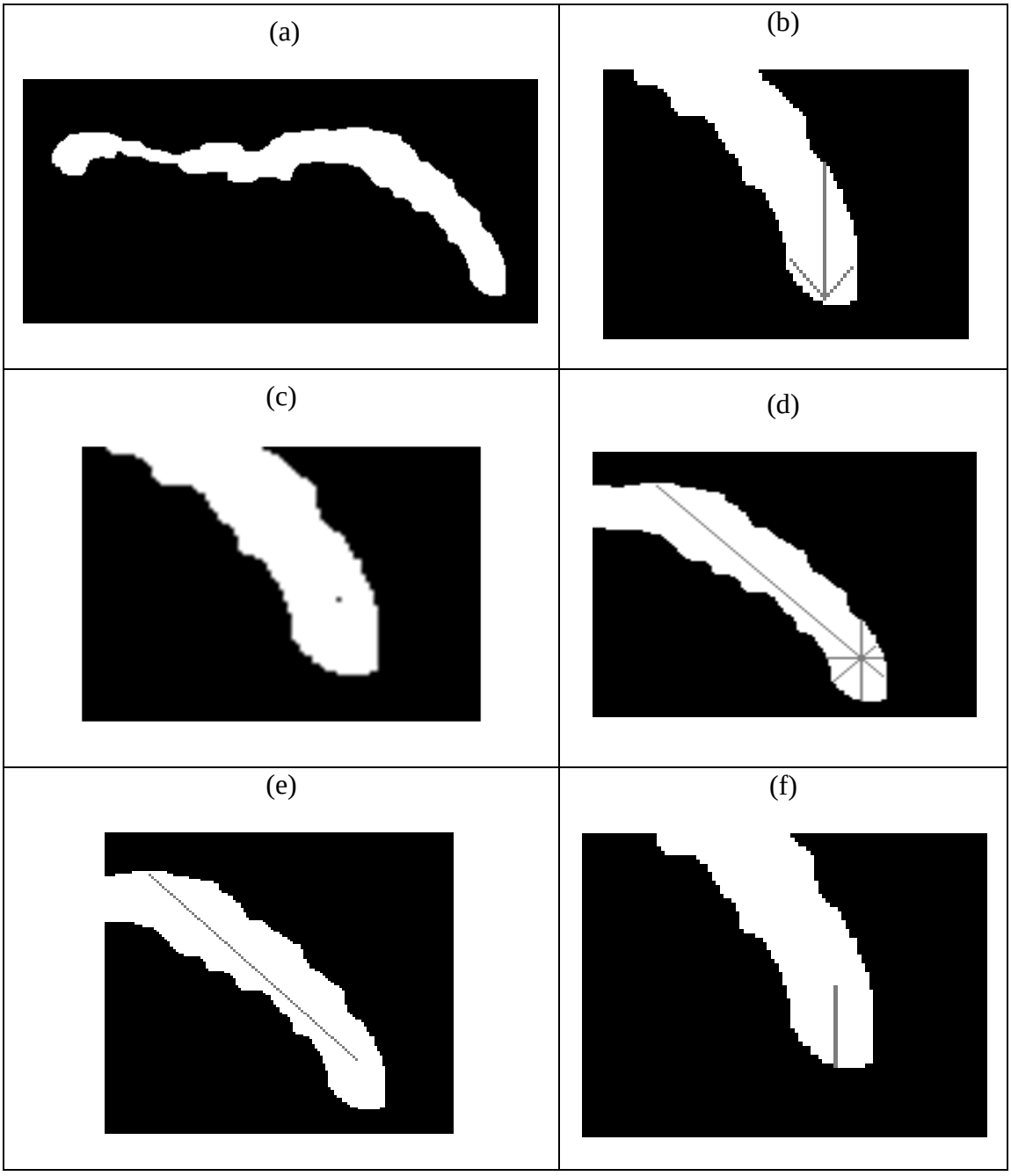


Figure 4.5. The steps implemented for spine determination. (a) Solar filament. (b) The three line segments. (c) The middle of the longest line segment. (d) The full four line segments that pass through the seed pixel. (e) The longest line segment. (f) The opposite longest one of e.

Step 1: Determine the initial seed pixel.

Step 2: Determine the starting Vertex.

Step 3: Determine the four line segments passing through this starting vertex.

Step 4: Find the longest line segment.

Step 5: Projecting and averaging the perpendicular line segments.

Step 6: Maintain the continuity of the drawn spine.

A seed pixel is first determined by moving from the lower left to the lower right boundary of the rectangle that encloses the filament like that shown in Figure 4.5.a. When a white pixel is found, three line segments are drawn from it towards the filament body as shown in Figure 4.5.b. Finally, the middle point of the longest line segment is chosen to be the seed pixel for drawing the spine, as shown in Figure 4.5.c. Moving to the next vertex requires drawing the four line segments in the directions shown in Figure 4.6 below.

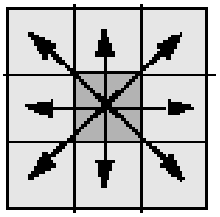


Figure 4.6. The four line directions.

These lines pass through the seed pixel (the first vertex determined in Figure 4.5.c) towards the boundary of the filament as shown in Figure 4.5.d and the longest is chosen as shown in Figure 4.5.e. This segment guides the part of the spine opposite the longest line segment as shown in Figure 4.5.f, which will determine the starting part of the spine to complete the drawing of the full spine.

The next projection and averaging phase of the algorithm proceeds as follows. For each point of the longest line segment, a perpendicular line is drawn and the mid-point found between the points where the perpendicular intersects the filament boundary, as

shown in figures 4.7.a and 4.7.b. The result of the projection process applied to the initial part of the spine is shown in Figure 4.7.c.

This process is repeated and to prevent the redrawing of the previous line segment step, the previous drawing direction is preserved as a reference for the next movement. This requires ignoring any direction that will guide the spine to go back to its previous starting position as shown in Figure 4.7.d. The final curve is shown in Figure 4.8.

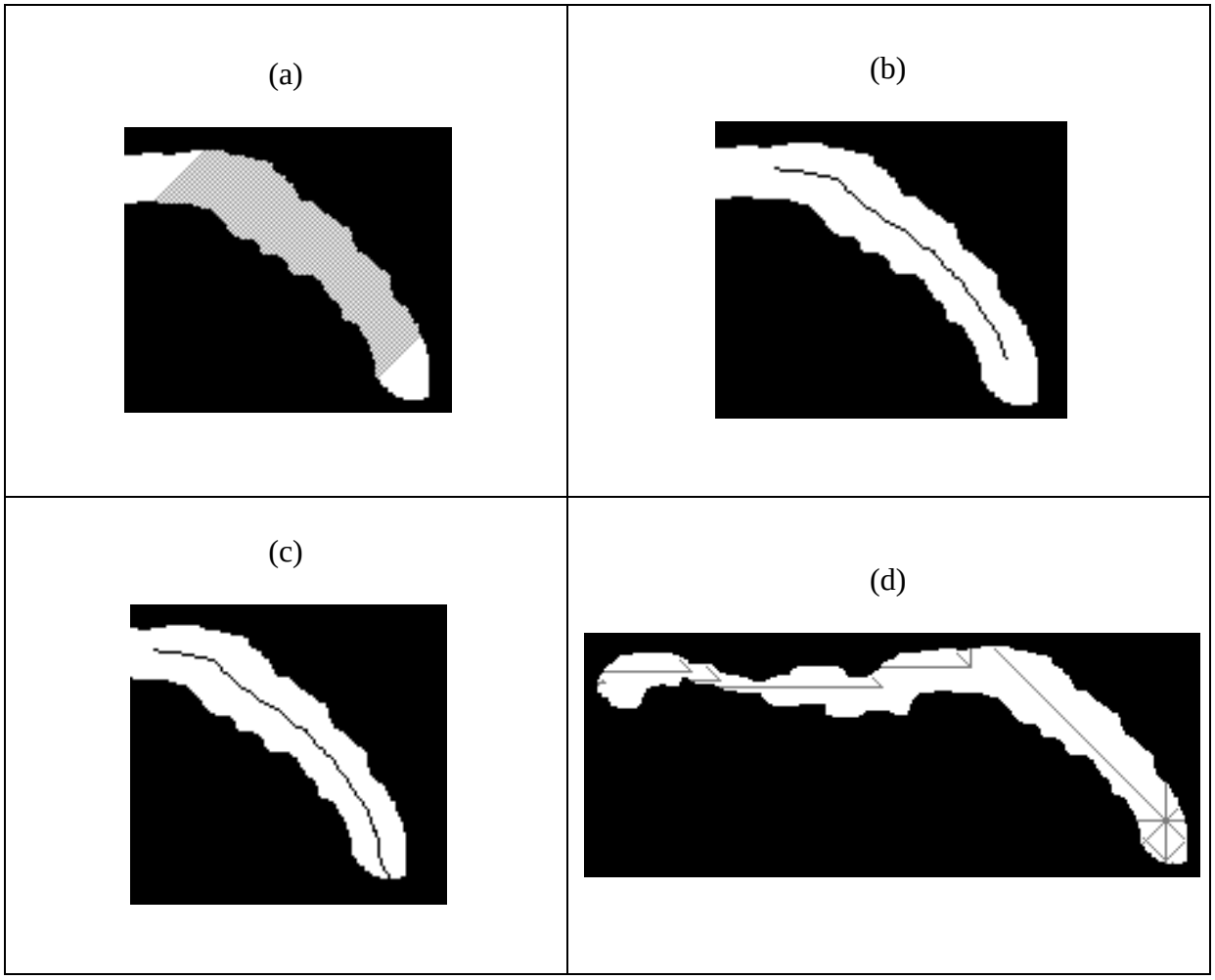


Figure 4.7. Projecting and averaging phase of determining the spine. (a) The projection result. (b) The averaging result. (c) The spine of the first longest line segment and its opposite longest one. (d) The continuity while preserving the previous orientation to determine the next movement.

The resulting discrete curve is represented as discontinuous points. Some of these points are missed because finding the mid-point of the perpendicular line segments that

intersects the largest line segment produces consecutive discrete points as shown in 4.9.a, for which at the end needs to be connected so as to maintain the continuity of the drawn spines.

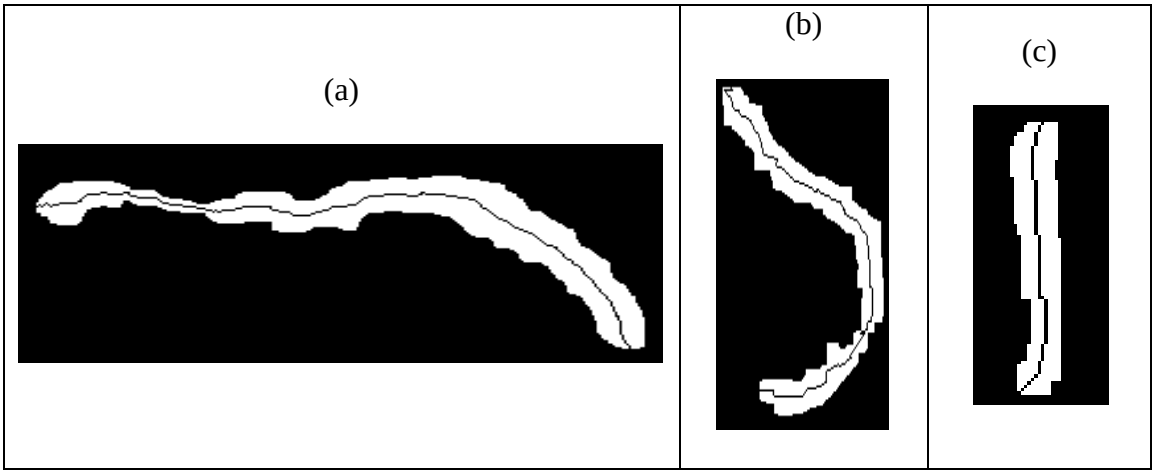


Figure 4.8. The complete filament spine: (a) Full spine of a horizontally aligned filament. (b) Full spine of a filament in an Hα image observed at BBSO observatory on February 9, 2002. (c) Full spine of a vertically aligned filament.

An algebra slope-intercept algorithm (Hearn *et al*., 1997) is used to connect these separated points by drawing lines between them as shown in Figure 4.9.b**.** In this algorithm any straight line on the co-ordinate plane can be described by the equation $ax+by+1=0$.

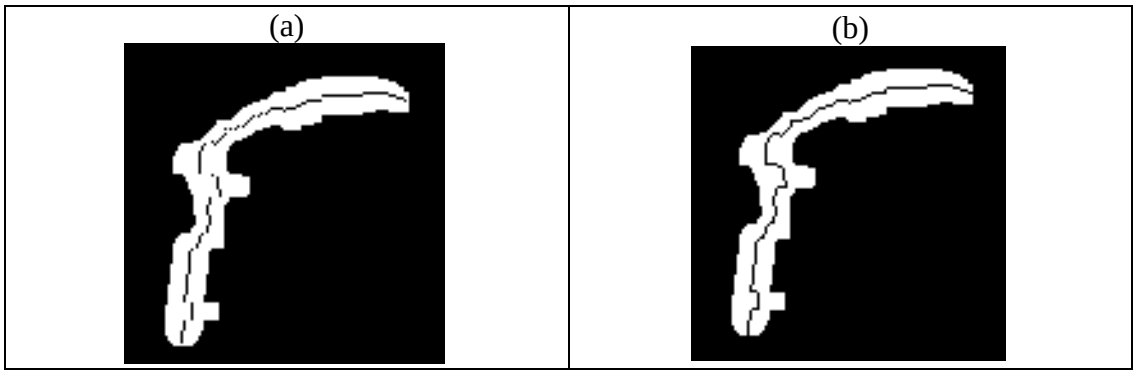


Figure 4.9. The result of applying the slope-intercept algorithm. (a) Before applying the algorithm. (b) After applying the algorithm.

4.4 Critical Evaluation of the Implemented Algorithm

In order to demonstrate the performance of the Atoum spine drawing algorithm, results were compared with those from Bernasconi's algorithm (Bernasconi *et al*., 2005). The two algorithms were applied to ninety eight filaments of various sizes collected empirically from solar images obtained on the dates: 1/1/1999, 2/1/1999, 3/1/1999, 2/1/2001, 3/1/2001, 4/1/2001, 29/7/2001, 30/7/2001 and 31/7/2001. It was found that the Bernasconi filament spines are more adapted for large filaments. This is demonstrated by setting a threshold for the length of the line segment to be twenty pixels in the Bernasconi algorithm. It is also apparent that the Atoum filament spines are more convoluted precisely because they accurately follows the bodies of the filaments, as they goes through the middle of the features as shown in Figure 4.10.

Figure 4.10.a shows an original filament and Figure 4.10.b shows the spine obtained by running the Atoum code drawn over the original filament. The figures 4.10.d, 4.10.f and 4.10.h show the spines drawn using the Bernasconi algorithm with different starting lines. The first guessed line is determined by Bernasconi as explained in (Bernasconi *et al*., 2005) by running roughly parallel to the longest side of the box that just encloses the filament under consideration. Figure 4.10.d shows the output if the first guessed line is chosen as in Figure 4.10.c where $\overline{AB}$ is the longest side of the box (one of the longest sides of the box that encloses the filament), while Figure 4.10.f shows the drawn spine if the first guessed line is chosen as in Figure 4.10.e where $\overline{CD}$ is the longest side of the box that encloses the filament. If the first guess line is chosen as in Figure 4.10.g, the line produced by going from c to d row-by-row going up until we get one end-point and going from d to b column-by-column going back until we get the other end-point , then the output will be as shown in Figure 4.10.h. These results illustrate the better accuracy

of the Atoum algorithm which is a characteristic of crucial importance for the automatic determination of the filament spine.

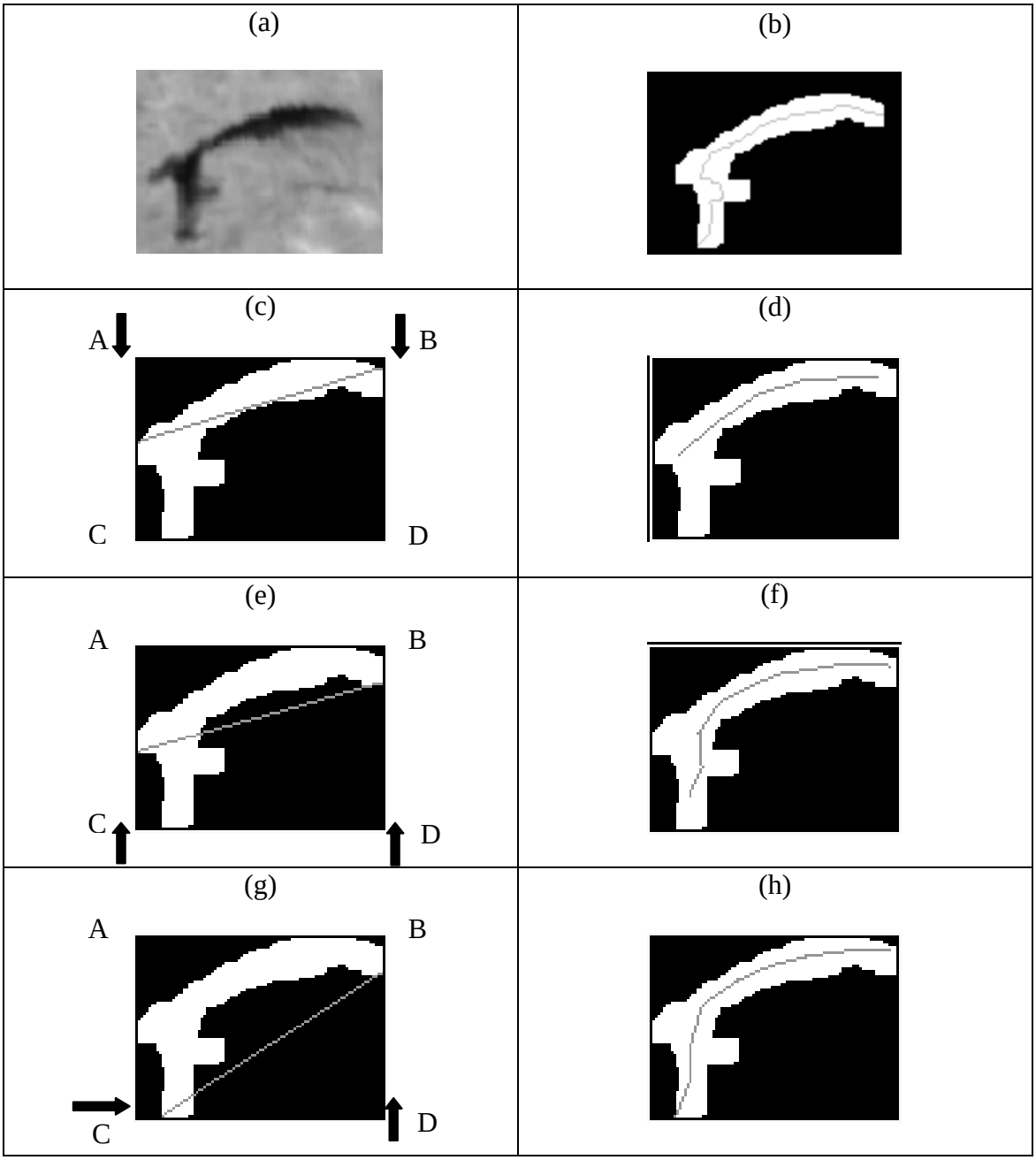


Figure 4.10. Samples illustrating the filament tortuousness and accuracy of the spines from Atoum (b) and Bernasconi (d, f, h) algorithms. (a) Original filament. (b) Spine produced using Atoum algorithm. (c) The first guess Line $\overline{AB}$. (d) The spine produced using the (c) guessed line. (e) The first guess line $\overline{CD}$. (f) The spine produced using the (e) guessed line. (g) The first guest line produced by row and column processing (h) The spine produced using the (g) guessed line.

4.4.1 Efficiency of the algorithm

The Atoum algorithm introduced here is computationally less complex compared to the Bernasconi algorithm. The former on average takes 0.098 s to extract a filament spine, while the latter takes 0.341 s. The computer used in this work was a PC with an Intel Core Duo CPU P8700, operating at 2.53 GHz, running the Windows Vista 32-bit operating system. Figure 4.11 shows the processing times for the full set composed of ninety eight filaments.

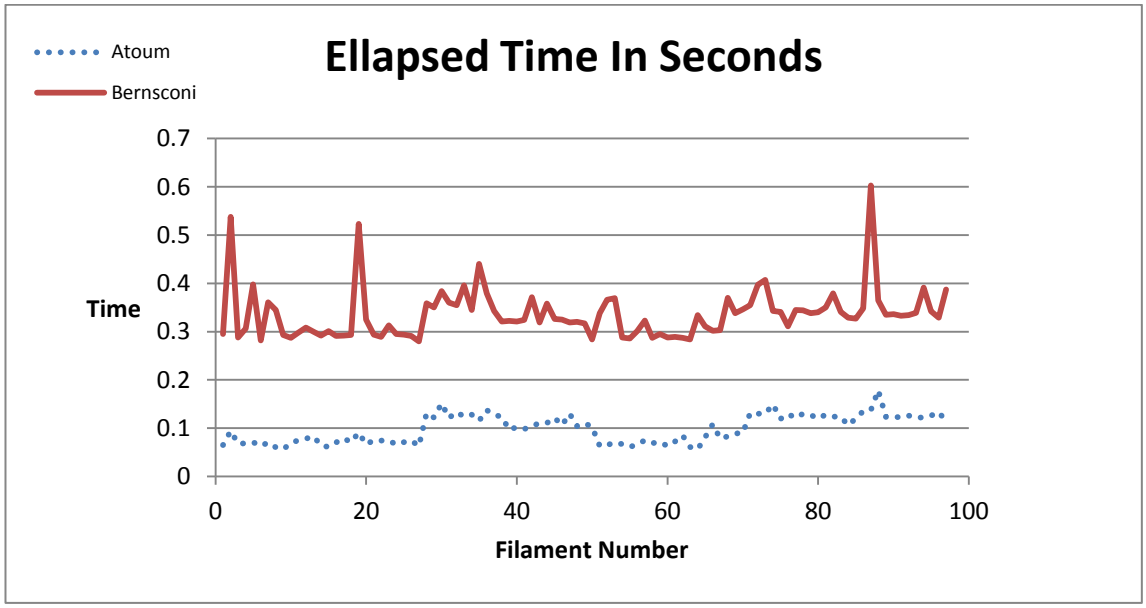


Figure 4.11. Elapsed time in seconds per feature for the two algorithms.

Additionally, the Atoum implementation results in a longer spine because it tracks the actual filament backbone more accurately as shown in Figure 4.10.b. This result is confirmed by the longer pixel count for the Atoum spines as shown in Figure 4.12, compared against (Bernasconi *et* al., 2005) results.

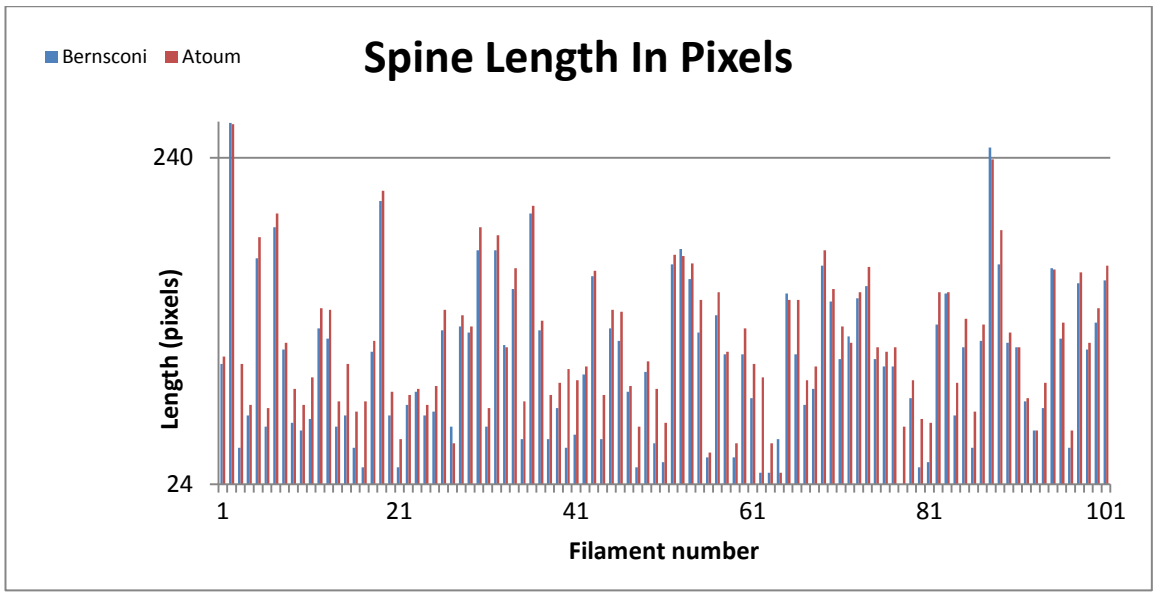


Figure 4.12. Spine length in pixels per feature.

4.5 Conclusions

In Section 4.1 is shown the development of an adaptive thresholding technique for segmenting Hα solar images to get foreground segmented filaments and a non-RoI background. Based on false acceptance rate and output images with well-defined filaments the segmentation process provides quality pre-processed images for machine vision techniques.

Section 4.2 describes the fully automatic detection algorithm for solar filaments which involves calculating statistical parameters and morphological operations. This detection process avoids using empirical values. The shape of the detected features is represented by determining its spine geometry. The resulting Atoum spin drawing algorithm gives the means to extract all the filament morphological features required, such as: filament length, filament centre, filament head-end, filament tail-end and filament boundary. The algorithm is valuable as part of a real-time system for detecting and tracking solar filaments. The comparison with Bernasconi shows that the Atoum algorithm represents the filaments more accurately and is also computationally faster, which could lead to a more precise tracking practice in real-time.

There are still, however, some problems yet to be resolved, these being incomplete spines due to errors in choosing the largest line segment through the automated drawing process.

CHAPTER FIVE

5 ADAPTIVE ALGORITHM FOR MERGING BROKEN FILAMENTS USING NEURAL NETWORK

5.1 Introduction

After the solar detection phase, some filaments were found to be not detected at all or not segmented properly due to variations in intensity along the filament; this results in these filaments being displayed as discrete segments. These broken filaments will affect post processing practices, such as detecting a complete spine for the whole filament. This spine, which is used to define the attributes of the underlying filaments, will, if broken, have an effect on detecting filament disappearances. Thus, a merging algorithm is needed to combine broken sections. In this chapter, a merging algorithm is proposed that merges broken filaments by using a NN classifier. This classifier is trained using numerical and statistical values extracted from the image filaments after drawing their spines. The algorithm considers every two neighbouring filaments, checking the merging possibility. Any merged entity is reconsidered as a new single filament and rechecked for possible further merges with other components. Thus, the algorithm will achieve single or successive merging processes according to the results of the merging criteria.

This chapter is organized as follows: Section 5.2 introduces the idea behind using the NN in the merging process. Section 5.3 describes the merging process. Section 5.4 presents an evaluation of the whole merging process. Conclusions on the associated findings and comparisons with previous work are discussed in Section 5.5.

5.2 Neural Networks

The NN approach is used in this research for data classification purposes. The strength of neural networks comes from their ability to learn a classification function, given a suitable training set, choice of feature vector and NN topology (Qahwaji and Colak, 2006). In addition to their ability to overcome the drawbacks of the classical algorithms used in problem solving, NNs have the attribute of adaptive learning which allows dynamic changes to their structures during the training phase. Moreover, a crucial advantage of these networks is their ability to classify new patterns after training because of their generalization capability (Mao and Jain, 1995 and Lerner *et al*., 1999).

The NN approach has been used before in the field of solar imaging. An NN technique is used in (Zharkova and Schetinin, 2003) to extract solar filaments automatically from Hα solar images. In another example, the results of an intensity filtering technique in (Qahwaji and Colak, 2005) are fed to a NN to verify the detected regions.

In this study, the approach adopted to merge broken filaments uses values found from the extracted filaments as inputs to a NN. The NN has an input layer with number of neurons equal to the number of values. This input layer is connected to a hidden layer and this hidden layer is connected with a one neuron output layer. The NN is trained on five numerical and statistical values extracted from the filament spines under consideration. For each direction, as determined in Section 5.3.1, five different normalized values, as shown in Table 5.1, are computed and fed to the NN. The five input values are specified in the second column, the method used in the normalization process is defined in column three and finally the value of the output is specified in the fourth column. Because the neural network is used to classify the filaments into two cases, merged or not merged, the output layer has one neuron.

Table 5.1. The inputs, normalizing methods and output used by the NN.

No.	Inputs	Normalizing Method	Output
1	The angle subtended by the centroids of two consecutive filaments at the mid-point of the line connecting the filament end points.	Dividing by 180º.	Merged =0.9 Not Merged =0.1
2	The distance between each spine end-point of each filament and the other filaments spine end-points.	Dividing by the maximum distance computed in the image.	
3	The difference between the mean intensities of the two consecutive filaments.	Dividing by the maximum mean value of the filaments in the image.	
4	The difference between the mean intensity of the first filament and the mean intensity of the area in-between.	Dividing by the maximum in-between mean intensity.	
5	The difference between the mean intensity of the second filament and the mean intensity of the area in-between.	Dividing by the maximum in-between mean intensity.	

The first criterion in Table 5.1 is that the two filaments to be merged should have appropriate relative orientation to each other. The second criterion is that two filaments are required to be in close proximity. The third criterion is that the two filaments should be merged if the difference between the mean filament intensity of the two consecutive filaments is small. The fourth and fifth criteria are that the two filaments should be merged if their average intensity values are very close to the in-between average intensity value.

The merging process starts when the NN returns an output value that is greater than 0.5. Then filled circles will be drawn along the line segments that connect these broken

filaments which ultimately recovers the approximate shape of the original filament as shown in Figure 5.1.

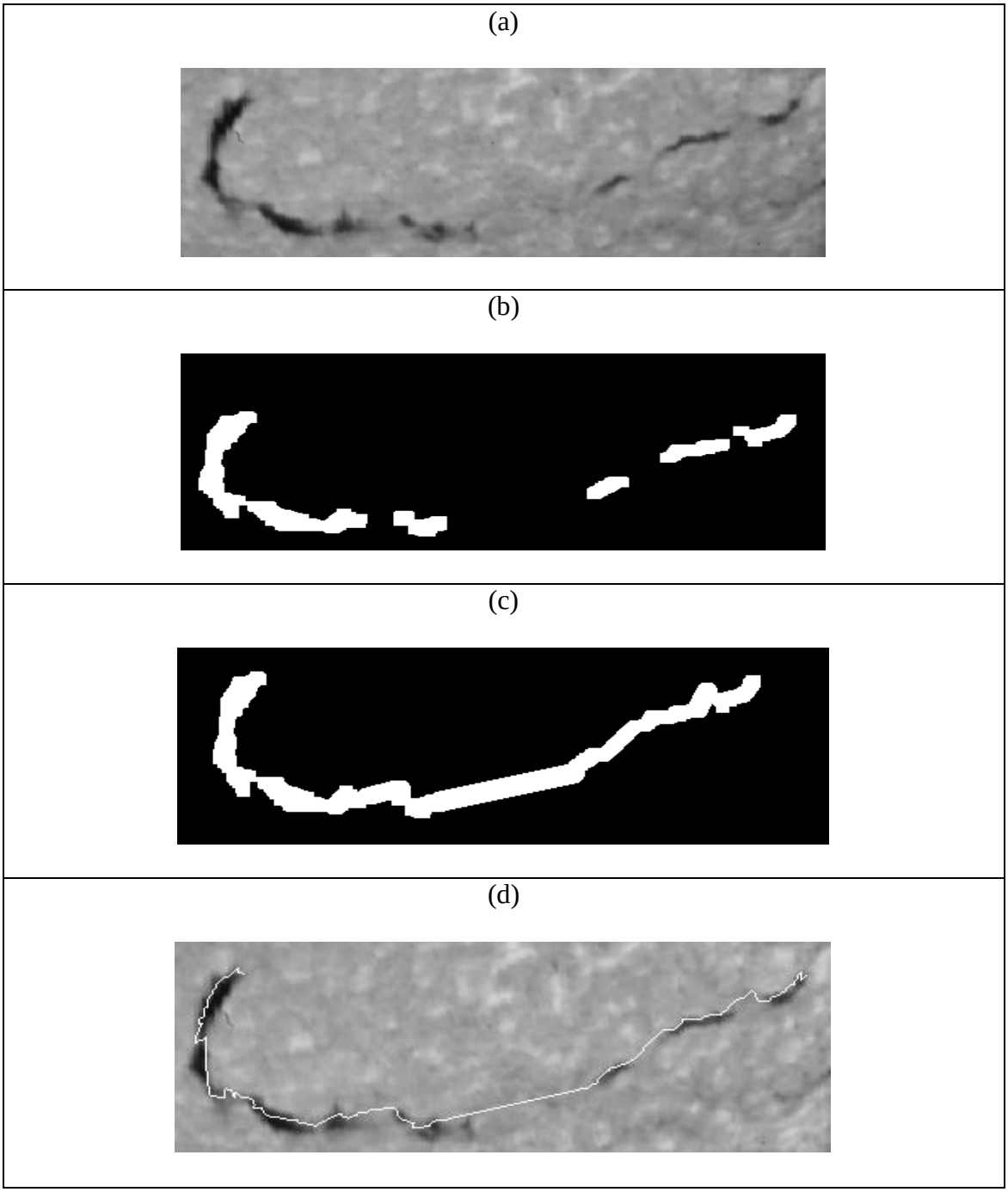


Figure 5.1. Results of the merging process: (a) Original filament. (b) Segmented filament. (c) Merged filament. (d) The filament spine after merging.

The NN was optimized by finding the minimum Mean Squared Error (MSE) during training with different NN topologies as described by Colak (Colak and Qahwaji, 2009).

The training set consisted of 145 samples, extracted from 20 different segmented Hα images from the months, January 1999, January 2001, and February 2001 and the testing set was 31 samples extracted from the same images.

Training experiments were carried out changing the number of nodes in the hidden layer from one to twenty. For every new experiment the MSE of the training stage was recorded and the number of hidden nodes with the least MSE was chosen. The network was optimized using seven nodes for the hidden layer; as shown in Figure 5.2, which shows the structure of the optimised NN.

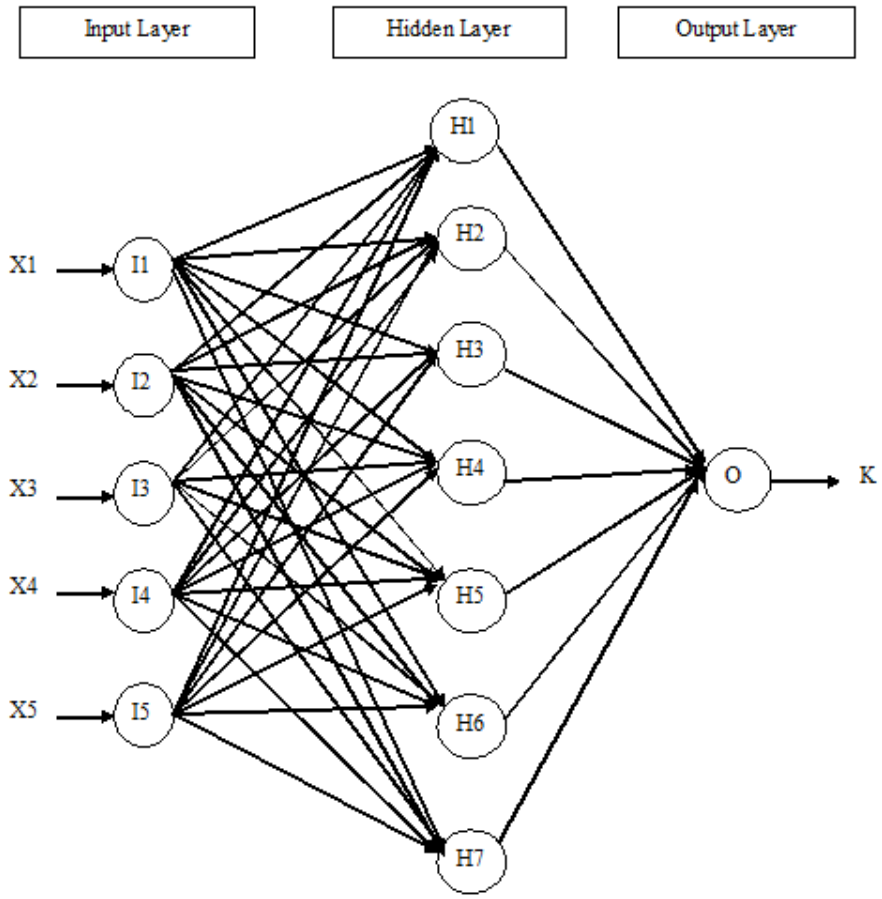


Figure 5.2 . The structure of the optimized neural network.

5.3 Merging Process

5.3.1 Determining the Direction

Determining the most appropriate merging direction from the end of each filament depends on finding the shortest distance between pairs of filaments. To do this, four different distances from each filament to all other filaments are determined. The first distance is between the spine start points of two filaments (SSDistance) as shown in the Figure 5.3.

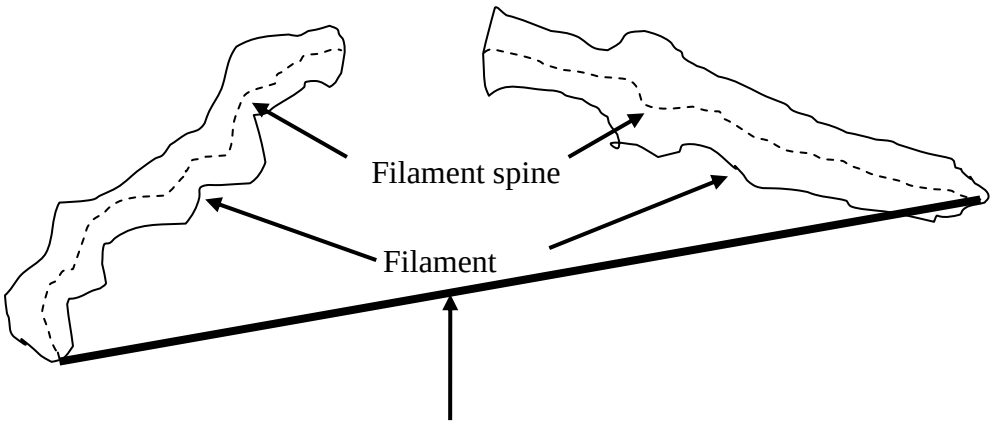


Figure 5.3. The distance between the start points of two successive filaments (SSDistance).

The second distance is between the spine start point of one filament and the spine end point of the other filament (SEDistance) as shown in the Figure 5.4.

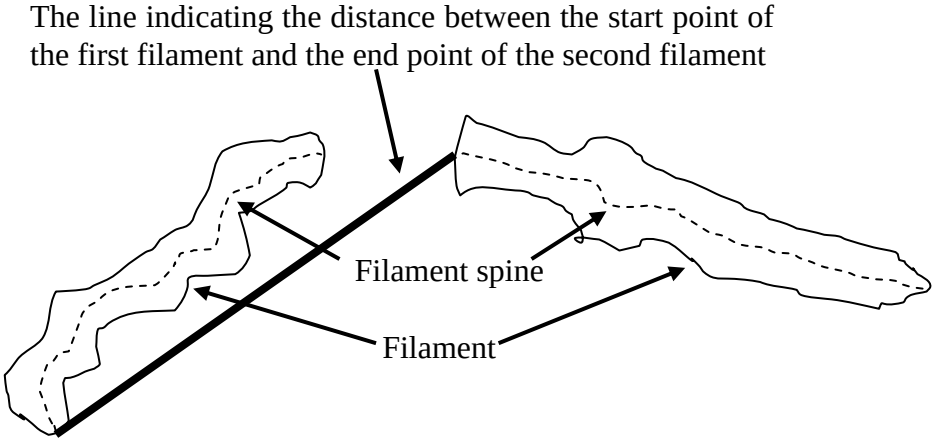


Figure 5.4. The distance between the start point of the first filament and the end point of the second filament (SEDistance).

The third distance is between the end point of one filament and the start point of the other filament (ESDistance) as shown in the Figure 5.5.

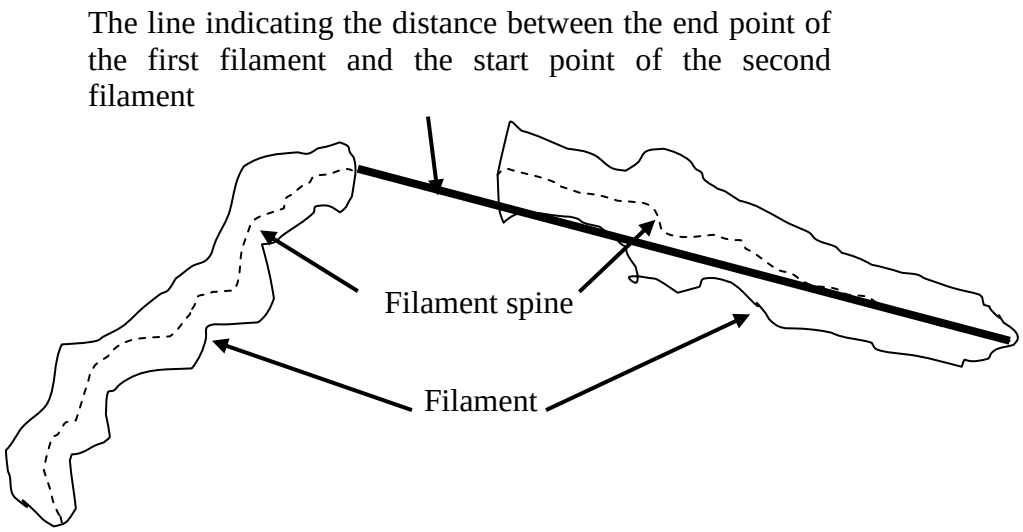


Figure 5.5. The distance between the end point of the first filament and the start point of the second filament (ESDistance).

The fourth distance is between the end points of the two filaments (EEDistance) as shown in the Figure 5.6.

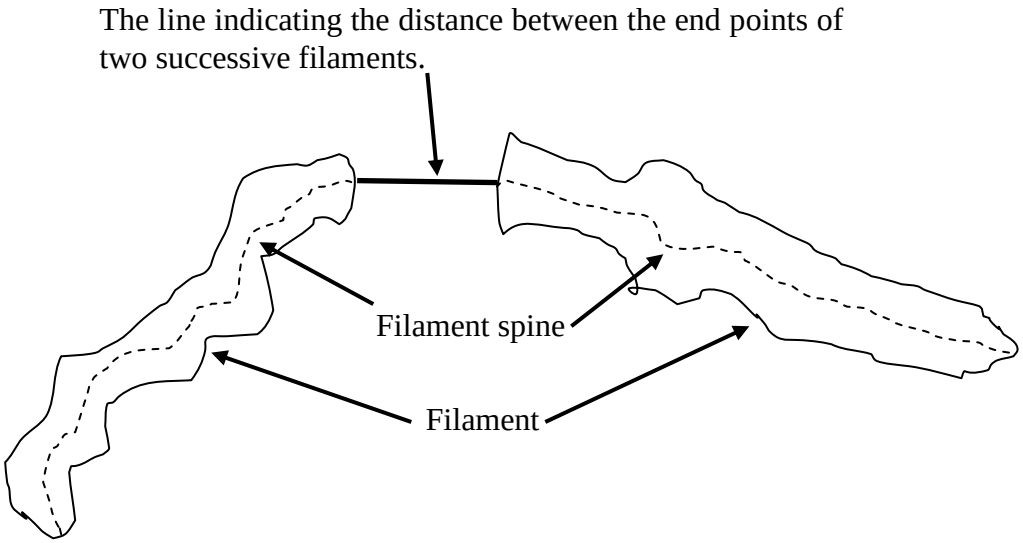


Figure 5.6. The distance between the end points of two successive filaments (EEDistance).

5.3.2 Previous Study

There are two basic criteria specified in the literature to merge broken elongated objects as given by Chen (Cheng *et al.*, 2007) and Ingrid (Ingrid *et al.*, 2002). The

broken features should be close enough and should satisfy some relative orientation requirements. These merging rules were used Bernasconi (Bernasconi *et al*., 2005) who defined constant values to be used with the two criteria. Two constants are used with the distances between each spine end-point of each filament and all other filaments as shown in Figure 5.7.

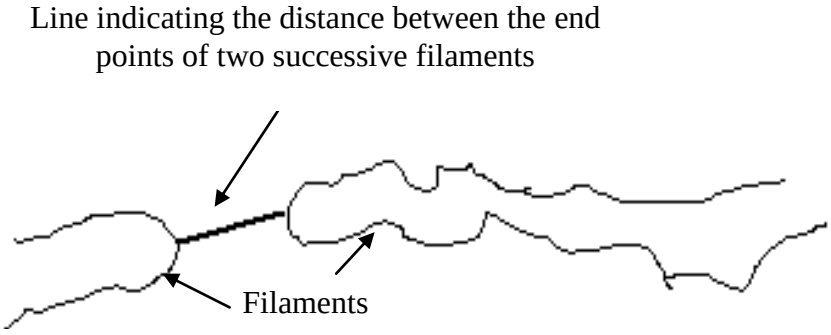


Figure 5.7. The distance between the two filaments.

Another constant value is used with the angles between the spines of the last two filament segments as shown in Figure 5.8 relative to the line connecting the two end points.

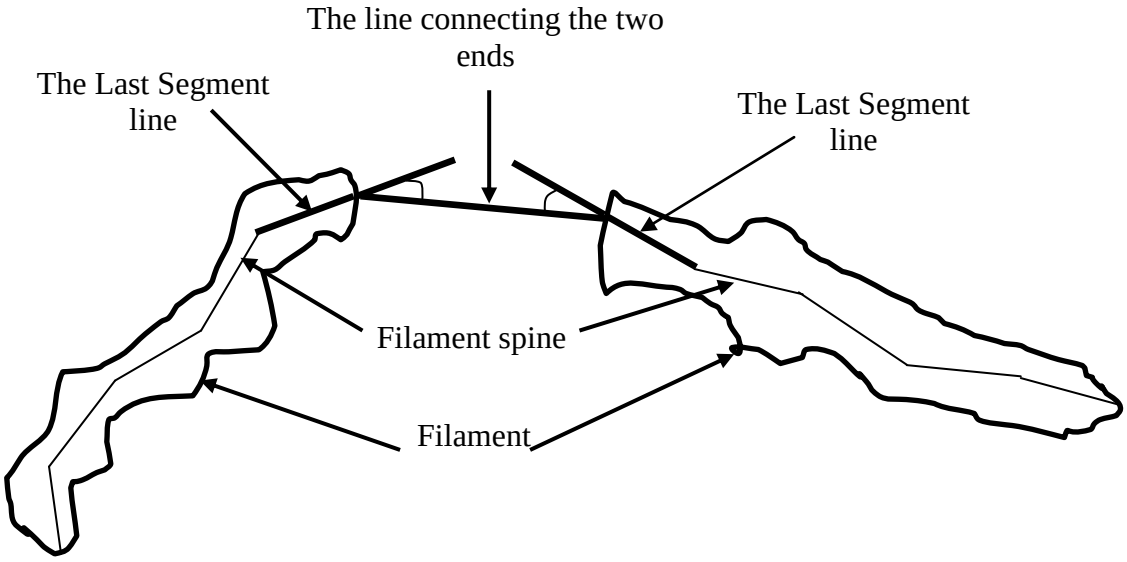


Figure 5.8.The angles between the last two segments and the line connecting the two end-points of the two filaments.

If the distance is less than 25 pixels then the two filaments are merged. If it is greater than 25 pixels and less than 100 pixels then the angles are calculated; if for both

segments this angle is less than 22.5° then the filaments are merged. This merging technique was not considered perfect by the author. It must be noted that this merging process is implemented after characterizing the solar filament. In this chapter the same two criteria (without the constants), with the addition of another three criteria, are used as an input vector for the NN.

5.3.3 The Atoum Approach

5.3.3.1 NN Input Vector

Before determining the input vector and depending on the determined co-ordinates of the start and end-points of each filament and all other filaments in the solar image the distance between each filament centroid and all other filaments centroid are calculated (DCC) as shown in Figure 5.9.

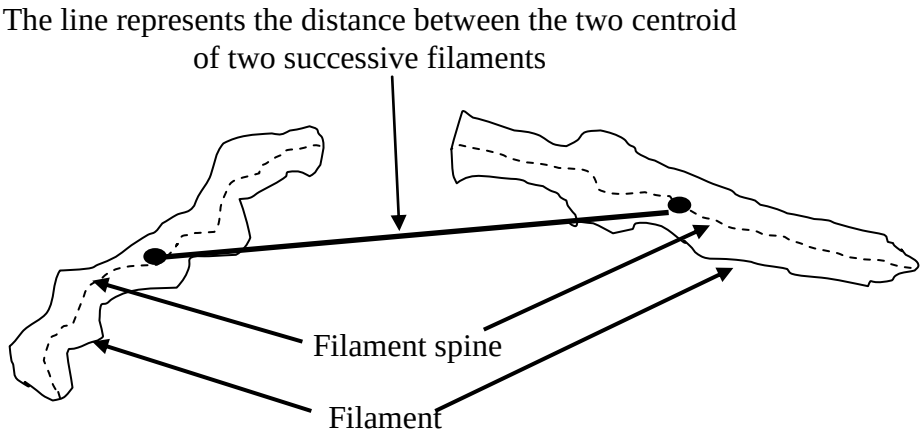


Figure 5.9. Line segment showing the distance between the centroids of two successive filaments.

Thereafter the code computes the distances of each filament spine end-point to all other filaments spine end-points to find the smallest (DEE). If the DEE is less than half the value of the DCC then the direction between each filament and all other filaments in the same image were determined as explained in Section 5.3.1. Then for each direction five different normalized values as shown in Table 5.1 and Figure 5.2, I1 to I5, are computed and fed to the NN. Figure 5.10 illustrate the first criterion in Table 5.1;

which is the angle θ subtended by the centroids of the consecutive filaments at the mid-point of the line that connecting the two filament end-points.

The angle between the two lines that connect the centroid of each filament and the mid-point of the line that connect the two end points.

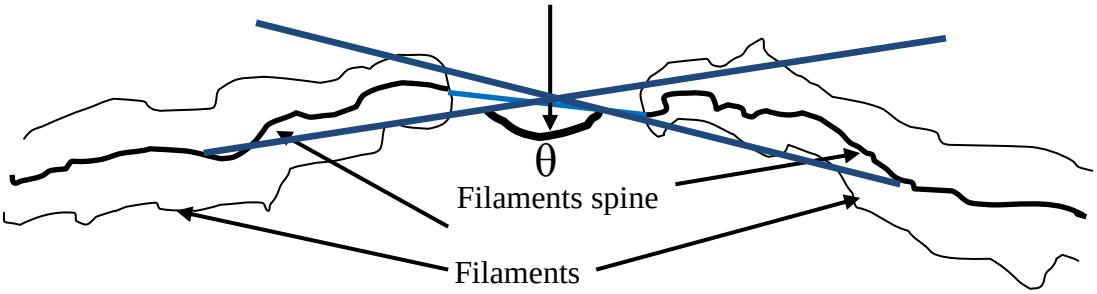


Figure 5.10. The angle between the centroids of two consecutive filaments and the mid-point of the line that connect the two end-points.

5.3.3.2 Filling Gaps by Drawing Circles

Drawing successive circles between the broken filaments can recover approximately the non-broken filaments. These circles takes as diameter the last line segment used in the projection and averaging phase of the process of drawing the filament spine. This is illustrated in figures 5.11 and 5.12.

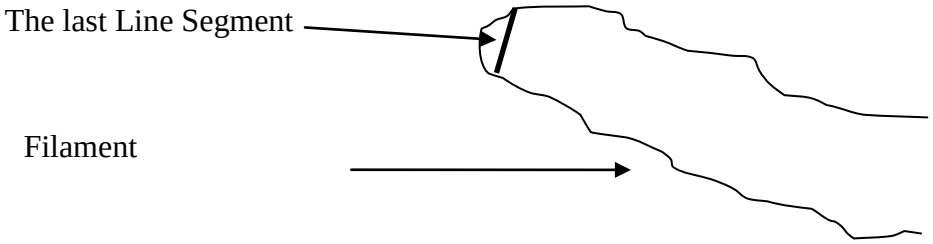


Figure 5.11. The last line segments used to draw the spine of each filament.

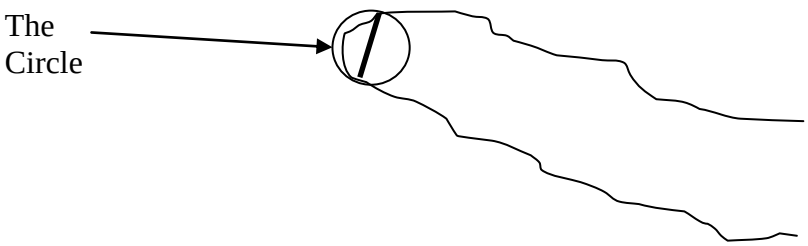


Figure 5.12. The circle drawn by using the last line segments as diameter.

These circles take the points on the line segment that connects the two filaments as centres, as shown in Figure 5.13.

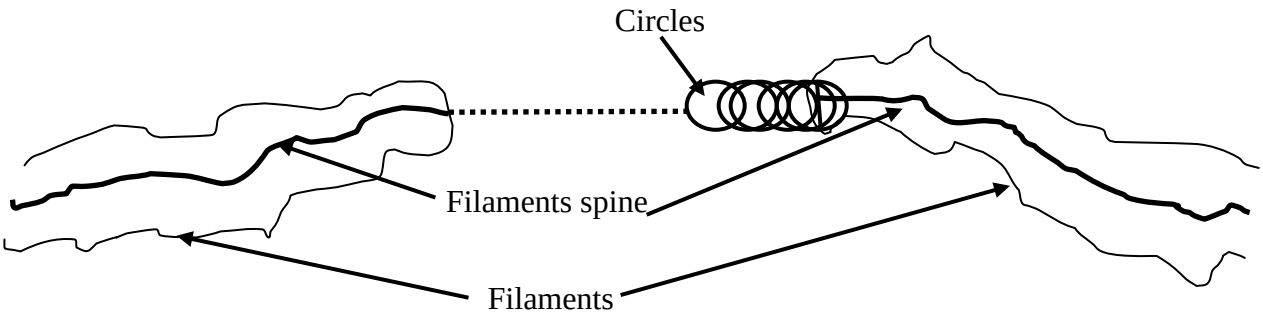


Figure 5 13. Drawing circles by taking the last line segment as diameter and taking the points on the line that connects the two filaments as centres of these circles.

Using these points will adjust the position of the drawn circles according to the orientation of the broken filaments as shown in Figure 5.14 and in Figure 5.1.

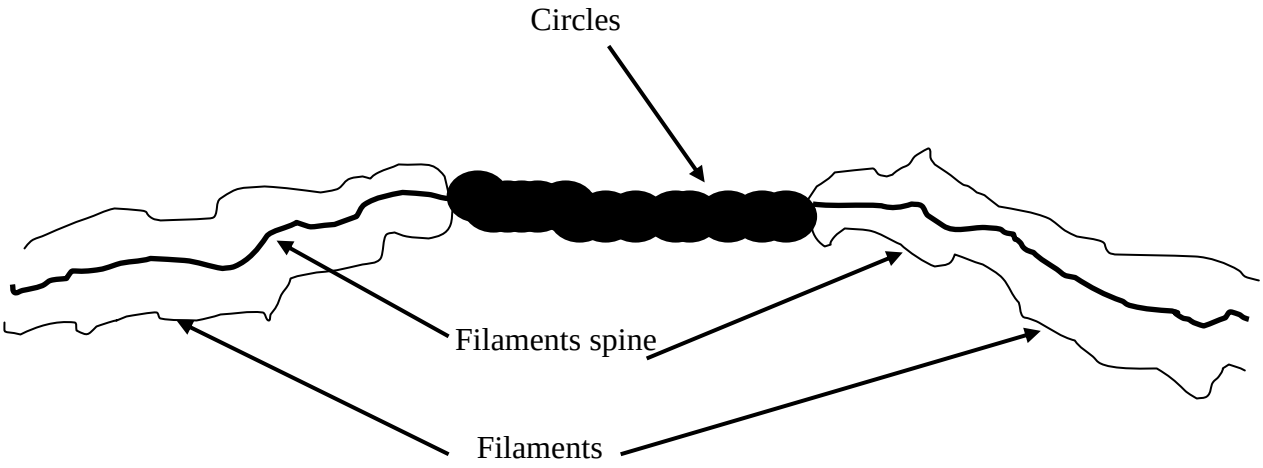


Figure 5.14. Drawing filled circles to fill the gap.

The merging process starts when the NN shown in Figure 5.2, returns an output value K that is greater than 0.5. The merging algorithm draws filled circles along the line segments that connect these broken filaments which ultimately recovers the approximate shape of the original filament as shown in Figure 5.1.c.

5.4 Evaluation of the Implemented Algorithm

The main aim of the work presented in this chapter was the creation of computer software that can achieve the merging of broken filaments. The merging method described in this chapter is novel and has not been previously used by any other

researchers as far as is known by the author. The algorithm is the first that uses an adaptive Neural Network classifier to merge the segmented filaments. It also appears to give a higher filament merging percentage than that of Bernasconi, as indicated below. Table 5.2 shows results of applying the Atoum merging technique on a selected set of Hα images.

Table 5.2. Results of applying the merging process on a set of Hα images.

	Image	Observed Number of Broken Filaments (BF)	Number of Broken Parts (NBP)	Merged Parts (MP)	Not Merged Parts (NMP)	Falsely Merged Parts (FM)
1	mh010304	4	8	8	0	4
2	mh010307	4	8	8	0	2
3	mh010308	2	5	5	0	0
4	mh010309	4	9	7	2	2
5	mh010312	1	5	5	0	0
6	mh010321	6	16	14	2	0
7	mh010401	5	12	10	2	2
8	mh010402	1	2	2	0	0
9	mh010403	3	8	6	2	2
10	mh010404	3	6	6	0	0
11	mh010408	5	12	8	4	0
12	mh010410	5	10	8	2	0
13	mh010412	2	5	2	3	0
14	mh010413	1	3	2	1	0
15	mh010418	1	3	3	0	0
16	mh010419	1	2	2	0	0
17	mh010423	3	6	4	2	0
18	mh010501	5	16	14	2	0
19	mh010502	1	2	2	0	0
20	mh010508	1	2	2	0	0
21	mh010509	3	6	6	0	0
22	mh010510	3	6	6	0	0
23	mh010512	2	4	2	2	0
24	mh010513	2	4	2	2	0
25	mh010514	1	2	2	0	0
26	mh010515	3	6	6	0	2
27	mh010516	2	5	4	1	0
28	mh010517	2	8	7	1	0
29	mh010518	3	7	6	1	0
30	mh010519	4	10	10	0	0
Total		83	198	169	29	14

The name of the image which also shows the observing date is given in column two. Column three shows the number of broken filaments observed in the image (BF), column four shows the number of broken parts in all the broken filaments in the image (NBP), column five shows the number of correctly merged parts (MP), column six shows the number of parts that should be merged but are not (NMP) and column seven shows the number of parts which are merged but should not be (FM). From the totals in Table 5.2 the following performance measures result:

Merging rate: MP/NBP = 85.35%.

Non-merging rate: NMP/NBP=14.65%.

False-merging rate: FM/NBP=7%.

The merging rate was 85.3%, and is achieved without the use of thresholds unlike the traditional techniques. According to the author's knowledge, this is the first time such detailed results have been presented, which thus makes it not possible to compare these results with previous researchers. Bernasconi (Bernasconi *et al*., 2005), mentions a merging percentage of 70% but no supporting information, like Table 5.2, is provided. It should be noted that the present result of 85% is superior to that reported by Bernasconi.

5.5 Conclusions

A novel merging method for combining the broken filaments is introduced in this chapter. This method exploits the findings from the extraction of the filament spine, the resulting start-points and the end-points accordingly and other statistical values that are computed for the filaments which are candidates for merging process.

This merging algorithm achieves the ability to merge the broken filaments by using a NN classifier trained with feature values extracted from a set of sample image filaments. After extracting filament spines, every two consecutive filaments, is considered to

check the merging possibility. Any merged entity is reconsidered as a new filament and rechecked for possible further merges with other filament. Thus, the algorithm will achieve single or successive merges according to the results of the merging criteria.

The algorithm achieves a higher merging percentage than the method of Bernasconi, but there is still a percentage of false merges remaining. It should be noted that this work highlights this merging challenge of false merging - often overlooked by other researchers. The merging algorithm exceeds the work of others by using the adaptive nature of using an NN and by avoiding the use of empirical threshold values for different filament merging attributes.

CHAPTER SIX

6 AUTOMATIC TECHNIQUE FOR TRACKING SOLAR FILAMENTS

6.1 Introduction

Object tracking is a general imaging term which represents an important task in the field of computer vision. Tracking can be defined as the problem of estimating the path of a moving or moved object in the image plane, as needed for further analysis (Yilmaz *et al.*, 2006). The lack of automatically generated tracking catalogues, the general need for automatic object analysis and representation and the need to know the reasons behind the occurrences of CMEs, together with the increasing power of computers and the availability of high quality instruments, make opportune the development automatic tracking methods for solar features and activities that may affect life on Earth.

Filament disappearances are generally monitored by observing and analysing successive solar H-alpha images. In this chapter, after the filament regions are obtained from individual H-alpha images, the results of two consecutive images are compared to detect these filament disappearances. Significantly, a novel efficient tracking technique for solar filaments is presented in this chapter, taking advantage of Heliographic Carrington Maps (HCMs) obtained by converting solar images to Heliographic Carrington co-ordinates. The relatively fixed positions of solar filaments over consecutive HCMs facilitate the process of tracking solar filaments. This study is directed towards large, stable filaments; small and unstable filaments are disregarded. A NN classifier is used to categorize the detected filaments as DFs or mis-detected filaments (MDFs). Features such as: Area, Length; intensity statistics Mean,

Standard Deviation, Skewness, and Kurtosis, are extracted and fed to this network to get a confidence level of at least 80 %.

The Atoum technique is composed of three main stages: detecting and representing solar filaments, generating HCMs and implementing an automatic tracking method.

The technique presented here gives results in close uniformity with those obtained manually, presented in the NGDC catalogue. The author could not find an evaluation or comparisons by other researchers like the one undertaken here. Furthermore, the only automatic algorithm that could found was provided by SWRL. However, it does not seem to perform very well when compared with the NGDC catalogue as indicated in Section 6.4.

So this study is most likely the first in distinguishing the actual disappeared filaments and the miss-detected filaments by applying ANNs.

The research presented here has concentrated on devising a novel solar detection and tracking technique based on an extensive literature review. The resulting algorithm does this by the use of an NN classifier with no empirical variables, which is one novelty of the algorithm. In the previous literature, all such methods appear to have used empirical values in their algorithms.

This chapter is organized as follows: Section Two describes the Carrington Heliographic co-ordinate system. Section Three illustrates the tracking technique. Results and evaluation are presented in Section Four and finally the conclusions are presented in Section Five.

6.2 Carrington Heliographic Co-ordinate System

Heliographic co-ordinates are a spherical co-ordinate system that uses the two variables, latitude and longitude to locate a point on the solar surface (Colak *et al.*, 2011). The longitude co-ordinates are converted to the Heliographic Carrington Co-ordinates to identify the positions of solar features in a manner independent of the solar rotation. Figure 6.1.a shows a H-alpha image with its segmented form in Figure 6.1.b and the segmented filaments are

shown in the HCM in Figure 6.1.c in the position corresponding to the date of the image in (a) .

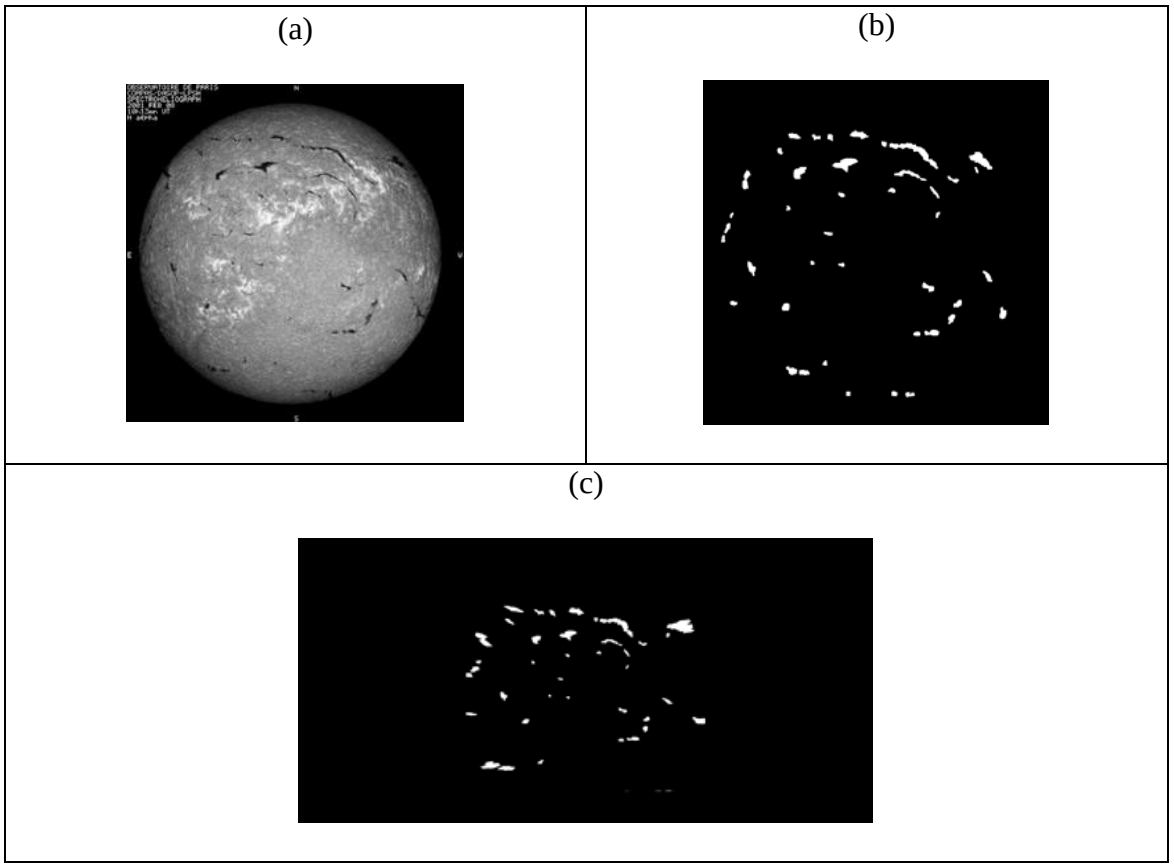


Figure 6.1. The HCM created using an H-alpha Image observed at the Meudon Observatory on 8th February, 2001. (a) Original image. (b) Segmented image. (c) Heliographic Carrington Map.

The method used in this study for filament tracking purposes to create HCMs is the one described by Colak (Colak *et al.*, 2011) as part of work on 3D representation of solar features. It starts by creating images for storing data in Carrington Heliographic Co-ordinates using a standard method of spherical astronomy described by Smart (Smart and Green, 1977). During this phase a resolution recalculation is performed to avoid information loss resulting from truncation. An image enhancement algorithm is then applied to determine pixels missing in the newly created heliographic images.

These maps are used, to take advantage of the relatively small movements of the filaments over such HCMs. This movement is easily accommodated using a rectangular tracking

window. This is used to determine whether a filament has disappeared or not by placing it in an identical position in the next day image. The presence of any white pixels means that the filament has not disappeared. The use of this technique means that it does not matter if the filament moves slightly in any direction as long as a single pixel remains in the search window because all that is of interest is the binary decision as to the presence or non-presence of the filament.

6.3 Filament Tracking Technique

Filament disappearance may be associated with filament eruptions, in which case at least 50 percent of the ejected material disappears over a 24 hour period (Jing *et al.*, 2004). In order to detect the filament disappearances, a new technique is developed to compare the filament detection results obtained in every two consecutive (daily) images.

The technique starts by checking that the images under consideration are successive images like those shown in Figure 6.2. Figures 6.2.a and 6.2.b show two successive solar images, the corresponding segmented images are shown in figures 6.2.c and 6.2.d and finally their corresponding HCMs are shown in figures 6.2.e and 6.2.f.

To exclude small erupted filaments, a condition is applied, as described below, that still keeps the algorithm fully automated by not using a specific filament size. For each filament in the first image, a search process for white pixels is carried out in the second image within the rectangle that encloses the first image filament. If no white pixel is detected and if the area of the first day filament is greater than the average filament area, it is considered a disappeared filament (DF), otherwise it is considered a stable filament (SF).

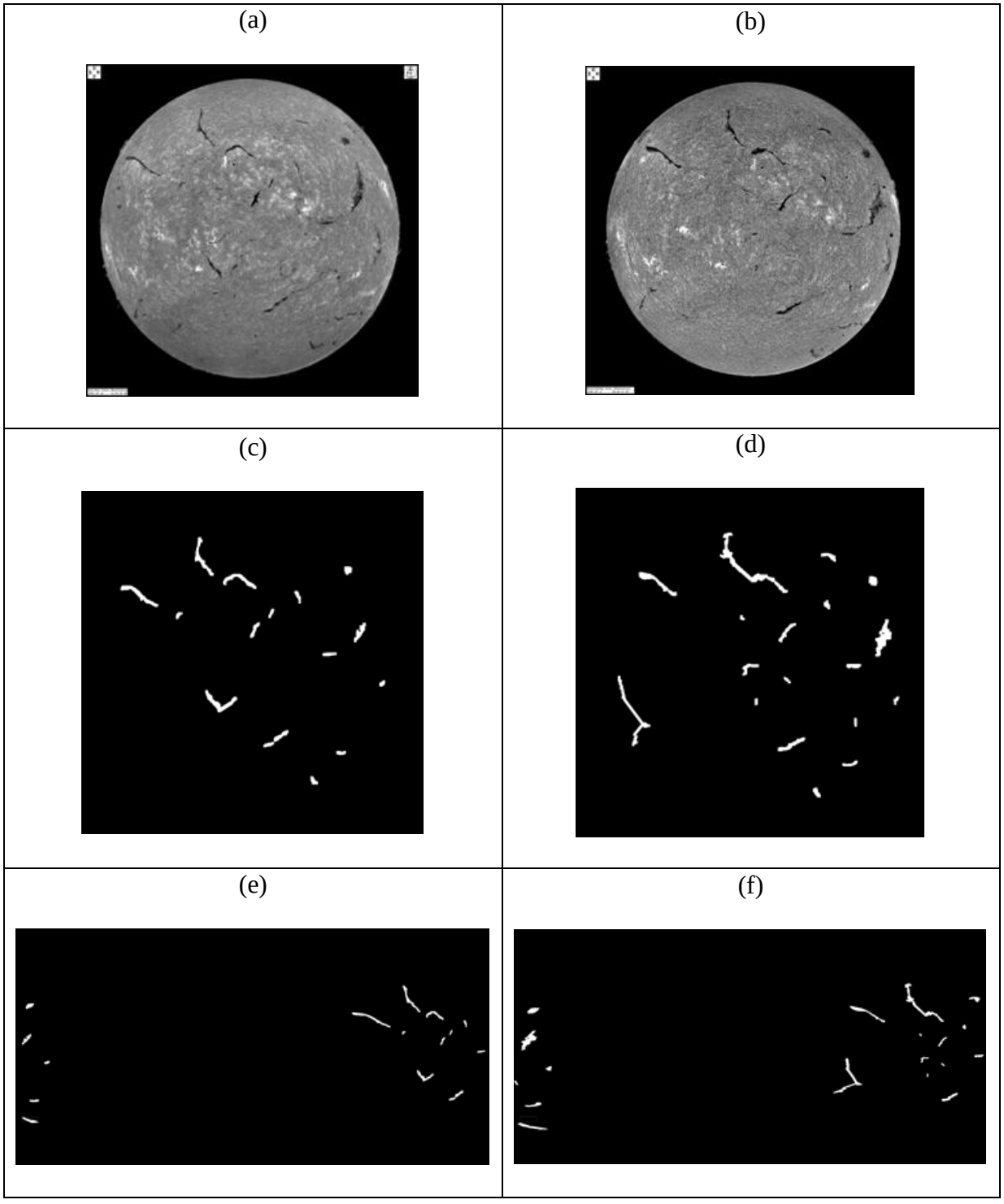


Figure 6.2. Example for two consecutive solar images. (a) H-alpha image observed at the BBSO observatory on 16th February, 2001. (b) H-alpha image observed at the KANZ observatory on 17th February, 2001. (c) Segmented image for (a). (d) Segmented image for (b). (e) HCM for (c). (f) HCM for (d).

A flow chart for the whole algorithm is given in Figure 6.3.

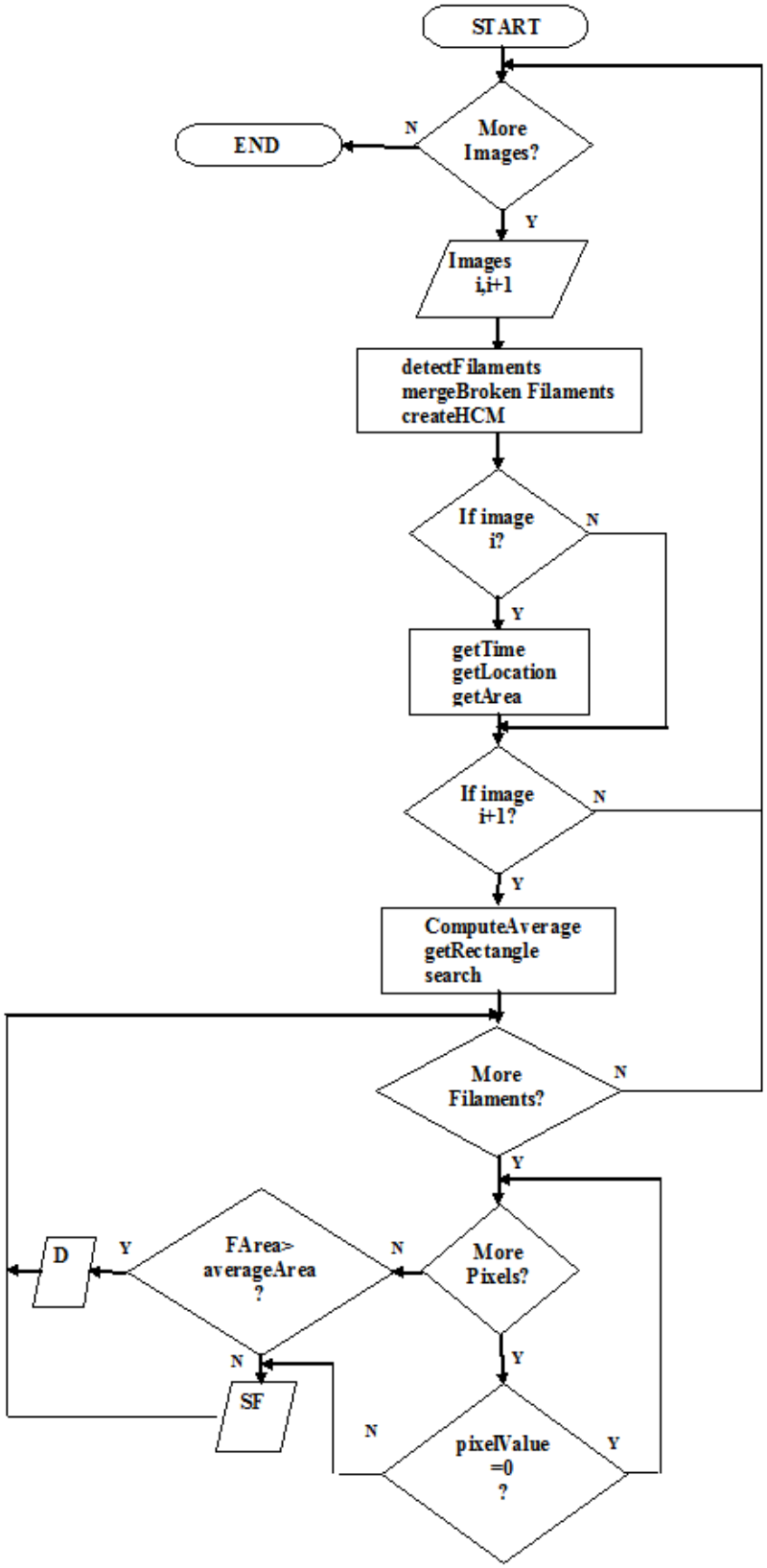


Figure 6.3. Flow chart of the tracking algorithm.

Here filaments are treated as two-dimensional projections of their 3D structure in line with other researchers but it is acknowledged that filaments are 3D dimensional structures and their movements are also in 3D and this extended analysis is reserved for future study.

Although the current method was found to give results closer to the manual NGDC catalogue than other work as shown in Section 4, there are still some filaments which are identified as disappeared whilst in fact they still exist according to manual observation. These are called miss-detected filaments.

The Atoum tracking algorithm was tested using solar maps for the whole of year 2001, which contained 6414 solar filaments. The results after running the code are shown in Table 6.1.

Table 6.1. The number of filaments detected in one year classified as actually disappeared filaments or miss-detected filaments.

Month	Total Number of Filaments	Total Number of Candidate Disappeared Filaments	Total Number of Miss Detected Filaments	Total Number of Disappeared Filaments
Jan	444	12	9	3
Feb	404	8	4	4
March	437	16	11	5
April	474	16	6	10
May	570	13	6	7
June	569	14	9	5
July	581	14	2	12
Aug	514	19	7	12
Sept	628	13	8	5
Oct	490	19	15	4
Nov	690	20	12	8
Dec	613	12	11	1
Total	6414	176	100	76

Table 6.1 contains the filament data divided into twelve rows labeled by month in the first column, the second column shows the total number of filaments detected in this month. The total number of disappeared filament candidates found is shown in column three. Column four shows the total number of miss-detected filaments disappearances and column five shows the total number of disappeared filaments. This total number of DF does not include the small filaments which have areas less than the average filament area. It can be seen in

Table 6.1 that the algorithm found 176 possible disappeared filaments; while by observation, 76 was the actual number of disappeared filaments and 100 were miss-detected filaments.

A Neural Network classifier was applied to reduce the number of these miss-detected filament disappearances. This NN was trained using six statistical values extracted from the filament segments under consideration as input to the NN together with the required decision. During training, the NN structure was optimized as to number of hidden nodes. Because the NN is used to classify the filaments as DFs or MDFs the output layer was chosen to have one neuron. The training set was 128 samples randomly selected from the 176 possible disappeared filaments, which correspond to about 72 % of the whole set. Table 6.2 shows part of this set, where the units of the input features, length, mean, Standard Deviation, skewness and kurtosis are pixel based while the unit of the area is pixel squared.

Table 6.2. Part of the dataset used for training the NN.

Area	Length	Mean	Standard Deviation	Skewness	Kurtosis	Classification
1	0.729412	0.64234	0.531212	0.386208	0.502009	0.1
0.783124	0.529412	0.822351	0.825098	0.542343	0.3975	0.1
0.289439	1	0.35511	0.776182	0.715797	0.623172	0.1
0.173395	0.455621	0.58242	0.866962	1	1	0.1
0.344664	1	0.410778	0.86509	0.381441	0.48789	0.1
0.454819	0.956522	0.746017	0.619736	0.176456	0.413408	0.1
0.458573	0.541667	0.836633	0.89782	0.539348	0.433797	0.1
1	0.978022	0.0902762	0.524678	1	1	0.1
0.454416	0.862745	0.615368	1	0.369439	0.300593	0.1
0.345745	0.921053	1	0.496952	0.0366197	0.254274	0.1
0.436374	0.77027	0.56723	0.730467	0.387564	0.24478	0.1
0.730387	0.433566	0.860784	0.771631	0.454088	0.394722	0.1
0.502905	1	0.499613	0.478322	0.360384	0.207216	0.1
0.32581	0.331984	0.504304	0.683344	0.450343	0.423496	0.1
0.621635	0.668966	0.156056	1	0.64531	0.82462	0.9
0.595061	0.877193	0.755683	0.643527	0.326519	0.239541	0.9
0.515694	0.428571	0.338145	0.757243	1	1	0.9
0.446492	0.512563	0.727908	0.900381	0.0304002	0.0473946	0.9
0.574199	0.409639	0.712091	0.714645	0.398278	0.310655	0.9
1	0.9625	0.371204	0.742038	0.868716	0.828327	0.9
0.435517	1	0.555018	1	0.540473	0.436595	0.9
0.549085	0.452381	0.793325	0.676178	0.306355	0.606939	0.9
0.696041	0.333333	1	0.612786	0.609897	0.504179	0.9
1	0.309091	0.467397	0.510517	0.584846	0.319598	0.9
0.410437	0.618812	0.552634	0.723398	0.266687	0.304684	0.9
0.388782	0.360656	0.581624	0.637538	0.069281	0.100327	0.9

The testing set was 20 out of the full set of 176 samples, as shown in Table 6.3.

Table 6.3. The dataset used for testing the NN.

Area	Length	Mean	Standard Deviation	Skewness	Kurtosis	Classification
0.937741	0.357771	0.792345	0.69617	0.090797	0.237719	0.1
0.375629	0.832335	0.534466	0.474628	0.354462	0.431021	0.1
1	0.315412	0.43658	0.665064	0.142854	0.086056	0.1
0.371573	0.57764	1	0.522092	1	1	0.1
0.503686	0.836207	0.604092	0.689848	0.0177343	0.129573	0.1
0.331754	1	0.477922	0.641134	0.364544	0.270865	0.1
0.487463	0.391608	0.763976	0.434676	0.910812	1	0.1
0.313131	0.62037	0.540518	0.682912	0.0186854	0.0686932	0.1
0.111579	1	0.345268	0.484412	1	0.75767	0.1
0.223774	0.490741	1	0.597617	0.219854	0.224239	0.1
0.631246	0.547038	0.665606	1	0.235759	0.160758	0.9
0.350589	0.4	0.638337	0.681501	0.27138	0.189881	0.9
0.212212	0.745946	0.869229	0.590798	0.0678205	0.195042	0.9
0.34753	0.421941	0.596294	0.630554	-0.0687134	0.0536176	0.9
0.570649	0.302752	0.542807	0.489823	0.362108	0.604928	0.9
0.175296	1	0.131282	0.556901	0.512134	0.486516	0.9
1	1	0.219364	0.721113	1	1	0.9
0.77796	0.4	0.840255	0.568293	-0.691428	0.570762	0.9
1	0.458937	0.743717	0.521591	0.743333	0.914094	0.9
0.22876	0.399679	0.651975	1	0.55278	0.675372	0.9

In the columns in Table 6.2, the last two input features are skewness, and kurtosis, which are calculated as shown in Equations 6-1 and 6-2 respectively (*NIST/SEMATECH, 2012):*

$$skew = \frac{\sum_{n=0}^{n}[X(n)-\mu]^3}{N\sigma^3} \tag{6-1}$$

$$kurtosis = \frac{\sum_{n=0}^{n}[X(n)-\mu]^4}{N\sigma^4} \tag{6-2}$$

Where μ is the mean, σ is the standard deviation, N is the total number of data points, and $X(n)$ is the pixel value. The classification value is shown in the last column where 0.1 means stable filament and 0.9 means disappeared filament.

6.3.1 Optimizing the Neural Network

Several training experiments were carried out while changing the number of nodes in the hidden layer from one to twenty. For every experiment, the MSE of the training stage was

recorded and the number of hidden nodes with the least MSE was chosen as the optimized one. The network was optimized with nine nodes in the hidden layer as shown in Figure 6.4.

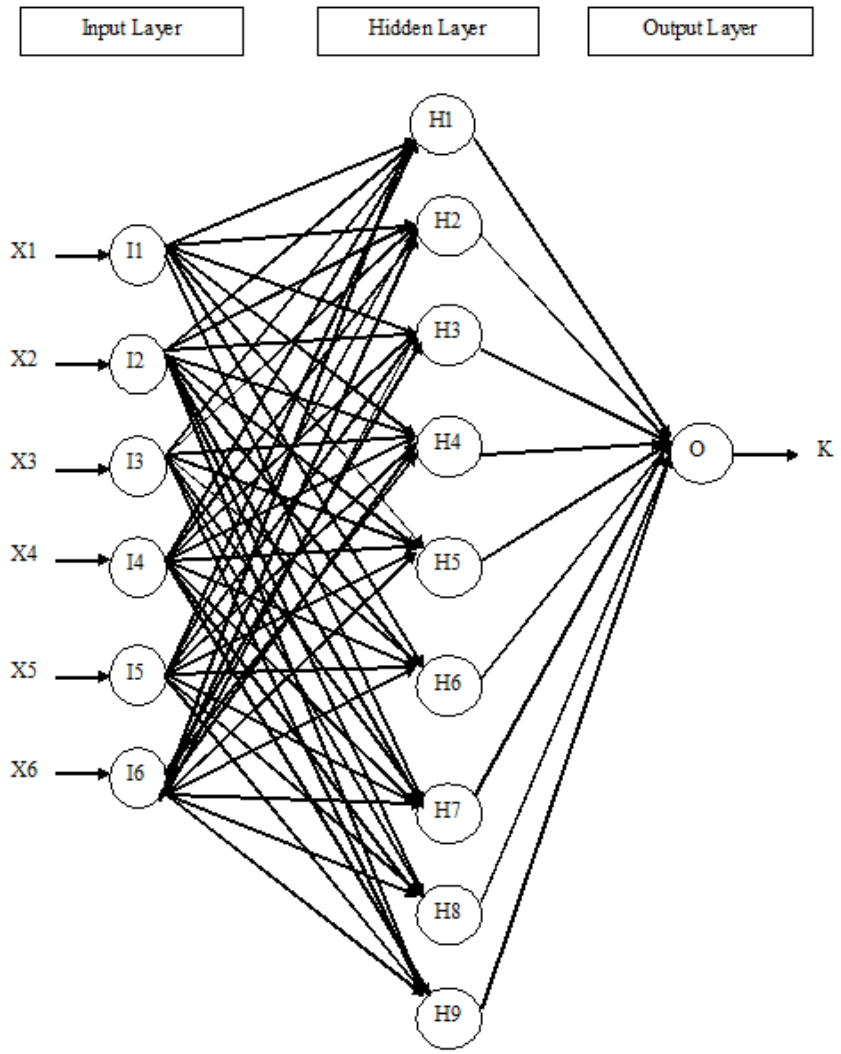
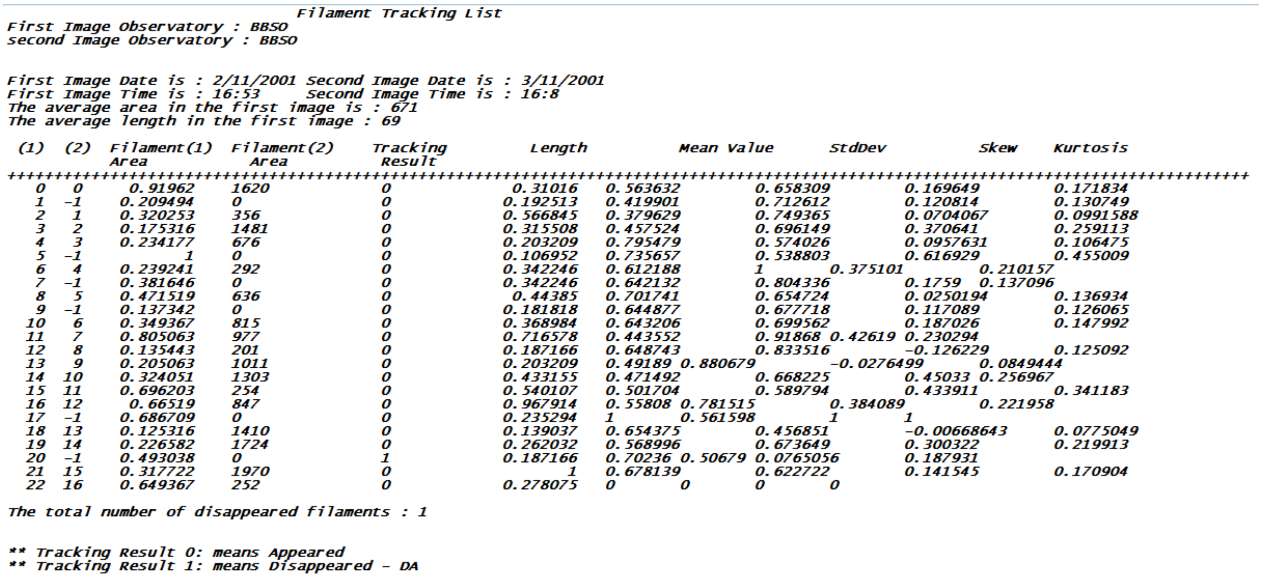


Figure 6.4. The structure of the optimised neural network.

Detecting a filament disappearance starts by checking that the area of the filament under consideration is greater than the average area of all filaments in the first day image. Next, the NN begins working with the values of the six parameters and returns the classification value of the filament under consideration. If this value is greater than 0.5 then the underlying filament is classified as having disappeared; otherwise it is considered as stable filament. The disappearance detection algorithm is considered unique because it is the first time that an adaptive NN classifier has been used to detect the filament disappearance as far as is known

by the author. Not only that, it gives a high detection rate of 80 %, as shown by comparing results in Tables 6.3 and 6.4.

A screen dump showing the output from the algorithm displayed in Figure 6.5, below, shows at the top, the name of the first and second solar image observatory, the date, time, the average filament area and average filament length. Column 1 shows the number of the filament in the first day image, column 2 shows the number of corresponding filament in the second day image; if it is considered as disappeared then this number is shown as -1, column 3 shows the normalized filament area, column 4 shows the area of the corresponding filament in the second day image. Column 5 contains the tracking results; where 1 means disappeared and 0 means stable.



```
                               Filament Tracking List
First Image Observatory : BBSO
Second Image Observatory : BBSO

First Image Date is : 2/11/2001 Second Image Date is : 3/11/2001
First Image Time is : 16:53     Second Image Time is : 16:8
The average area in the first image is : 671
The average length in the first image : 69

 (1)  (2)  Filament(1)  Filament(2)     Tracking        Length          Mean Value      StdDev          Skew    Kurtosis
           Area           Area           Result
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
    0   0     0.91962   1620            0             0.31016   0.563632        0.658309        0.169649        0.171834
    1  -1    0.209494   0               0            0.192513   0.419901        0.712612        0.120814        0.130749
    2   1    0.320253   356             0            0.566845   0.379629        0.749365        0.0704067       0.0991588
    3   2    0.175316   1481            0            0.315508   0.457524        0.696149        0.370641        0.259113
    4   3    0.234177   676             0            0.203209   0.795479        0.574026        0.0957631       0.106475
    5  -1           1   0               0            0.106952   0.735657        0.538803        0.616929        0.455009
    6   4    0.239241   292             0            0.342246   0.612188        1       0.375101        0.210157
    7  -1    0.381646   0               0            0.342246   0.642132        0.804336        0.1759  0.137096
    8   5    0.471519   636             0             0.44385   0.701741        0.654724        0.0250194       0.136934
    9  -1    0.137342   0               0            0.181818   0.644877        0.677718        0.117089        0.126065
   10   6    0.349367   815             0            0.368984   0.643206        0.699562        0.187026        0.147992
   11   7    0.805063   977             0            0.716578   0.443552        0.91868 0.42619 0.230294
   12   8    0.135443   201             0            0.187166   0.648743        0.833516        -0.126229       0.125092
   13   9    0.205063   1011            0            0.203209   0.49189 0.880679        -0.0276499      0.0849444
   14  10    0.324051   1303            0            0.433155   0.471492        0.668225        0.45033 0.256967
   15  11    0.696203   254             0            0.540107   0.501704        0.589794        0.433911        0.341183
   16  12     0.66519   847             0            0.967914   0.55808 0.781515        0.384089        0.221958
   17  -1    0.686709   0               0            0.235294   1       0.561598        1       1
   18  13    0.125316   1410            0            0.139037   0.654375        0.456851        -0.00668643     0.0775049
   19  14    0.226582   1724            0            0.262032   0.568996        0.673649        0.300322        0.219913
   20  -1    0.493038   0               1            0.187166   0.70236 0.50679 0.0765056       0.187931
   21  15    0.317722   1970            0                   1   0.678139        0.622722        0.141545        0.170904
   22  16    0.649367   252             0            0.278075   0       0       0       0

The total number of disappeared filaments : 1

** Tracking Result 0: means Appeared
** Tracking Result 1: means Disappeared - DA
```

Figure 6.5. Screen dump of the result of the code execution for 2nd and 3rd Feb., 2001.

Table 6.4 contains the results from testing the NN. It shows the classification values that are compared with the required values given in Table 6.3 in order to determine the disappearance detection rate. Agreement is found when the values in Table 6.4 are greater than 0.5 while the values in Table 6.3 are 0.9 and also the values that are less than 0.5 in Table 6.4 while the values in Table 6.3 are 0.1. This comparison gives 16 matching results out of 20, which equals an 80% confidence level.

Table 6.4. Test results from the NN.

Area	Length	Mean	Standard Deviation	Skewness	Kurtosis	Classification
0.937741	0.357771	0.792345	0.69617	0.090797	0.237719	0.0294087
0.375629	0.832335	0.534466	0.474628	0.354462	0.431021	0.375232
1	0.315412	0.43658	0.665064	0.142854	0.086056	3.55E-05
0.371573	0.57764	1	0.522092	1	1	0.255865
0.503686	0.836207	0.604092	0.689848	0.0177343	0.129573	0.456521
0.331754	1	0.477922	0.641134	0.364544	0.270865	0.132397
0.487463	0.391608	0.763976	0.434676	0.910812	1	0.268917
0.313131	0.62037	0.540518	0.682912	0.0186854	0.0686932	0.72257
0.111579	1	0.345268	0.484412	1	0.75767	3.03E-07
0.223774	0.490741	1	0.597617	0.219854	0.224239	0.475688
0.631246	0.547038	0.665606	1	0.235759	0.160758	0.999993
0.350589	0.4	0.638337	0.681501	0.27138	0.189881	0.654153
0.212212	0.745946	0.869229	0.590798	0.0678205	0.195042	0.50409
0.34753	0.421941	0.596294	0.630554	-0.0687134	0.0536176	0.591079
0.570649	0.302752	0.542807	0.489823	0.362108	0.604928	0.999908
0.175296	1	0.131282	0.556901	0.512134	0.486516	1
1	1	0.219364	0.721113	1	1	0.781505
0.77796	0.4	0.840255	0.568293	-0.691428	0.570762	3.11E-14
1	0.458937	0.743717	0.521591	0.743333	0.914094	6.60E-06
0.22876	0.399679	0.651975	1	0.55278	0.675372	0.435708

6.4 Evaluation Results

The results of the Atoum algorithm are compared with the results from Bernasconi (Bernasconi *et al*., 2005) by applying the method to the Hα solar images from the Big Bear Solar Observatory (BBSO), for the same period, from July 6, 2000 to January 9, 2005, that was examined by Bernasconi. They presented their results as the number of appeared filaments. After running their code, 9,459 disappeared filaments were computationally found, out of the 19,211 filaments processed. This means that the percentage of disappeared filaments found by this tracking approach was equal to approximately 51%. Using the NN algorithm, the number of disappeared filaments was found to by 418 out of 22,213 processed for the same period. This means that the percentage number of disappeared filaments found by the NN algorithm is approximately 2%. These Bernaconi and Atoum results show a great variation. Both results can be compared with the number of disappeared filaments registered during the same time period in the NGDC manually compiled catalogue. Counting the

number of disappeared filaments shown in the NGDC catalogue gives 691. This is much closer to the 418 disappeared filaments resulting from the NN method than the 9752 resulting from Bernasconi method which shows the former is much more consistent with the manual results recorded in the NGDC catalogue, which was the aim of the method**.** Since both the Bernasconi and Atoum methods detect broadly similar numbers of filaments, 19,211 and 22,213, over this period the difference may be due to the different treatment of small filaments.

6.4.1 Evaluation of Excluding Small Filaments

To exclude the small erupted filaments from consideration by the NN, four different conditions were compared to find the one that gave the closest results to the manual NGDC result. These conditions exclude filaments with less than half the average area of the whole filaments in the first image, filaments with less than half the average length of the whole filaments in the first image, combine the former two conditions together and finally apply no condition. The disappeared filament results for the whole year of 2001 were collected and compared with the results of the NGDC and the SWRL catalogues for the same year. The average number of disappearances found applying each of the conditions was compared with the catalogued values. The area condition gave average value nearest to the NGDC average, as shown in Table 6.5.

Table 6.5. Disappeared filament results from applying the four conditions to exclude small filaments for comparison with NGDC and SWRL. The average of the numbers in each column is shown at the bottom of each column.

	Month	Length + Area	Area	Length	Without	NGDC	SWRL
1	Jan_2001	12	24	43	130	25	4
2	Feb_2001	6	18	51	139	8	0
3	March_2001	18	31	47	112	17	1
4	April_2001	18	33	49	112	32	5
5	May_2001	7	24	44	123	16	3
6	June_2001	12	30	54	150	29	5
7	July_2001	11	33	56	148	14	2
8	August_2001	19	30	64	137	27	3
9	Sept_2001	9	27	61	157	7	1
10	Oct_2001	17	37	61	152	34	2
11	Nov_2001	14	39	63	154	31	2
12	Dec_2001	12	26	61	146	36	1
	Average	12.91666667	29.33333	54.5	138.3333	23	2.416667

To show the convergence or divergence between what has been discovered after applying the four different conditions and what is found in the NGDC and SWRL catalogues, several charts where drawn for the month of January 2001. Figure 6.6 compares the data from the NGDC observations and the Atoum algorithm after applying the area condition.

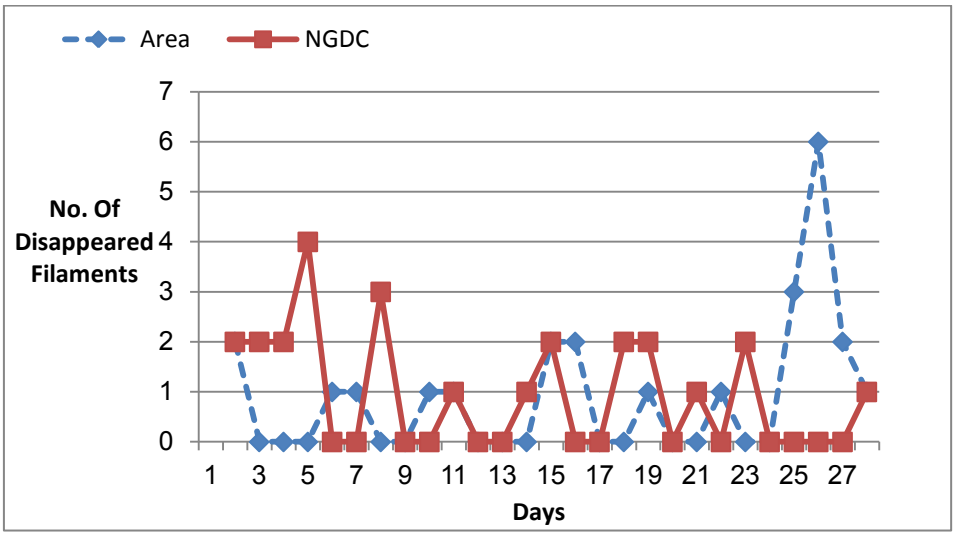
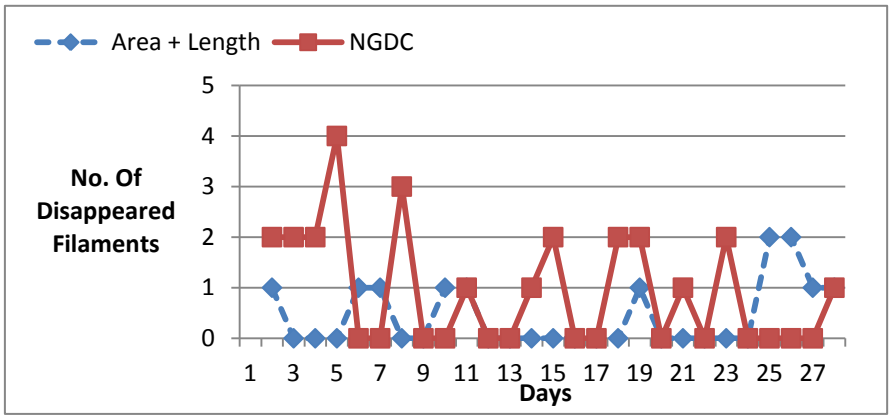


Figure 6.6. The comparison between the NN algorithm results with the area condition and the NGDC Catalogue results for January, 2001.

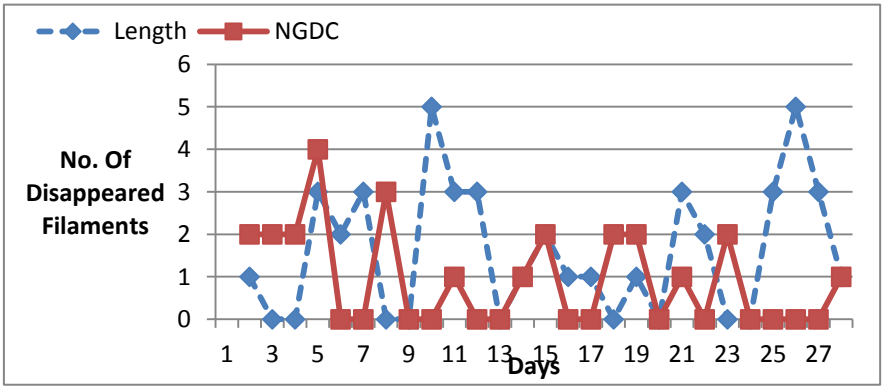
The days of the month are displayed on the horizontal axis and the numbers of disappeared filaments are displayed on the vertical axis in this, and the next four charts. Charts for the other three conditions are shown in Figure 6.7.

Figure 6.7.a shows the relation between the average values of all the disappeared filaments in the NGDC catalogue and the average value of all the disappeared filaments from the NN algorithm, after excluding the filaments with area less than the average area value and with length less than the average length value in the first day image. This figure shows that the daily number of DFs in NGDC differs from the number given by the NN algorithm in both number and date. The latter may be due to different sensitivities to detecting disappearances of the manual and machine methods. Figure 6.7.b shows the relation between the average values of all the disappeared filaments in the NGDC catalogue and the average value of all the disappeared filaments from the NN algorithm, after excluding the filaments with length less than the average length of all the filaments in the first day image. Again, this figure shows high divergence between the two results. Figure 6.7.c shows the relation between the average values of all disappeared filaments in NGDC and the NN algorithm without applying any condition. This figure shows higher divergence between the numbers in the two results which applying the condition is meant to reduce. None of these figures show detailed convergence between the NN algorithms and the results of the NGDC but the area condition is the closest overall.

(a)



(b)



(c)

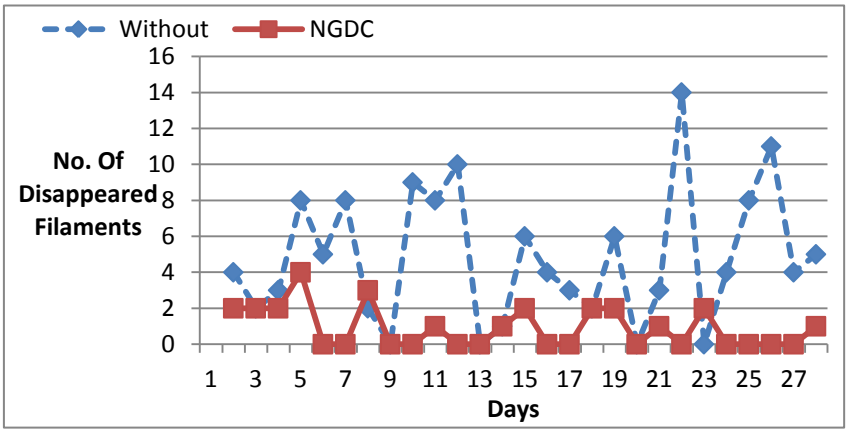


Figure 6.7. Three charts showing the NGDC results for January, 2001 in comparison with those from applying the NN NN algorithm with the three conditions (a) area+length, (b) length, (c) no condition.

Taking the NGDC manual catalogue as a reference, the graph shown in Figure 6.8 was drawn to compare the numbers of disappeared filaments with those from the SWRL automatic catalogues. This shows a similar divergence in positions between the two sets of results as seen in figures 6.6 and 6.7 but the numbers are much less than those in the NGDC catalogue.

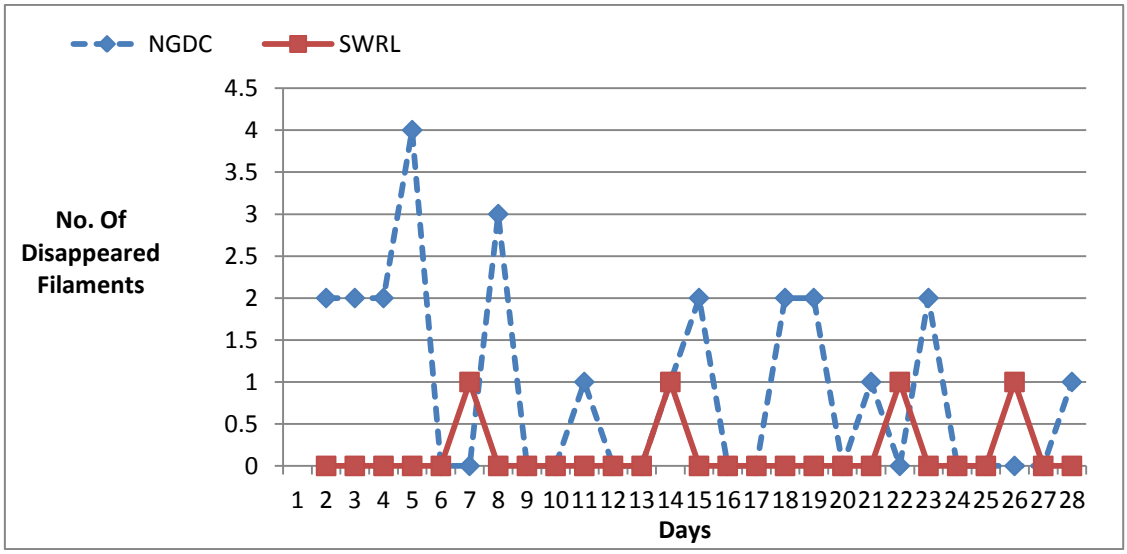


Figure 6.8. The average numbers of all disappeared filaments in the NGDC manual catalogue and the SWRL automatic catalogue for January, 2001.

6.5 Conclusion

A Neural Network classifier is used in solar filament tracking to categorize the detected filaments as DFs or MDFs. Features including Area, Length, Mean, Standard Deviation, Skewness and Kurtosis are extracted from the underlying filament and fed to the NN. This NN is applied to reduce the number of the filaments miss-detected when a simpler binary decision rule was used. The network is optimized by using nine nodes for the hidden layer.

The result is a novel filament tracking algorithm which aims to automatically detect solar filament disappearance in Hα full disk images obtained from different solar observatories. Experimental results show that there is a significantly but small number of miss-detected filament disappearances phenomenon which still needs to be solved.

The algorithm has improved on the previous detection process by now being able to accept, handle and process solar images with different contrasts from several observatories. A solar image is first converted to the HCM format and then filament tracking is accomplished by comparing the contents of the rectangle that encloses a filament in the first image with those in the successive image. The technique continues by distinguishing miss-detected filament disappearance from actual filament disappearance by using the NN with the input features: area, length, mean, standard deviation, skewness, and kurtosis. The result after applying the NN is an 80% detection level.

The new tracking technique improves on the work of others by representing the tracking process more plainly, by considering the miss-detection filaments and by utilizing the adaptive nature of an NN. All of these advantages could lead to a more precise knowledge of the association expected between disappeared filaments and CMEs.

It has, however, to be noted that if there were many middle sized filaments with only one very big filament, then the middle filaments will not be taken into consideration by the algorithm. This is actually a rare case and is not a major problem under normal solar phenomena. The author appears to be the first researcher to actually consider the possibility of this occurrence having any effect on the tracking algorithm; hence no references could be found for this scenario in the literature. This case will be addressed in future work on the NN algorithm.

CHAPTER SEVEN

7 CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

7.1 Conclusions

7.1.1 Overall Conclusions

The main achievement of the research presented in this thesis can be described as a collection of algorithms related to the field of automatic detection of solar filament disappearances, which have been developed using machine learning based technologies. These involved several Neural Network topologies trained using different attributes of solar filaments. It is believed that this work is important because for the first time machine learning-based techniques have been adopted for merging broken filaments and detecting the filament disappearances. There are additionally two aspects that have importance for filament detection techniques, namely the implementation of a fully automated system without using any predetermined empirical thresholds and the real time functionality of the algorithm. These findings are considered an important step towards creating automated and reliable prediction systems for CME.

7.1.2 Detailed Conclusions

The concluding remarks on this research are listed as follows.

- Segmentation of solar filaments constitutes an important and difficult concern in the field of solar image processing. Therefore, an adaptive thresholding technique was devised that introduced minimal noise whilst also being able to detect more unambiguous filaments, with the constraint of avoiding the use of any empirical

values in Hα solar images. This work was described in Chapter Four with performances measured in terms of the FAR, which is the probability of a non-filament being detected as a filament. The resultant FAR value was found to be 9% when compared with another segmentation algorithm over one month of solar images. Additionally reducing the FAR value means reducing the proportion of the presence of other solar features like sunspots. This value is lower than the work of others but is still greater than zero, which means that there are still some non-filaments detected as filaments. This segmentation technique could be used to highlight the dark features (like filaments in Hα images) in any cleaned image and eliminate the background.

- A fully automated spine detection technique is presented in Chapter Four. It is apparent that the Atoum algorithmically derived spine is more convoluted, precisely because it accurately follows the body of the filament. It is less computationally complex compared to Bernasconi (Bernasconi *et al*., 2005). On average the Atoum algorithm was 3.5 times quicker than Bernasconi, taking only 0.098 s to extract the filament spine. Additionally, the results of the Atoum algorithm show longer spine lengths because it tracks the actual filament backbone more accurately. Despite the positive results from the detection technique, there is still a need to reduce the unwanted pixels and the small holes. Furthermore, there are incomplete spines existing due to errors in choosing the largest line segment through the automated drawing process. The technique could be adopted in future to extract the spine of different features in different types of images; especially elongated objects such as the spine of neuronal dendrites.

- It was clear from the literature that the author had to include the spatial and orientation relationships between the regions of the segmented filaments so as to

determine if they should be merged or not. There was one technique developed by Bernasconi (Bernasconi *et al.*, 2005) which considered these issues in merging the broken filaments but there was a challenge to avoid using empirical thresholds for the angles and distances used in this method. The Atoum merging algorithm achieves 15% greater merging of broken filaments than the Bernasconi algorithm. This improved performance was obtained by using a neural network approach and also avoids the thresholds used in traditional techniques. Although there is a percentage of false merges still remaining, it should be noted that this work has highlighted this challenge of the existence of false merging - often overlooked by other researchers. The same method could be used in future to merge broken structures in different images.

- A novel tracking method was implemented in Chapter Six that achieved a fully automated real-time method; coping with the movement of filaments upward and downward and most importantly addressing the problem of the presence of filaments that are considered disappeared, but are not in fact. All these challenges were solved by using a Neural Network approach. The results of the NN algorithm were compared with the results of Bernasconi (Bernasconi *et al.*, 2005) and in the NDGC catalogue. The NN approach under detects the number of disappeared filaments by 40% while Bernasconi's algorithm over detects by over 1400%. This clearly shows the superiority of the NN algorithm in reproducing the NDGC results. Although the NN algorithm has given improved results in terms of accuracy, there still exist a small percentage of filaments that are detected as disappeared when actually they have not and also filaments that have actually disappeared but are not detected as such.

Given time, a CME prediction system could be implemented based on the findings of this work. The results of detecting filament disappearances could then be compared with the data of NGDC manual catalogue of filament and the data in the SOHO/LASCO CME manual catalogue. This could be achieved through combining the present work with that of Al-Omari (Al-Omari *et al.*, 2010).

7.1.3 Knowledge Discovery Resources

In this research, a wide range of experience was gained in through making full use of the following resources:

- Many sources of solar data were considered for use. Images from several ground-based observatories including the Meudon and BBSO Observatories were used as well as the NGDC and SWRL catalogues of solar filament disappearances.
- Many different features of Visual Studio C++ were used. All the algorithms developed in this research were designed using the C++ programming language.
- The adaptive nature of the developed algorithms was produced using Neural Network algorithms.

7.2 Suggestions for Future Work

Some of the challenges that still need to be overcome with suggested solutions and some ideas for further research are included in the following list.

- As concluded in Chapter Four, there is a problem which has emerged while extracting the filament spine. The final phase of the spine algorithm is to achieve the task of averaging and projection. For each point in the largest line segment, the mid-point of the perpendicular line segments that intersects this line is found; i.e. if the first drawn line is horizontal then the projection will be achieved by finding all the mid-points of the vertical lines that intersect this line. This phase sometimes

produces and incomplete or a ‘stuck spine’ due to errors in choosing the largest line segment through the automated drawing process.

- Although the merging algorithm gives a high merging percentage; there is still a percentage of false merging remaining. Either these broken filaments represent one filament and they are combined wrongly or they are parts of two different filaments and are combined together incorrectly.
- Another challenge which has been observed while implementing the tracking algorithm is the existence of miss-detected filament. These filaments which are classified as having being disappeared have not actually disappeared at all because the detection algorithm has miss-identified them; there is no ideal filament detection technique. The present study seems to be the first in distinguishing the actual disappeared filaments and the miss-detected filaments by applying ANNs. Even so, a percentage of these filaments are detected as disappeared.
- One further challenge in the area of detecting filament disappearances is applying a backward tracking method by the construction of temporal evolution patterns. This could be implemented by comparing the current day image with the previous day image to detect newly appeared filaments.
- An additional challenge is improving the system by developing an automatic prediction technique for CME-filament association. The technique will start with comparing the disappeared filaments with the detected CMEs manually from a CME catalogue. If there is an association between them then a machine learning algorithm will be developed to extract the features of these filaments which will then be processed by the machine learning algorithm to predict the occurrence of a CME.

References

Aboudarham, J., Scholl, I., Fuller, N., Fouesneau, M., Galametz, M., Gonon, F., Maire, A., and Leroy. Y., (2008) Automatic Detection and Tracking of Filaments for a Solar Feature Database, *Annales Geophysicæ*, **26** (2), pp. 243-248.

Alejandro L., (2008) The Source Region of Coronal Mass Ejections, *The Astrophysical Journal*, **688** (1), pp. 647-655.

Al-Omari M., Qahwaji R., Colak T. and Ipson S., (2010) Machine Learning-based Investigation of the Associations between CMEs and Filaments, *Solar Physics*, **262** (2), pp. 511-539.

Atoum, I. A., Qahwaji, R. S., Colak, T., and Ahmed, Z. H., (2009) Adaptive Thresholding Technique for Solar Filament Segmentation, *Ubiquitous Computing and Communication Journal*, **4** (4), pp. 91-95.

Attrill, G., Nakwacki, M. S., Harra, L. K., Van Driel-Gesztelyi, L., Mandrini, C., H. Dasso, S. and Wang, J., (2006) Using the Evolution of Coronal Dimming Regions to Probe the Global Magnetic Field Topology, *Solar Physics*, **238** (1), pp. 117-139.

Bernasconi, P. N., Rust, D. M., and Hakim, D., (2005) Advanced Automated Solar Filament Detection and Characterization Code: Description, Performance and Results, *Solar Phy*sics, **228** (1-2), pp. 97-119.

Cheng, J., Xhou, X., Miller, E., Witt, R. M., Zhu, J., Sabatini B. L., and Wong, S. T., (2007) A Novel Computational Approach for Automatic Dendrite Spines Detection in Two-photon Laser Scan Microscopy, *J. Neurosci. Methods*, **165** (1), pp. 122-34.

Colak T. and Qahwaji R., (2007) *Automated Prediction of Solar Flares Using Neural Networks and Sunspots Associations,* Advances in Soft Computing, *Springer,* **39**, pp. 316-324.

Colak, T. and Qahwaji, R., **(**2009) *Automated Solar Activity Prediction: A Hybrid Computer Platform Using Machine Learning and Solar Imaging for Automated Prediction of Solar Flares, Space Weather,* **7** (6), S06001, doi: 10.1029/2008SW000401.

Colak, T., Qahwaji, R., Ipson, S. and I. Ugail, H., (2011) Representation of Solar Features in 3D for Creating Visual Solar Catalogues, *Advances in Space Research,* **47** (2), pp. 2092-2104.

Denker, C., Johannesson, A., Marquette, W., Goode, P. R., Wang, H. and Zirin, H., (1998) Synoptic Hα Full-Disk Observations of the Sun from BigBear Solar Observatory, *Solar Physics,* **184** (1), pp. 87-102.

Efford, N., (2000) Digital Image Processing, A Practical Introduction Using Java, 1st edn. Pearson Education Limited, Addison Wesley, USA.

Ermolli I., Solanki S. K., Tlatov A. G., Krivova N. A., Ulrich R. K. and Singh J., (2009) Comparison among CA II K Spectroheliogram Time Series with an Application to Solar Activity Studies, *The Astrophysical Journal*, **698** (2), pp. 1000-1009.

Freeland, S. L., and Handy, B. N., (1998) Data Analysis with the SolarSoft System, *Solar Physics*, **182,**(2), pp. 497-500.

Fuller, N. and Aboudarham, J., (2004) Automatic Detection of Solar Filaments versus Manual Digitization, *Knowledge-Based and Intelligent Information & Engineering Systems, Lecture Notes in Computer Science*, Springer Berlin Heidelberg, **3215**, pp. 467-475.

Fuller, N., Aboudarham, J. and Bentley, R. D., (2005) Filament Recognition and Image Cleaning on Meudon Hα Spectroheliograms, *Solar Physics,* **227** (1), pp. 61-73.

Gao, J., Wang, H. and Zhou, M., (2002) Development of an Automatic Filament Disappearance Detection System, Solar Physics, **205** (1), pp. 93-103.

Gopalswamy, N., (2006) Coronal Mass Ejections of Solar Cycle 23, *J. Astrophys Astr*., **27** (2), pp. 243–254.

Gopalswamy, N., Shimojo, M., Lu, W., Yashiro, S., Shibasaki, K. and Howard, R. A., (2003) Prominence Eruptions and Coronal Mass Ejection: A Statistical Study Using Microwave Observations, *Astrophysical Journal*, **586**(1), pp. 562-578.

Hearn, D. and Baker, M.P., (1997) Computer Graphics C Version, 2nd edn., USA, Prentice Hall.

Hong L. and Jain A., (1998) Integrating faces and fingerprints for personal identification, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **20** (12), pp. 1295-1307.

Ingrid Y., Koh, Y., Lindquist, W. B., Zito, K., Nimchinsky, E. A. and Svoboda, K., (2002) An Image Analysis Algorithm for Dendritic Spines, *Neural Computation*, **14** (6), pp. 1283–1310.

Jing, J., Yurchyshyn, V. B., Yang, G., Xu, Y. and Wang, H., (2004) On the Relation between Filament Eruptions, Flares and Coronal Mass Ejections, *Astrophysical Journal*, **614** (2), pp. 1054-1062.

Joshi, A. D., Srivastava N. and Mathew, S. K., (2009) Automated Detection of Filaments and their Disappearance using Full-Disc Hα Images, *SolarPhysics*, **262** (2), pp. 425-436.

Kom, G., Tiedeu, A. and Kom, M., (2007) Automated detection of masses in mammograms by local adaptive thresholding, *Computers in Biology and Medicine*, **37** (1), pp. 37-48.

Lener, B., Guterman, H., Aladjem, M. and Dinstein, I., (1999) A Comparative Study of Neural Network Based Feature Extraction Paradigms, *Pattern Recognition Letters*, **20** (1), pp. 7-14.

Mao, J. and Jain, A., (1995) Artificial Neural Networks for Feature Extraction and Multivariate Data Projection, *IEEE Transactions on Neural Networks*, **6** (2), pp. 296-317.

NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/, [Last accessed 10th April, 2012].

Moon, Y., Choe, Y. G., Haimin, Wang, Park Y. D. and Gopalswamy N., (2002) A Statistical Study of Two Classes of Coronal Mass Ejections, *The Astrophysical Journal*, **581** (1), pp. 694–702.

National Solar Observatory, (1996) *Ca K and H Alpha Images Explained*, [ONLINE] Available at: http://eo.nso.edu/MrSunspot/answerbook/cak_ha_expl.html, [Last Accessed 9th April, 2012].

Pojoga, S. and Huang, T. S., (2003) On the Sudden Disappearances of Solar Filaments and their Relationship with Coronal Mass Ejections, *Adv. Space Res.*, **32** (12), pp. 2641-2646.

Qahwaji, R. and Colak, T., (2005) Automatic Detection and Verification of Solar Features, *Int. J. Imaging Systems Tech.*, **15** (4), pp. 199-210.

Qahwaji, R. and Colak**,** T., (2006) *Neural Network-Based Prediction of Solar Activities, 3rd International Conference on Cybernetics and Information Technologies,Systems and Applications*, pp. 192-195, Florida.

Qu, M., Shih, F. Y., Jing, J. and Wang, H., (2005) Automatic Solar Filament Detection Using Image Processing Techniques, *Solar Physics*, **228** (1-2), pp. 119-135.

Robbrecht, E., Berghmans, D. and Van der Linden, R. A. M., (2009) Automated LASCO CME Catalog for Solar Cycle 23: Are CMEs Scale Invariant?, *The Astrophysical Journal*, **691** (2), pp. 1222-1234.

Schmieder, B., (2006) Magnetic Source Regions of Coronal Mass Ejections, *J. Astrophysical Astr*., **27** (2&3), pp. 139-149.

Shih, F, Y. and Cheng, S. (2004). Adaptive Mathematical Morphology for Edge Linking. *Information Sciences*. **167**(1-4), pp. 9-21.

Shih, F. Y. and Kowalski, A. J., (2003) Automatic Extraction of Filaments in H-alpha Solar Images, *Solar Phys*ics, **218** (1-2), pp. 99-122.

Smart, W.M, and R. M. Green., (1977) *Textbook on Spherical Astronomy*, 6th ed., England, Cambridge University Press.

U. S. Air Force Fact Sheet, (2010) *2nd Weather Squadron, Solar Observing Network*, [ONLINE] Available at: http://www.afweather.af.mil/library/factsheets/factsheet_print.asp?fsID=16522&page=1, [Last Accessed 9th April, 2012].

Yilmaz A., Javed O., Alper Y. and Shah M., (2006). Object Tracking: A Survey, *ACM Computing Surveys (CSUR)*, **38** (4), pp. 45.

Yuan, Y., Shih, F. Y., Jing, J., Wang, H. and Chae, J., (2011) Automatic Solar Filament Segmentation and Characterization, *Solar Physics*, **272** (1), pp. 101-117.

Zharkova V. V. and Schetinin V., (2003) A Neural Network Technique for Recognition of Filaments in Solar Images, Knowledge-Based Intelligent Information & Engineering Systems Lecture Notes in Computer Science, Springer, Berlin, Heidelberg. **2773**, pp. 148-154.