National Observatory of Athens
Department of Informatics and Telecommunications -
University of Peloponnese

Master of Science in Space Science Technologies and Applications.

Eirini Angeloudi
MSc Candidate
2022201803001

Development of Automated methods for the detection of solar vortex flows

Master Thesis

Advisory Committee

Konstantinos Tziotziou
Konstantinos Koutroumbas
Georgia Tsiropoula

Athens, 2020

Abstract

The Sun's surface is very dynamic, even on periods of weak solar activity, producing a variety of magnetic structures, amongst them structures that can cause plasma to exhibit vortex-like motions. These so-called solar vortex flows have been widely observed on the Sun's photosphere and chromosphere as well as on upper solar atmospheric levels. Their existence is highly connected with the creation of magnetohydrodynamic (MHD) waves, a primary candidate for channeling energy and plasma from the photosphere up to the low corona. Therefore, the automatic detection of solar vortex flows is fundamental for proper investigation of the lower solar atmosphere dynamics and the estimation of upwards energy transfer. Previous detection approaches utilized a criterion, like vorticity, based on the frequently inconsistent and often erroneous derivation of the velocity field from the solar observations. The aim of the present thesis is the implementation of an automated methodology for the detection of solar vortex flows that is based on the morphological characteristics they present.

The approach followed is the application of a series of pre-processing methods aiming to remove non-dynamic or almost linear structures not related to vortex flows from a solar data set (consisting of a sequence of frames depicting a specific solar region). Such methods include the Robust Principal Component Analysis (RPCA), in order to retrieve the dynamic parts of the solar data as well as the removal of components that relate to almost linear structures and thus are not part of a vortex flow. Subsequently, the output produced by the sequential application of the pre-processing methods is aggregated and evaluated so that the primary candidate region for vortex flow existence is detected.

The evaluation of the produced results on an actual solar data set proves that the proposed methodology succeeds in identifying the primary solar vortex flow region. More importantly, a promising analysis of the distribution of dynamic, vortex-like structures on the observational sequence is produced that can be used, by further future extensions of the current work, for the detection of multiple structures within the investigated field of view. It is concluded that the presented methodology constitutes an efficient alternative way for the identification of solar vortex flows, especially after all proposed improvements and extensions are properly implemented.

Acknowledgements

I would like to thank Dr. Konstantinos Tziotziou for giving me the chance to work on such an interesting topic as well as for his constant and valuable guidance. I would also like to express my gratitude to Dr. Konstantinos Koutroumbas for his precious feedback and the thoughtful comments and recommendations throughout this MSc thesis. I am extremely thankful to both of them for always being available to discuss and provide their continuous ideas on how to move forward.

I would also like to thank my family and friends for their unconditional support and their constant encouragement in this very intense academic year.

Contents

Abstract **i**

Acknowledgements **ii**

1 Introduction **1**

1.1 Motivation 1

1.2 Aims and Objectives 2

1.3 Overview of the structure 3

2 Physical Problem **5**

2.1 Overview of Solar Vortex Flows 5

2.2 Formation of vortex flows 6

2.3 Observational detection of vortex flows 9

2.4 Observations of vortex flows 12

2.5 Simulations of vortex flows 14

2.6 Characteristics of the Studied Vortex flow 15

3 State of the Art **17**

4 Proposed Detection Methods **24**

4.1 Robust Principal Component Analysis (RPCA) 25

4.2 Filtering 29

4.2.1 Canny Edge Detector 29

4.2.2 Sobel Operator 32

4.2.3 Other filters 33

4.3 Removal of “almost linear” structures 34

4.4 Evaluation mechanism 37
4.5 Toy models 39
4.6 Computational tools and characteristics 40

5 Evaluation results **41**
5.1 Robust PCA 41
5.2 Filters 48
5.3 Removal of "almost linear" structures 53

6 Swirl detection methodology **56**
6.1 Detection mechanism 56
6.2 Toy models 67

7 Discussion and Future Work **72**

8 Conclusions **76**

Bibliography **77**

List of Figures

2.1 The solar structure 6
2.2 Solar granulation 8
2.3 Vortex flows formation 9
2.4 Sun at different wavelengths 10
2.5 Photospheric vortex flows 13
2.6 Chromospheric vortex flows at Ca II 854.2 nm 14
2.7 Studied Vortex Flow 16

3.1 Enhanced vorticity method and vorticity strength method comparison . . . 19
3.2 LAVD vortex detection method 20
3.3 Vorticity strength method, LAVD and Γ-detection comparison 21

4.1 Diffence of Gaussians (DOG) filter 34
4.2 Rolling window 38

5.1 Frames of the original data depicting the conspicuous vortex at various stages. The primary vortex flow is highlighted for easier identification. . . . 42
5.2 The original frame 174 as well as the Sparse and Low component that were resolved after the RPCA application. 44
5.3 The original frame 302 as well as the Sparse and Low component that were resolved after the RPCA application. 45
5.4 Frames depicting the Sparse component as it was resolved after the RPCA application on the original data set frames of Figure 5.1. 46
5.5 Frames depicting the Low-rank component as it was resolved after the RPCA application on the original data set frames of Figure 5.1.. 47

5.6 Frames of the data (similar to those of Figure 5.1) after applying the Canny filter depicting the conspicuous vortex at various evolutionary stages. 49
5.7 Frames of the data (similar to those of Figure 5.1) after applying the Sobel filter depicting the conspicuous vortex at various evolutionary stages. 50
5.8 Frames of the data (similar to those of Figure 5.1) after applying the DOG filter depicting the conspicuous vortex at various evolutionary stages. 51
5.9 Frames of the data (similar to those of Figure 5.1) after applying the Mad-max filter depicting the conspicuous vortex at various evolutionary stages. . 52
5.10 Frames of the Canny processed data before (left column) and after the "almost linear" removal procedure (right column). 54
5.11 Frames of the Sobel processed data before (left column) and after the "almost linear" removal procedure (right column). 55

6.1 Sample frames (similar to those in Figure 5.1) of the Sparse component of the Canny processed data after the "almost linear" removal procedure. . . . 58
6.2 Sample frames (similar to those in Figure 5.1) of the Sparse component of the Sobel processed data after the "almost linear" removal procedure. . . . 59
6.3 The mean frame F_{mean} of the Canny Sparse component's frames (after the "almost linear" removal). 60
6.4 The mean frame F_{mean} of the Sobel Sparse component's frames (after the "almost linear" removal). 60
6.5 The Probability Distribution Image from the Canny mean frame for various window sizes and shapes 61
6.6 The Probability Distribution Image from the Sobel mean frame for various window sizes and shapes 62
6.7 The region corresponding to the maximum square window value for the Canny F_{mean}. 63
6.8 The region corresponding to the maximum circular window value for the Canny F_{mean}. 64
6.9 The region corresponding to the maximum square window value for the Sobel F_{mean}. 65

6.10 The region corresponding to the maximum circular window value for the Sobel F_{mean}. 66
6.11 Frames of the toy model featuring a ring rotating and swinging diagonally. . 68
6.12 Frames of the toy model featuring a vortex rotating and swinging diagonally. 69
6.13 Frames of the toy model featuring 2 vortices and dynamic linear structures. 70
6.14 Comparison of the solar data sparse component after RPCA and a vortex toy model. Both are visualized in tall image depictions (see text), where each column holds the pixels of a single frame. The patterns expected from the toy model are also visible in the actual solar data. 71

Chapter 1

Introduction

1.1 Motivation

Without any doubt, the Sun constitutes the most extraordinary object in our solar system. The importance of our nearest star has long been recognized and acknowledged as vital for the creation of life on Earth. Throughout human history, people have been unceasingly observing the Sun and attempting to explain its numerous mysteries. Even today, the Sun remains one of the most fascinating fields of study in Astrophysics and constitutes a cosmological laboratory for understanding the physics that also govern the dynamics of other stars.

Upon the discovery of the first telescopes, scientists were able to observe the Sun from a closer view. In recent years, more detailed images were obtained with solar telescopes and dedicated instruments launched into space. A wild and unpredictable environment was revealed by examining the Sun in different wavelengths also beyond the visible spectrum. Numerous complex phenomena on the Sun surface and atmosphere have been widely observed and the underlying mechanisms that create them have been thoroughly investigated. Such phenomena include the well-known Sunspots in active regions of the Sun and the most recently observed and less studied solar vortex flows that constitute part of the quiet Sun fine solar structure.

However, the Sun still holds secrets that scientists seek to unlock. One of the greatest ones is widely known as the coronal heating problem and revolves around the mechanisms that contribute to the rapid increase of the Sun's corona temperature to millions of degrees, a

temperature 200 to 500 times higher than that on the surface. There exist many theories attempting to explain the coronal heating suggesting several factors that contribute to this sudden increase of temperature. Among them, it has been widely suggested that a considerable amount of energy is transferred to the corona through magnetohydrodynamic (MHD) waves. A substantial fraction of such waves could be produced by ubiquitous solar vortex flows existing on the Sun's surface and its chromosphere.

Solar vortex flows have been observed through several, mainly ground-based, instruments viewing the Sun at different wavelengths and exhibit different (mainly small) spatial scales and lifetimes (see following chapter for details). Their exact occurrence rate is still unknown and differs in most studies. More importantly, solar vortex flows are observed both on the Sun's photosphere and chromosphere and in some cases also spanning to higher atmospheric layers up to the low corona. This is very intriguing since vortex flows could therefore provide a mechanism for transferring energy, mass and momentum from the subsurface levels of the Sun higher up.

It is thus clear that solar vortex flows constitute a very important part of the Sun's fine-scale environment. Due to their smaller scale in comparison e.g. to Sunspots, their detection is a far more difficult task. Nonetheless, it is imperative that solar vortex flows on the Sun's surface and atmosphere are more broadly detected, so that an in-depth study of their characteristics as well as their contribution to the solar small-scale dynamics can be better established.

1.2 Aims and Objectives

Datasets including time series of solar observations on a specific wavelength for a certian region of the Sun, where vortex structures sometimes can be detected even by visual inspection, can be acquired from a variety of instruments attached on solar telescopes. However, due to the huge volume of such data, the detection of vortex structures by visual inspection is infeasible. Thus, the need for automated processing methods arises. The aim of the current thesis is the development of an automated methodology for the detection of solar vortex flows.

The proposed automated method receives as an input a time series consisting of grey-

scale images depicting a specific region of the Sun. The grey-scale data correspond to the amount of photons at the observed wavelength, that were collected by the detection instrument for a restricted period of time. In chromospheric lines, such as H-alpha and Ca, the data are very dense, complex and detailed, containing information that does not facilitate the detection of vortical motions as they include several other fine-scale structures and dynamic phenomena.

To this end, the first and most important step on the proposed detection method is the removal of excessive information from the dataset that is irrelevant to vortex flows. This is achieved by utilizing a number of techniques including robust principal component analysis (RPCA), filtering and removal of linear features from the data set, as it will analysed in detail further on. The processed data set is then evaluated and the primary candidate region for the existence of a vortex flow is detected.

It is worth noting that, unlike the majority of solar vortex detection methods that have been proposed up to this date, which depend on velocity criteria (e.g. [1], [2], [3]), the proposed method extracts information using the morphological characteristics of vortex flows.

1.3 Overview of the structure

In the current subsection, a quick overview of the thesis structure is provided.

The second chapter contains a brief description of the physical problem of solar vortex flows. The mechanism that produces such vortex flows is discussed. The characteristics of the main wavelengths that solar vortex flows have been observed so far are described shortly and an overview of all observations up to this date is presented. Lastly, the characteristics of the specific vortex flow used as input for the developed detection methods are briefly discussed.

The third chapter provides a brief overview of the state of the art concerning automated vortex flow detection methods and of the different practices that have been employed so far. Also, a short comparison between those methods and the method proposed in the current thesis is provided.

In the fourth chapter, the various methods utilized for processing and extracting informa-

tion from the solar data are discussed. The combination of the different techniques in order to create a step by step detection methodology is described in detail. Also, the evaluation of the proposed swirling detection methodology is discussed in this chapter.

The fifth chapter includes the results produced from the application of each one of the knowledge extraction techniques discussed on the previous section on the original data set. The pre-processing methods are also combined and run sequentially following the proposed swirling detection methodology. The results produced as well as the output of the presented evaluation mechanism are presented in the sixth chapter.

In the seventh chapter, the proposed methodology for the detection of solar vortex flows and the results it produces are discussed in detail. Lastly, thoughts on further extensions of the work are also included in the seventh chapter. The eight chapter contains the conclusions and final thoughts on the present thesis's aims and contributions.

Chapter 2

Physical Problem

2.1 Overview of Solar Vortex Flows

The surface of the Sun, known as the photosphere, is characterized by constant motion, emergence of small-scale and larger scale magnetic fields and a variety of often magnetic-related plasma flows. The most common features observed in the photosphere even by naked eye are Sunspots that correspond to large-scale magnetic structures appearing on the solar surface [4]. More detailed views of the Sun's surface also reveal granular motions, which originate from the convection zone of the Sun that lies below the photosphere. Bright granules represent motions of hot plasma rising from the convection zone that cools down when it reaches the solar surface and then descends in the intergranular lanes often in a vortex-like motion [5]. This phenomenon, also known as the bathtub effect, produces intergranular vortex flows, that upon the right conditions (which will be further analyzed in a following section) can expand to the chromosphere, transition region, and even the low corona in the form of atmospheric vortex flows [6].

Such vortex flows are present both on active regions of the Sun but mainly on regions that demonstrate weak solar activity, widely known as the quiet Sun. They have been detected at different layers of the Sun's atmosphere at a variety of temporal and spatial resolutions [7]. Such vortex flows also appear in most simulations of the Sun surface and atmosphere, as it will be discussed later on and their existence has been also anticipated by theory [8].

The study of solar vortex flows in the photosphere and the chromosphere is important as it

can reveal an alternative mechanism of channeling energy from the lower layers of the solar atmosphere to its upper ones. The exact mechanisms that contribute to the rapid increase of the temperature to millions of degrees on the Sun's corona, a temperature 200 to 500 times higher than the one on the surface, are still very poorly understood. Consequently, the detection and study of such vortex flows could shed further light on the heating of the Sun corona and provide a complementary explanation on this mystery.

2.2 Formation of vortex flows

The Sun can be divided in the following regions in terms of structure: the core, the radiative zone, the convection zone, the visible surface (photosphere) and the atmosphere. The solar atmosphere comprises of the chromosphere, the solar transition region and the corona. A schematic representation of the solar structure can be found in Figure 2.1.

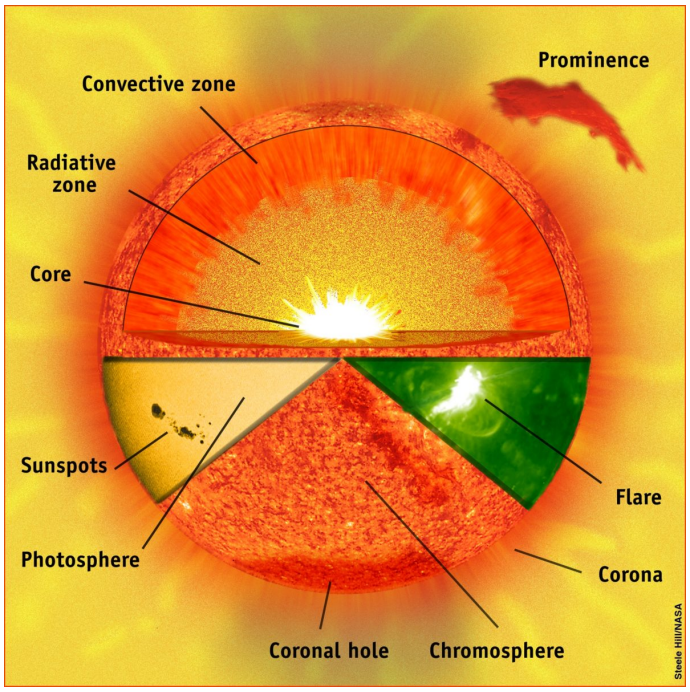


Figure 2.1: The solar structure

In the core of the Sun, where the temperature is about 15 millions degrees Celsius, thermonuclear fusion takes place. The energy produced in the core is transferred via photons taking about 170,000 years to get from the core to the top of the convection zone. This is due to the high density of the radiation zone that results in bouncing effects by continuous absorption and re-emission of the photons.

The convection zone is the outer-most layer of the interior of the Sun, where convective currents develop and transfer energy towards the solar surface. Plasma-β, i.e., the ratio of the plasma pressure ($p = nk_BT$) to the magnetic pressure ($p_{mag} = B^2/2\mu_0$) and is high in the convection zone. Moving towards the surface and continuing through the atmospheric levels, the value of plasma-β gradually drops, as the density of plasma decreases far more quickly than the corresponding decrease of the magnetic field. Plasma-β is a fundamental physical quantity concerning stellar atmospheres as it is an indicator of the balances between magnetic versus thermodynamic processes on the plasma motion [9].

The process of solar vortex flow formation begins when material at the transition layer between the radiation and the convection zone heats up and expands, rising towards the surface. As a result, thermal cells are produced that move upwards and finally reach the photosphere. Those cells are known as granules and are responsible for the grainy appearance of the solar photosphere. A high resolution image, where the aforementioned granulation on the solar surface is apparent, taken by the Daniel K. Inouye Solar Telescope (DKIST) on January 20, 2020 can be found in Figure 2.2.

Within the granules, warm plasma ascends to the solar surface, where it cools radiatively just beneath the photospheric surface and its density increases, causing it to sink back into the convection zone within the intergranular lanes. There, it heats up again and the process repeats, effectively creating the solar convective cycle [10].

The formation of solar vortex flows is tightly connected with the aforementioned process. As it is described in [11], the conservation of angular momentum dictates that when cool plasma sinks down through the intergranular lanes, a mechanism producing a vortex flow is often initiated that highly resembles the swirling patterns the water manifests on a sink of a bathtub. Due to this resemblance, this hydrodynamic phenomenon is commonly referred to as the “bathtub” effect. For further reference, the hydrodynamics concerning the solar convection have been analyzed in depth in [12].

Such solar vortex flows that mainly expand vertically have been called intergranular vortex flows (IVFs) in [11] and also referred to as “inverted tornadoes” in [12]. They are expected to appear with greater probability at the vertices of integranular lanes due to the convergence of plasma from neighboring granules. At those vertices, both horizontal and vertical vortex flows are generated. However, the horizontal flows frequently sink

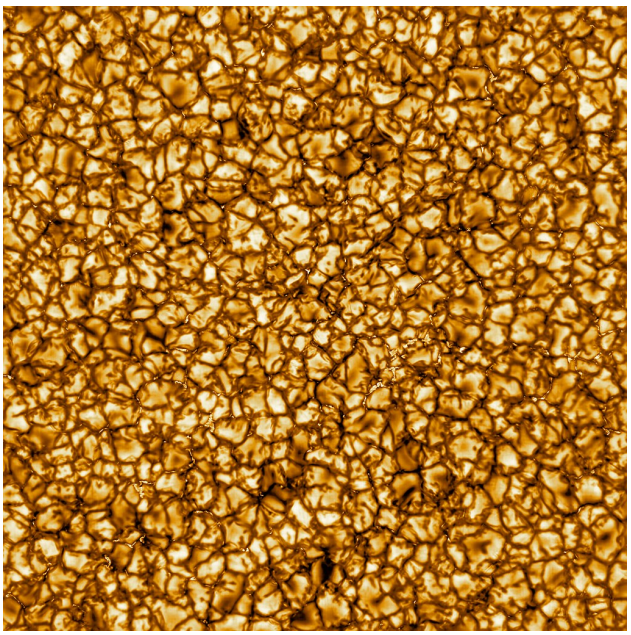


Figure 2.2: High-resolution image of the Sun's surface featuring solar granulation taken by the 4-meter Daniel K. Inouye Solar Telescope (DKIST) in Hawaii, USA.

through the intergranular lanes and are converted to vertical flows by the force of the different kinds of downflows that surround them [13].

These kind of vortex tubes play a fundamental role in solar magnetic flux dynamics and are a key factor in the generation of other kinds of solar vortex flows higher up in the solar atmosphere. More specifically, atmospheric vortex flows can be generated, should an intergranular vortex flow co-exist with the footprint of a magnetic field for long enough. As it was previously mentioned, in the upper layer of the solar convection zone the gas is greatly ionized and the value of plasma-β is high. That essentially means that the plasma dominates over the magnetic field and drags it along. As a result, intergranular vortex flows and magnetic lines often coincide locally. Exact co-location can result in the swirling of magnetic field structures that expand in the chromosphere and can even reach the upper levels of the solar atmosphere and the corona. On higher atmospheric heights, the value of plasma-β is low, since the plasma density decreases rapidly. There, the rotating magnetic field drags the plasma with it and chromospheric swirls are formed.

In Figure 2.3, a schematic representation of the aforementioned mechanism is illustrated. In this vortex system, particles move upwards and downwards. The magnetic field lines are "frozen-in" and plasma can propagate between the two vortex systems.

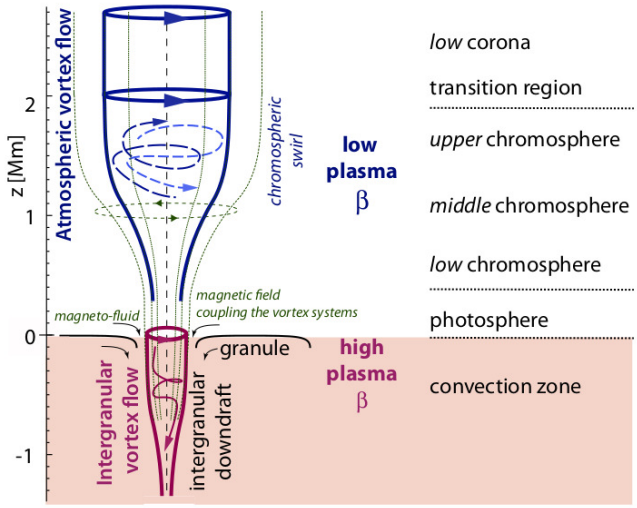


Figure 2.3: Schematic double-nature of vortex flows [11]

2.3 Observational detection of vortex flows

In order to study the different elements and structures on the Sun's surface and atmosphere, observations at different wavelengths take place on a wide range of the electromagnetic spectrum. A composite picture of the temperature ranges in the material around the sun is possible by using specialized organs like spectrometers on solar telescopes.

More specifically, the Sun contains a variety of atoms and ions, that emit light (photons) at specific wavelengths upon reaching a particular temperature. By observing the photons arriving at a solar instrument, one can identify the temperature on the region they originated from and in a sense the corresponding vertical layer of the solar structure.

Each element and its ions have unique patterns of spectral lines that characterize them [14]. In a uniform and continuous spectrum, a spectral line can be observed as a dark or bright line, depending on whether it has resulted from an absorption or an emission of light in that narrow frequency range.

The main spectral lines at which solar vortex flows have been observed are the following:

- **H-alpha**

  H-alpha (Ha) is a spectral line with a wavelength of 656.28 nm or 6562.8 Å [15], which corresponds to visible deep-red light. It is the first spectral line in the Balmer

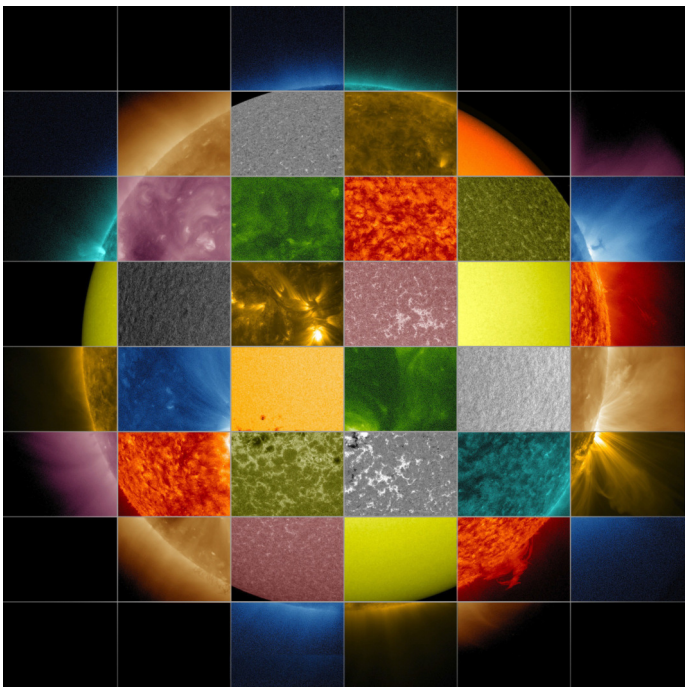


Figure 2.4: This collage of solar images from NASA's Solar Dynamics Observatory (SDO) shows how observations of the sun in different wavelengths help to highlight different aspects of the sun's surface and atmosphere. Credits: NASA/SDO/Goddard Space Flight Center.

series, which compromise a set of spectral lines describing the emissions of the atom of Hydrogen. When an electron transitions from a state with a principal quantum number of $n \geq 3$ to $n = 2$, which is the closest orbit to the nucleus, photons of different wavelength are emitted. H-alpha light occurs when an electron falls from the third ($n = 3$) to the second ($n = 2$) energy level and it is the brightest hydrogen line in the visible spectral range.

The H-alpha line is one of the most popular candidates for studying the solar chromosphere, since it is one of the most prominent spectral lines originating from that atmospheric layer [16]. The average formation height of the H-alpha spectral line is debatable and it is generally considered as a weak measure of the actual formation of the H-alpha line core. Among other estimations, it is considered to be between 1500 km to 2700 km in [17] while a more recent analysis places it between 750 km and 1500 km [16].

- **Ca**

The Ca spectral lines are ones of the most commonly used spectral lines for observation of the solar chromosphere. Among them, Ca K and H, located at 3933

Å and 3968 Å, are two of the deepest and broadest absorption lines of the solar spectrum. They were first observed by the German physicist Joseph von Fraunhofer (1787–1826) and provide great insight on the structure and properties of the chromosphere [18].

Additionally, the Ca II spectral line, located at 8542 Å and emitted from singly ionized Calcium, is another candidate for studying the chromosphere. It occurs when an electron falls from the forth ($n = 4$) to the third ($n = 3$) energy level in the ionized Calcium. Ca II 8542 Å constitutes one of the best observational tools for investigating the lower levels of the sun's atmosphere. Its average formation height is estimated to be around 1200 km from the sun's surface [19]. It is worth noting that the Ca II 8542 Å core is formed in a lower layer of the atmosphere than the H-alpha line core.

- **G-band**

  The G-band is another one of the spectral absorption lines observed and named by Joseph von Fraunhofer, while examining the solar spectrum. The G-band is a broad spectral line in the wavelength range 429.5 nm – 431.5 nm. It mainly occurs due to absorption by carbon-hydrogen (CH) molecules as well as by neutral iron (Fe) [20].

  This particular band is the main candidate for observing magnetic bright points, since their contrast appears higher in those wavelenghts [21]. More specifically, magnetic bright points are structures on the solar surface that are thought to trace kG magnetic concentrations [22] and are abundant in the integranular lanes. They appear as bright areas and are good tracers of magnetic elements [23]. At the photospheric level, vortex flows are primarily detected by studying magnetic bright points at the surface of the Sun with the G-band.

- **Mg II**

  The magnesium spectral lines at 2795.53 Å and 2802.71 Å are called Mg II k and Mg II h respectively. They occur when singly ionised magnesium experiences 3s-3p transitions from upper levels to the ground state [24]. Mg II k and h are placed in the near Ultraviolet (NUV) spectral range and demonstrate a complex line profile. They consist an excellent tool for observing multiple layers of the solar atmosphere

and their systematic observation has greatly contributed in unveiling structures in the transition region.

2.4 Observations of vortex flows

Solar vortex flows have been observed widely both in the Sun's photosphere and the chromosphere. In 1988, Brand et al. reported observations of a vortex structure with a 1.5 hour duration in the intergranular lanes of the photosphere, by examining bright points' motions [25]. The data were obtained by the Swedish 1-m Solar Observatory on La Palma (Canary Islands).

More recently (2008), Bonet et al. [26] identified vortexes of a smaller scale ($\leq 0.5Mm$) between the intergranular lanes in the photospheric layer that were reported to manifest a logarithmic motion behavior as they sink back in the convection zone. They estimated that the small scale vortex flows on the photosphere have a lifetime of around 5 minutes and an occurrence rate of 1.8×10^{-3} vortex $Mm^{-2}min^{-1}$. Images in the G-band (a 10.8 Å wide filter centered at 4305.6 Å) and in Ca H core and continuum obtained with the Swedish Solar Telescope (SST) [27] were used for this research. In [28], Bonet et al. enhanced their previous occurrence rate estimations by a factor of 1.7 using images obtained with the Imaging Magnetograph eXperiment (IMaX) mounted on the 1-m balloon-borne SUNRISE telescope[1] in 2010.

In 2009, Attie et al. [29] discovered larger scale vortex flows in the photosphere, utilizing blue continuum and G-band images from the Solar Optical Telescope (SOT) on-board the Hinode satellite[2]. They discovered wider vortex flows ($\sim$ 20 Mm) that were reported to have a lifetime of more than an hour.

In 2011, Vargas Domínguez et al. [30] analyzed the nature of vortex flows observed in G-band (centered at 4305.6 Å) images obtained with the Swedish 1-m Solar Telescope in 2007. They estimated that vortex flows exist in the photosphere with an occurrence rate of $\sim 1.5 \times 10^{-3}$ vortex $Mm^{-2}min^{-1}$.

Requerey et al. (2018) in [30], examined images at the Na I D1 589.6 nm line obtained with the Hinode in 2010 using the Narrowband Filter Imager (NFI) on board it. They

[1] `http://www2.mps.mpg.de/en/projekte/sunrise/`
[2] `https://www.nasa.gov/mission_pages/hinode/index.html`

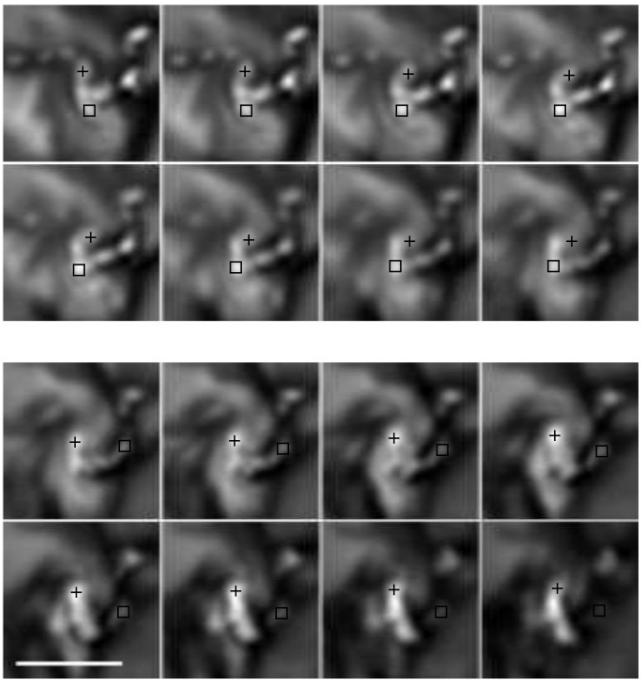


Figure 2.5: Sequence of snapshots corresponding to the event analyzed in detail in Bonet et al. [26]. The snapshots are separated by 15 s, and the bar on the lower left image corresponds to 1 Mm on the Sun. Two sets of two nearby BPs are marked, so that it is easy to visualize how the two BPs rotate with respect to each other. The bright structures in each snapshot depict a spiral pattern resembling the vortex derived from BP proper motions.

reported the detection of long lasting vortex flows ($>$ 24 hours) that exist at supergranular vertexes, with a diameter of around 5 Mm.

Vortex flows have also been widely observed at the chromosphere as well. In [31], Wedemeyer-Böhm and Rouppe van der Voort observed chromospheric flows on the Ca II 854.2 nm spectral line while analyzing images obtained with the CRisp Imaging Spectro-Polarimeter (CRISP) at the SST. They reported small scale formations (arcs, rings, spiral arms) with an average size of 1.5 Mm. In [6], it is estimated that 1.1×10^4 swirls exist in the chromosphere at any given time. This is based on observations obtained with the CRISP featuring 14 swirls with an average lifetime of $\sim$ 12 minutes. A close up of these Ca II 854.2 nm observations is presented in Figure 2.6.

In [32], Park et al. reported simultaneous observations of a small scale vortex flow in the spectral lines of H-alpha (for the first time), Ca II 8542 Å, and Mg II k. The observations were obtained on June 7, 2014 during a coordination between the ground-based CRISP/SST and the space-based Interface Region Imaging Spectrograph (IRIS) satellite

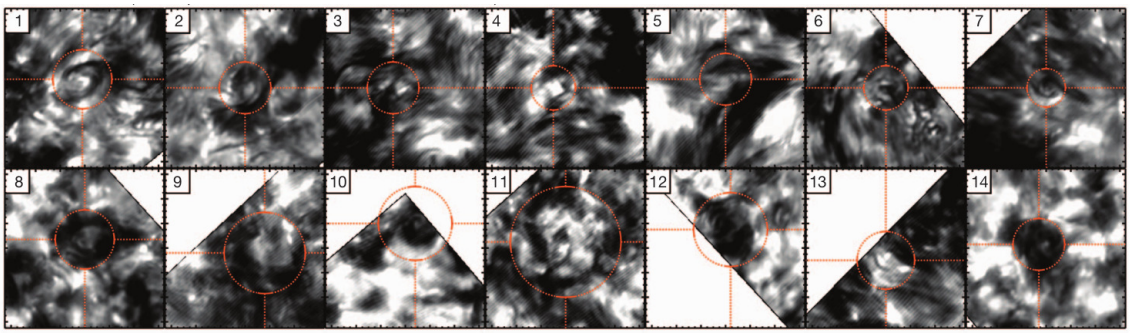


Figure 2.6: Close-ups (field of view $13'' \times 13''$) of the Ca II 854.2 nm line core maps for each of the 14 swirls described in [6].

that permitted the examination of a vortex flow spanning from the chromosphere to upper atmospheric levels. In the small scale vortex observed, two consecutive upflow events were also recorded during a ~ 5 minute period. Further H-alpha observations of a vortex flow, used during this study (see Section 2.6 for details), were more recently obtained by Tziotziou et al. [7] with CRISP/STT.

2.5 Simulations of vortex flows

The physics describing the plasma dynamics in vortex flows and its response/signature in photospheric and chromospheric lines is very complicated. A number of simulations attempting to adequately describe these complex structures of the Sun's surface and atmosphere have been developed over the years.

For instance, Kitiashvili et al. [13] utilized the 3D radiative MHD "SolarBox" code developed by Alan Wray et al. [33] at the NASA/Ames Research Center and the Stanford Center in order to study vortex tubes formation. Moll et al. [34] simulated the properties of MHD waves produced by solar surface convection, by utilizing the 3D magnetohydrodynamics simulation code (named MURaM code) that was implemented by Vögler et al. [35], in an attempt to facilitate the study of the solar photosphere and the convection zone. Wedemeyer et al. [6] analyzed the properties of vortex structures and their formation by using the CO^{5}BOLD simulation code implemented by Freytag et al. [36].

The aforementioned simulations, as other numerous numerical simulations not explicitly mentioned here, produce very realistic results that are in good agreement with observations as they take into account detailed physics governing the subsurface, solar surface and atmospheric layers, where vortex structures are generated and reside, in great depth.

As time progresses, their efficiency improves even further and nowadays they pose an invaluable tool for studying the dynamics of solar vortices.

2.6 Characteristics of the Studied Vortex flow

In the current thesis, a specific solar vortex flow is studied and used as input in the detection methods that will be discussed in the following sections. This particular vortex flow and its characteristics have been extensively described in [7].

The vortex flow was detected by utilizing a set of multi-wavelength data obtained with the CRisp Imaging SpectroPolarimeter (CRISP) instrument at the Swedish 1-m Solar Telescope (SST). A region of the quiet Sun in the south-west hemisphere was observed on June 7, 2014, between 07:32 UT–08:21 UT and 08:28 UT–09:16 UT. The retrieved data include spectral images in the H-alpha (6562.81 Å) and the Ca II 8542 Å line profiles, both sampled at seven different wavelengths. The location of the region of interest (ROI) and the observations in the aforementioned wavelengths at a selected point in time are presented in Figure 2.7.

The studied vortex flow has an almost circular shape and manifests a clockwise rotation. It has a variable centre around the position $(124.8'', -592'')$ of the solar plane. The particular vortex flow demonstrates some unique characteristics compared to previous studies that are analyzed in detail in [7]. The chromospheric flow has a radius of $\sim 3''$, which is at the large end in comparison to earlier observations, that in their vast majority include vortexes of a radius of $\sim 0.5'' - 2''$, as it was discussed above.

Moreover, the vortex flow presents a remarkably long duration of $\sim$ 100 minutes, which constitutes the longest duration ever reported for a small-scale solar vortex. As it was previously mentioned, typical solar vortex flows last around 15 minutes. In addition to its long duration, the studied vortex flow is also clearly visible for almost the entire $\sim$ 1.7 hour duration of the observations, rendering it very persistent.

Lastly, it is worth noting that the vortex flow presents an interesting substructure, as it can be revealed by examining more closely the time series of the spectral data. Inside the large vortex flow or in high proximity, four smaller swirls can be observed. Those swirls are not as persistent as the large vortex flow but tend to reappear approximately around the

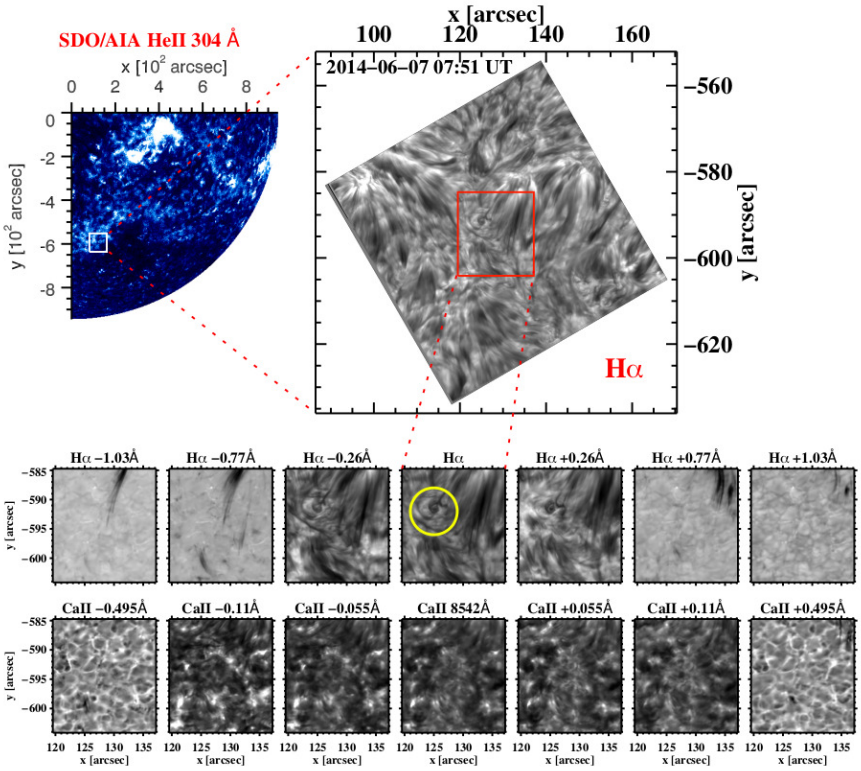


Figure 2.7: Top left: SDO/AIA He II 304 Å image with an overplotted rectangle marking the full FOV of the CRISP observations. Top right: co-temporal full FOV CRISP H-alpha line centre sample image with an overplotted rectangle marking the ROI. The bottom two rows show the ROI in the seven different wavelengths along the H-alpha and the Ca II 8542 Å profiles, while the overplotted yellow circle in the line-centre H-alpha panel indicates the location of the analysed vortex flow [7].

same locations. They have a shorter duration and a radius that is more consistent with the chromospheric swirl characteristics that have been reported by other observations.

For the purpose of this thesis, the time series of the described vortex flow in the center of the H-alpha spectral line will be utilized and further processed, in an attempt to develop an automated method for detecting solar vortex flows.

Chapter 3

State of the Art

This chapter provides a short literature review on the automated techniques for identifying solar vortex flows that have been developed up to this date, showing the state of the art in the solar vortex automated detection.

A variety of methods have been proposed so far, aimed to enable the easier and more efficient detection of solar vortex structures. Prior to the development of such techniques, vortex flows had been identified via visual inspection, a time-consuming, biased approach that fails to detect the majority of solar vortex flows present in the data. All methods are based on the development of a criterion to determine the existence of one or more vortex flows in a data set. The most common one is the "vorticity" ω, which is formally defined as the curl of the velocity field v:

$$\omega = \nabla \times v \tag{3.1}$$

Vorticity is a good indicator of the existence of vortex flows, as it is highly correlated with the flow's circulation. However, it has been shown that vorticity is not an ideal criterion for the detection of solar vortex flows, as it also obtains high values in the presence of shear flows that are often present in the Sun's atmosphere. Consequently, its usage can result in false detection of solar vortex flows [37].

In order to overcome this issue, a more sophisticated criterion that is not affected by the presence of shear flows has been used in more recent studies for the detection of vortex flows. The "vorticity strength" or "swirling strength", introduced by Zhou et al. [38], is the imaginary part resulting from the eigenanalysis of the velocity gradient tensor. The

swirling strength does not detect shear flows and has been succesfully used in a number of studies [34], [1], [3] etc. More specifically, Moll et al. in [34] identified vortices in simulations by using the swirling strength λ_{ci}. They detected vortices in regions where the eigenanalysis of the velocity field produced a pair of complex conjugate eigenvalues. High values of the imaginary part of those eigenvalues (λ_{ci}) indicated strong vorticity and were therefore identified as regions that contained a vortex. Moreover, they defined a criterion for categorizing the detected vortices to diverging or converging flows that presented either inward or outward motion based on the values of the eigenvalues.

Kato & Wedemeyer in [1] described and compared two automatic detection methods that identified solar vortex flows by using as the detection criteria an enhanced version of vorticity and the vorticity strength. Their approach was based on the determination of the velocity field by applying local correlation tracking (LCT) [39] on solar simulations. Using the derived velocity field, they extracted the vorticity, the vorticity strength and the streamlines present due to flows in the simulation data. The enhanced vorticity was then calculated as the convolution of vorticity with the extracted streamlines. Furthermore, a visualization technique known as Line Integral Convolution (LIC) was applied on the enhanced vorticity maps and the vorticity strength maps produced. LIC is a technique proposed by Cabral & Leedom in [40] that enables the visualization of turbulent flows and thus facilitates visual as well as automatic detection of vortices. Kato & Wedemeyer concluded that the vorticity strength method produced better results than the enhanced vorticity method in most aspects of their study. The vorticity strength method detected vortices with a higher success rate and provided more accurate results in terms of vortex localization as well as structure prediction. This is also apparent in Figure 3.1 produced during the application of their methods on a simulated isolated swirling motion. However, it is worth noting that the enhanced vorticity method was able to detect a number of rotating structures with low vorticity, discarded with the vorticity strength method, that might play an important role in certain case studies.

In [2], Giagkiozis et al. presented an automated method for the identification of vortex flows that was heavily based on a methodology known as Γ detection introduced by Graftieaux et al. [41] and primarily used for the detection of vortices in turbulent motions. This method is based on the definition of two functions, Γ_1 and Γ_2, that can identify vor-

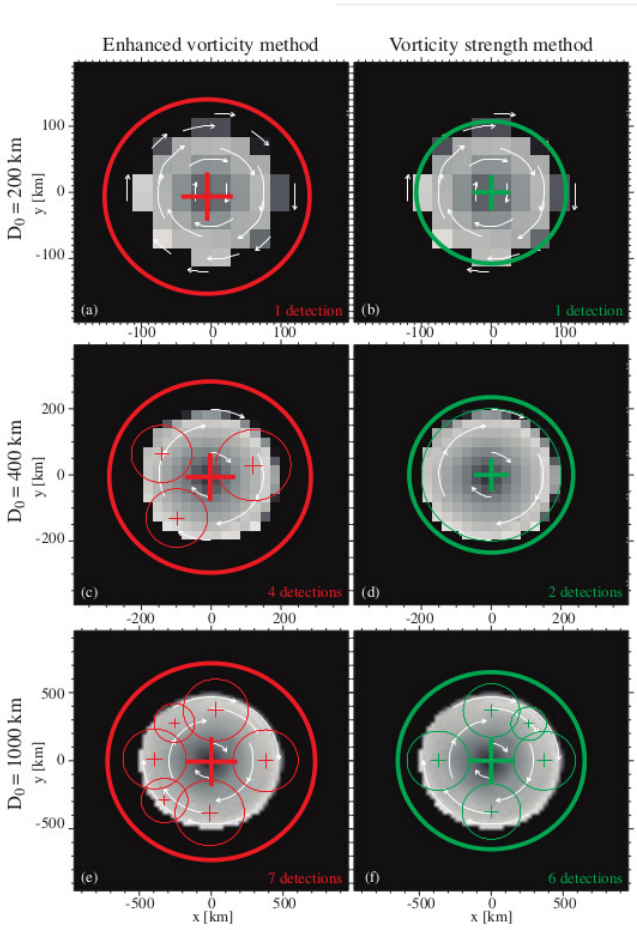


Figure 3.1: Examples of the detection procedure proposed by Kato et al. [1] for a vortex with a solid body rotation by using the enhanced vorticity method (left column) and the vorticity strength method (right column). The diameter of the rotating feature D_0 is 200 km for a) and b), 400 km for c) and d), and 1000 km for e) and f), respectively. Red and green crosses and circles indicate the location and the size of detected vortices, respectively. The vortex diameter estimated by the vorticity strength method is more accurate than the one produced with the enhanced vorticity method.

tices and estimate a number of their characteristics when applied on the velocity vector. In order to obtain the velocity field from actual photospheric intensity data, Giagkiozis et al. proceeded with a variety of preprocessing techniques, including local correlation tracking and image histogram equalization, in order to remove atmospheric effects that might be present in the observations. Using the Γ_1 and Γ_2 functions, the identification procedure produced results concerning the lifetime, diameter and location of the detected vortices. Giagkiozis et al. reported that the proposed method identified a greater number of vortices present in the data set than other previously-proposed methodologies. Nonetheless, they underlined that their findings were highly depended on the accuracy of the velocity field derivation, which is only an approximation of the solar dynamics and thus may not

reflect the actual solar environment.

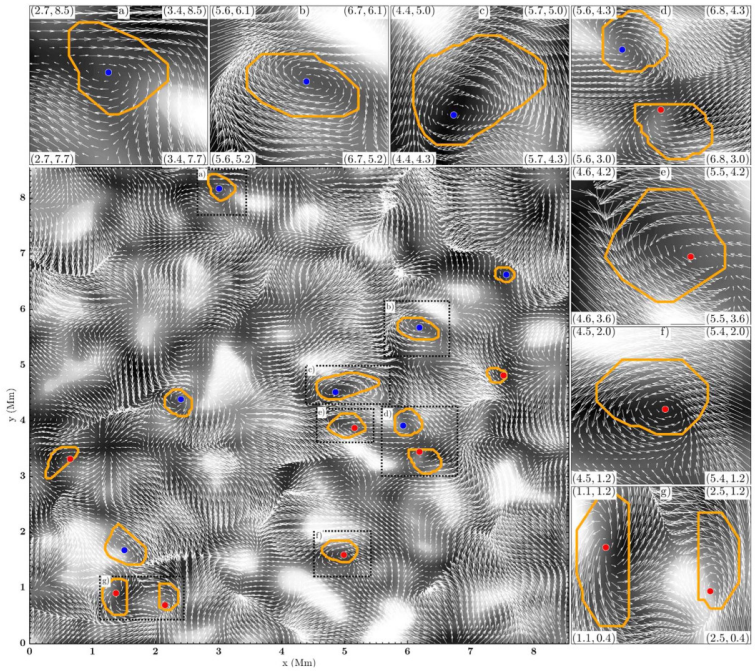


Figure 3.2: Giagkiozis et al. [2] estimation of velocity field based on the Fe I continuum (intensity shown in gray scale) using LCT. The identified vortices and their boundaries produced by the proposed LAVD detection method are also illustrated. The circles denote the vortex center, with the red circles referring to the counterclockwise vortices (positive) and the blue circles to the clockwise vortices (negative).

In [42], Silva et al. proposed an automatic detection method for solar vortex flows based on a technique developed by Haller et al. [43] known as Lagrangian Averaged Vorticity Deviation (LAVD). This method seeks to identify the boundaries of detected vortices in an objective manner. Silva et al. utilized LAVD in order to develop a method for vortex identification that takes into account the variance that the velocity field presents through time. The LAVD method employs the averaged vorticity deviation as the vorticity criterion that has as follows:

$$LAVD_{t_0}^{t_0+\tau}(x_0) = \int_{t_0}^{t_0+\tau} |\omega(x(t),t) - \langle\omega(t)\rangle| dt \tag{3.2}$$

where ω is the vorticity, $x(t)$ denotes the particle position function extacted from the velocity field and $\langle\cdot\rangle$ is the instantaneous spatial mean. Vortices are subsequently detected in the local maxima of the LAVD function. Moreover, based on the streamlines geometry of the displacement vector, a new parameter d was introduced in order to account for falsely identified vorticies when combined with the LAVD methodology. Silva et al. com-

pared the results of the proposed method with the results produced by two alternative methodologies based on the "vorticity strength" criterion and Γ detection respectively. They reported that the LAVD method was superior to the other two mainly due to its ability to differentiate persistent vortex flows from other non-lasting swirling structures. Their results, comparing the three methodologies, are shown in Figure 3.3. It is also worth noting that in a number of cases the vortices detected by LAVD were different in terms of amount and shape from the ones extracted by the other methods. This was attributed to the fact that LAVD mainly focuses on the dominant swirling structure present in every neighborhood.

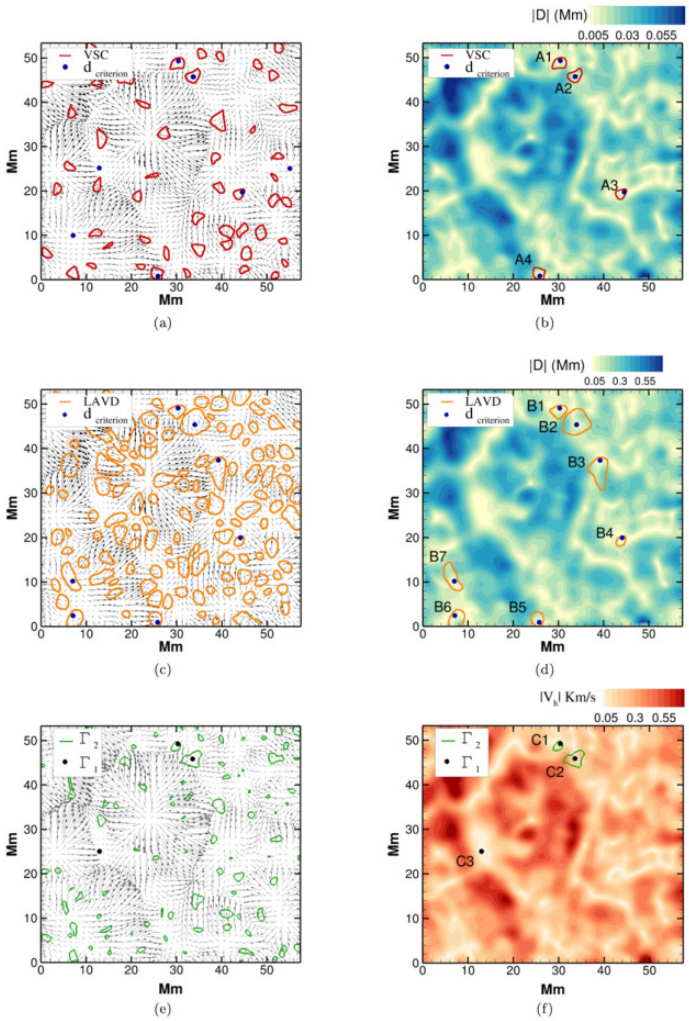


Figure 3.3: Vortex detection by the normal and corrected Eulerian and Lagrangian methods by Silva et al. [42]. Vortex detection by the vorticity strength criterion (VSC) and the corrected VSC+d method are shown in figures (a) and (b). Vortex detection by the normal LAVD method and the corrected LAVD+d method are shown in (c) and (d). Vortex detection by Γ-2 and by the Γ method after selection by the Γ-1 criterion are presented in (e) and (f) respectively. The LAVD+d method identifies more vortices than the other two methods.

More recently, Cuissa et al. proposed a novel dynamical equation for describing the swirling strength in [3]. The analytical evolution equation of the vorticity strength was derived from modifications to the momentum equation of the system. It was subsequently applied on numerical simulations of magnetohydrodynamic fields and it was reported to be more reliable in vortex identification than previous methods. Cuissa et al. also compared two methodologies based on the evolution equations of the vorticity and the vorticity strength criteria respectively. They concluded that the vorticity strength method proved more efficient in vortex detection, especially in simulations of the solar convection zone where shear flows are present that bias the results produced by the vorticity.

The study of Cuissa et al. focused on identifying vortices present in simulation data. Nonetheless, they reported that their method could also be applied on data derived from observations, by extracting the velocity field in the three-dimensions space using techniques such as local correlation tracking (LCT). However, they raised concerns that such an approach could lack in accuracy since multiple spectral data are required for the three-dimensional velocity field extraction and consequently the noise present in the data would highly increase. Nonetheless, they concluded that the proposed approach would still produce significant results even when applied on a two-dimensional velocity field in an attempt to reduce the noise in the data set.

Conclusively, it is clear that all the automated detection methods of solar vortex flows described in this chapter share a common ground. They all define a criterion, such as the vorticity or the swirling strength, that is based on the derivation of the velocity field present due to solar motions. Consequently, these detection methods require prior knowledge of the rotational velocity, which is a highly-challenging quantity to extract from solar observations.

It is therefore apparent that, while these velocity-dependent methods perform well on simulated data, their efficiency highly depends on the accuracy of the velocity vector estimation when applied on data sets of actual observations. As a result, they might produce inconsistent or unrelated results to the actual solar environment. To this end, in the current thesis a morphological method that does not require the knowledge of the velocity field is proposed. By eliminating the dependency of the velocity field derivation from the procedure, the development of a robust methodology that performs equally well on data

derived from synthetic profile simulations and actual observations is attempted.

Chapter 4

Proposed Detection Methods

A number of diverse processing methods have been utilized as ingredients in an attempt to create an automated procedure for detecting solar vortex flows, using as an input spectral data depicting the solar surface or chromosphere. The first step towards this aim is to remove redundant information from the data set that does not contribute to the detection of vorticity. To this end, Robust Principal Component Analysis (RPCA) is applied on the frames data set, in an attempt to extract information related to swirling motions from the dense and complex original data. RPCA is a statistical procedure that has been widely used as a tool to extract the foreground from the background in video applications, as it will be analyzed later on. In the current context, RPCA is applied on the data set in order to differentiate the slow-moving parts of the frame sequence from the faster-moving.

In addition to RPCA, selected edge detection filters applied on digital images are also utilized in an attempt to reduce the volume of information each frame contains. Such filters include the Canny edge detector and the Sobel filter, both of which retrieve the edges present in an image. It is anticipated that by extracting the edges in each frame, the vorticity-related information will be preserved while the complexity will decrease rapidly.

Moreover, in order to further eliminate structures that are unrelated to vortex flows, the "almost linear" components of each frame, which are the result of other motions manifesting on the solar photosphere and chromosphere, are identified and removed from the data set. This is achieved by first identifying candidate "almost linear" components and then

removing those of them that present high linearity, according to a linear regression-based criterion.

Lastly, apart from the major preprocessing methods discussed above, additional techniques are also utilized and combined in order to develop a step by step methodology for the automatic detection of solar vortex flows. All the techniques utilized for this purpose are described in detail in the following sections.

4.1 Robust Principal Component Analysis (RPCA)

In the last few decades, the volume of scientific information available for research has risen dramatically. Consequently, the need for efficient knowledge extraction out of very high volume data sets (big data problem) is nowadays more crucial than ever. A wide range of practices has been designed and implemented for this purpose, constituting an essential tool for processing a data set prior to further analysis. Amongst them, sparse signal recovery and low rank matrix completion, which seek for multiple subspace structures in the data, have been a primary candidate for knowledge extraction in an number of applications.

One of the most widely-used statistical procedures, the Principal Component Analysis (PCA) attempts to retrieve the closest low rank representation to a given data set. However, PCA can be quite brittle upon dealing with grossly corrupted observations. A more robust alternative called Robust PCA (RPCA), has been proposed in recent years, that performs ideally upon dealing with this kind of data. Specifically, RPCA aims to recover a low-rank component and a sparse component from the data set, with the sparse component corresponding to the corrupted measurements.

Robust PCA has been widely used in a broad range of image and video processing applications such as video surveillance, face recognition and image denoising. The main principle upon which it operates is the recovery of a low-rank matrix L and a sparse one S from a data matrix X, so that $X = L + S$. The RPCA problem is stated more formally as follows 4.1:

$$\min_{L,S} \quad rank(L) + \lambda \|S\|_0 \qquad subject\ to \quad X = L + S \tag{4.1}$$

where rank(L) denotes the rank of the matrix L, $\|S\|_0$ is the l_0 "norm" [1] of the sparse component, λ is a regularization parameter and $\{X, L, S\} \in R^{m \times n}$.

However, this is a non-convex, NP-hard problem [44]. One way to deal with this issue, is to consider a relaxed convex version of the above problem [45]. The convex relaxation can be achieved by relaxing the rank constraint $rank(L)$ in the optimization problem to the nuclear norm $\|L\|_*$ and the sparsity constraint $\|S\|_0$ to the ℓ_1 norm $\|S\|_1$. This relaxation problem is mostly known as Principal Component Pursuit (PCP) and it is described in equation 4.2:

$$\min_{L,S} \quad \|L\|_* + \lambda \|S\|_1 \qquad \textit{subject to} \quad X = L + S \tag{4.2}$$

where the nuclear norm $\|L\|_*$ is equal to the sum of the singular values $\sigma_i(L)$ of the matrix, that is, $\|L\|_* = \sum_i \sigma_i(L)$, $\|S\|_1$ is the ℓ_1 norm [2] of the sparse component and $\{X, L, S\} \in R^{m \times n}$. The parameter λ is a regularization parameter and choosing $\lambda = \frac{1}{\sqrt{max(m,n)}}$ ([45]) has been proven to work well for most applications using PCP. For the purposes of this work, this value will be used for λ, as well.

In order to solve the minimization problem described in 4.2, an optimization technique called Alternating Direction Method of Multipliers (ADMM) [46] can be used. ADMM is an algorithm that solves convex optimization problems efficiently by combining previously established optimization techniques like dual decomposition [47] and the Method of Multipliers [48]. The main principle upon which it operates is the optimization of a dual problem featuring the augmented Lagrangian of the original problem. ADMM works in an iterative manner, by solving a minimization problem per variable (or set of variables) in an alternating fashion.

ADMM is one of the main candidates for handling PCP and the procedure for solving the optimization problem in 4.2 is described in detail in [49]. The augmented Lagrangian for the PCP problem presented in equation 4.2, with a penalty parameter $\rho > 0$, is the

[1] The l_0 "norm" of a matrix is defined as the total number of its nonzero elements.
[2] The ℓ_1 norm of a matrix is defined as the sum of the absolute values of its columns.

following:

$$\begin{aligned}\mathcal{L}_\rho(L,S,Y) &= \|L\|_* + \lambda\|S\|_1 + \mathbf{Tr}(Y^T(L+S-X)) + \frac{\rho}{2}\|L+S-X\|_F^2 \\ &= \|L\|_* + \lambda\|S\|_1 + \frac{\rho}{2}\|L+S-X+W\|_F^2 + const\end{aligned} \tag{4.3}$$

where $Y \in R^{m\times n}$ is the dual variable and $W \in R^{m\times n}$ is defined as $W = \rho^{-1}Y$ for simplicity. The value of the penalty parameter ρ is more commonly selected as shown in equation 4.4 [50]. This value will also be used in the current thesis.

$$\rho = \frac{1}{4\overline{|X|}}, \qquad where \quad \overline{|X|} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}}{mn} \tag{4.4}$$

Initializing S and W, the updatings in each iteration are described by the following equations:

$$\begin{aligned}L^{(k+1)} &:= \underset{L}{argmin}\ \mathcal{L}_\rho(L, S^{(k)}, Y^{(k)}) \\ &:= \underset{L}{argmin}(\|L\|_* + \frac{\rho}{2}\left\|L + S^{(k)} - X + W^{(k)}\right\|_F^2)\end{aligned} \tag{4.5}$$

$$\begin{aligned}S^{(k+1)} &:= \underset{S}{argmin}\ \mathcal{L}_\rho(L^{(k+1)}, S, Y^{(k)}) \\ &:= \underset{S}{argmin}(\lambda\|S\|_1 + \frac{\rho}{2}\left\|L^{(k+1)} + S - X + W^{(k)}\right\|_F^2)\end{aligned} \tag{4.6}$$

$$W^{(k+1)} := W^{(k)} + (L^{(k+1)} + S^{(k+1)} - X) \tag{4.7}$$

In order to further reduce the complexity of the problem at hand, proximal operators [51] will be utilized as they pose an excellent tool for dealing with optimization problems. More specifically, equation 4.5 can be interpreted as a proximal operator [3]

of the nuclear norm of L, evaluated at the value $X - S^{(k)} - W^{(k)}$. Accordingly, equation 4.6 can be alternatively viewed as the proximal operator of the ℓ_1 norm of S evaluated at $X - L^{(k+1)} - W^{(k)}$.

[3]The proximal operator of a function f from a Hilbert space $\mathcal{X}$ to $(-\infty, +\infty]$ is defined as:

$$prox_f(u) = \underset{u\in\mathcal{X}}{argmin}\{f(u) + \frac{1}{2}\|u-x\|^2\} \qquad for\ any\ x \in \mathcal{X}. \tag{4.8}$$

These proximal operators (which will offer the updating of L and S in each iteration) can be calculated using Singular Value Decomposition (SVD) and a soft thresholding operator. SVD is a technique that factorizes a given $m \times n$ matrix M in a form $M = U\Sigma V^*$, where:

- U is an $m \times m$ (real or complex) unitary matrix
- V is an $n \times n$ (real or complex) unitary matrix
- Σ is an $m \times n$ diagonal matrix with non-negative diagonal entries

and the diagonal values $\sigma_i \equiv \Sigma_{ii}$ of Σ are called the *singular values* of the matrix M.

The soft thresholding operator ST_κ, where $\kappa > 0$, is defined as follows:

$$ST_\kappa(x) = \begin{cases} x - \kappa & \text{if } x > \kappa \\ 0 & \text{if } |x| \leq \kappa \\ x + \kappa & \text{if } x < -\kappa \end{cases} \tag{4.9}$$

In the light of 4.7 and 4.9, it can be proved that equations 4.5 and 4.6 can be rewritten using their proximal operator's interpretations:

$$L^{(k+1)} \quad := \quad U\ \mathbf{diag}(ST_{1/\rho}(\sigma_1),\ \ldots,\ ST_{1/\rho}(\sigma_n))\ V^T \tag{4.10}$$

where $\sigma_1,\ \ldots,\ \sigma_n$ are the singular values of $X - S^{(k)} - W^{(k)} = U\Sigma V^*$ and

$$S_{ij}^{(k+1)} \quad := \quad ST_{\lambda/\rho}(X_{ij} - L_{ij}^{(k+1)} - W_{ij}^{(k)}) \tag{4.11}$$

where S is computed element-wise for every $S_{ij}, \quad i = 1,\ \ldots,\ m \quad and \quad j = 1,\ \ldots,\ n.$

Overall, the algorithm that has been implemented for Robust PCA in the scope of the current thesis can be found in Algorithm 1.

RPCA is most commonly utilized for the detection of movement in videos, where the retrieved low-rank component corresponds to the static background of the video and the sparse component to the detected motion. In order to provide an appropriate input for the RPCA analysis, the transformation of each video frame to a vector is required so that a 2-D matrix with the frame vectors as columns is calculated. A similar procedure will also

Algorithm 1 ADMM for PCP

Input: $X,\ \rho > 0,\ \lambda > 0$

1: **Initialization** $k = 0, \quad S = S^{(0)}, \quad W = W^{(0)}$
2: **repeat**
3: Perform SVD on $X - S^{(k)} - W^{(k)} = U\Sigma V^*$
4: $L^{(k+1)} := U\ \mathbf{diag}(ST_{1/\rho}(\sigma_1),\ \ldots,\ ST_{1/\rho}(\sigma_n))\ V^T$
5: $S_{ij}^{(k+1)} := ST_{\lambda/\rho}(X_{ij} - L_{ij}^{(k+1)} - W_{ij}^{(k)})$
6: $W^{(k+1)} := W^{(k)} + (L^{(k+1)} + S^{(k+1)} - X))$
7: $k = k + 1$
8: **until** convergence

Output: $\hat{L} = L_{k+1}, \hat{S} = S_{k+1}$

by followed for the computation of a suitable input matrix for the application of RPCA on the selected solar data, as it will be discussed in the results section.

4.2 Filtering

Among the various filtering operations applied on a digital image, the one that identifies the edge pixels in it is of increased interest. Two of the most widely used filters of the this kind are the Canny edge detector and the Sobel filter, which will be briefly commented in the sequel.

4.2.1 Canny Edge Detector

The Canny Edge Detector, or Canny filter for simplicity, is a technique widely used in computer vision applications for the detection of edge pixels in images. The Canny filter was developed in 1986 by John F. Canny and first presented in [52].

Like other proposed edge detection algorithms, the Canny Edge Detector is an important tool for image processing applications that focus on object morphology and the detection of boundaries between components in an image. By retrieving the edge pixels and discarding the rest [4], the volume of the data is reduced and as a result processing is further accelerated, while the majority of the crucial structural information required for that kind of applications is preserved.

The Canny Edge Detector achieves the detection of edge pixels in an image with a multiple stage procedure that can be broken down to the following steps:

[4] Of course, this refers to applications where edges are of special interest.

- Noise reduction

  The first step required in the Canny Edge Detector procedure is the removal of noise that might be present in the image. In order to reduce the degrading noise impacting the data, a Gaussian filter is utilized for blurring the original image. The Gaussian filter is defined by a kernel (square window) of a size $(2k+1) \times (2k+1)$. The selection of the kernel size is an important factor as it can highly affect the performance of the edge detector. Selecting a large size value can reduce the edge detection accuracy in terms of localization, while small kernel sizes present lower efficiency in removing noise. A size of 5×5 (that is k=2) is reported to work well in most case scenarios. It should also be noted that in the process of smoothing the image in order to remove noise, the Gaussian filter also smooths the edges existing, which can result in oversight of weak edges. However, noise removal is a step that, in general, greatly improves the efficiency of the edge detection and certain adaptive Gaussian filters that have been proposed can minimize the aforementioned oversight of weak edges due to smoothing (see [53]).

- Gradient calculation

  After the blurring has been applied to the original image, the Canny edge detector proceeds with identifying edge pixels in four different orientations (horizontal, vertical and the two diagonals). In order to do so, the image is processed using an edge detection operator (eg. Sobel, Roberts etc.) for the extraction of the first derivative in the vertical (G_y) and the horizontal (G_x) direction at each pixel (more details are given in section 4.2.2 below). Using these values, the gradient of the image pixels as well as their direction can be calculated as follows:

$$G = \sqrt{{G_x}^2 + {G_y}^2} \tag{4.12}$$

$$\Theta = \arctan(G_y, G_x) \tag{4.13}$$

  The gradient direction, which is always perpendicular to the edge that it corresponds to in the image, is rounded to the closest one of the four angles (vertical, horizontal

and the two diagonals).

- Non-maximum suppression

  On the gradient intensity matrix produced from the previous step, an edge thinning technique called non-maximum suppression is applied. The main principle upon which this method operates is the identification of the local maximum in a region (in terms of intensity) and the suppression of the rest of the pixel values in that neighborhood and gradient direction. The resulting image after this procedure contains only thin edges that indicate the strongest intensity changing points in the data. Those edges represent a close approximation to the actual edges of the original image. However, there are still exist edges produced due to noise or variation of color in the image. Two more steps are required to identify the actual edges present in the data.

- Double threshold

  In order to filter out identified edges that do not correspond to actual edges of the original image, a double threshold approach is applied on the processed data. That is, a high and the low threshold are defined and pixels are categorized as follows regarding their gradient intensity value:

  - *Strong edge pixels*: pixels with a gradient intensity value over the high threshold
  - *Weak edge pixels*: pixels with a gradient intensity value between the low threshold and the high threshold
  - *Irrelevant pixels*: pixels with a gradient intensity value lower than the low threshold

  The strong edge pixels are instantly classified as edge pixels in the final output. The irrelevant pixels are suppressed and no longer considered for this procedure. Weak edge pixels are re-evaluated at the final step, in order to decide whether they will be identified as edge pixels or not.

  An important aspect of the double threshold step is the selection of the threshold values. At some applications, empirical determination of the threshold values takes place, that highly depends on the input image. In this thesis, the implementation of

the Canny filter uses a technique known as **Otsu's method** for the determination of the high threshold. Otsu's method is one of the most popular techniques for image thresholding in an automated manner. It was proposed by Nobuyuki Otsu and thoroughly explained in [54]. Having defined the high threshold, the low threshold is set equal to the half of the high threshold.

- Edge Tracking by Hysteresis

  In the last step of the Canny Edge Detector procedure, the weak edge pixels are evaluated to determine whether they should be preserved as edge pixels in the output. The criterion used for this decision is whether the weak edge pixels have at least one strong edge pixel in their neighborhood. More specifically, if there exists at least one strong edge pixel among the eight pixels surrounding the evaluated one, then the latter is preserved in the final output as an edge pixel. This is a good indicator since noise pixels, that are most probably unconnected and random, are suppressed by this criterion. The output of this step, which is the also the final output of the algorithm, is an image including the detected edge pixels.

4.2.2 Sobel Operator

The Sobel operator, or more appropriately the Sobel–Feldman operator, is an operator used in computer vision applications mainly for the extraction of edges in images. It was first presented by Irwin Sobel and Gary Feldman at a talk at Stanford Artificial Intelligence Laboratory in 1968. More formally, it has been described in detail in [55].

The Sobel operator produces an isotropic gradient estimate of the image. The gradient intensity at each pixel is approximated in a 3x3 neighborhood by calculating the convolution of the 3x3 part of the image centered at the pixel under consideration, with the 3x3 kernels defined for this purpose. At each point, the vector sum of the eight directional derivative vectors is estimated. The two following kernels in the horizontal and vertical direction respectively have been defined and their usage in the Sobel filter has been demonstrated

in [55]:

$$K_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \qquad K_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{4.14}$$

By defining the horizontal and vertical derivative approximations of the image as $G_x = K_x * A$ and $G_y = K_y * A$, where A is the original image, the gradient intensity at each pixel can be computed by the equations 4.12, 4.13 that were previously presented in the Canny filter section.

The operator is very computationally efficient since it is based on simple image convolution and thus it has been widely used in edge detection algorithms (e.g. the Canny edge detector mentioned in the previous section). It is worth noting, however, that the gradient intensity estimation that is generated is not always the sharpest, especially in images that present high-frequency variations [56].

4.2.3 Other filters

Apart from the above, two additional filters were applied on the data set, the Difference of Gaussians (DOG) filter and the Madmax filter.

The DOG filter is band-pass filter that is frequently utilized as a feature-enhancer. Its implementation involves the calculation of two blurred versions of the original image that are produced by convolving the image with distinct Gaussian kernels with a different standard deviation. The DOG filter output is then calculated by subtracting the one blurred image from the other, less blurred image. As a result, the information that is preserved in the output of the DOG filter lies between the two thresholds above which all frequencies are suppressed due to the convolution with each Gaussian kernel (low-pass filters). This procedure can also be visualized in Figure 4.1.

The Madmax filter is a multi-directional maximum gradient operator proposed by Koutchmy and Koutchmy [57]. Its implementation is based on modifying a second derivative spatial filter. Madmax has been mainly utilized for increasing the contrast of images depicting solar structures. Apart from enhancing the features present in an image, it also considerably

[5] https://micro.magnet.fsu.edu/primer/java/digitalimaging/processing/diffgaussians/index.html

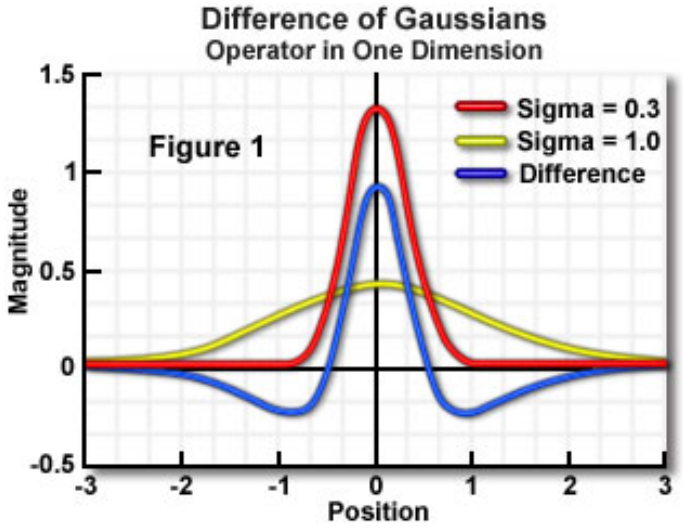


Figure 4.1: The plot of two Gaussian curves with different standard deviations and their difference. [5]

reduces the background noise [58].

4.3 Removal of "almost linear" structures

In order to further facilitate the detection of solar vortex flows from the processed data, a procedure aiming to remove the "almost linear" features present in each frame is proposed. These features represent in real observations mainly fibril-like structures (e.g. mottles/spicules) that are not completely removed by RPCA and filtering as they can also be quite dynamic. A quick overview of the steps comprising this procedure followed by a more in-depth description of each step is provided below.

Initially, the connected components of each frame (a definition will be given shortly) need to be determined. Subsequently, for each connected component it should be decided whether it is linear or non-linear. To this end, linear regression is applied on the set of the pixel coordinates that comprise each component in order to resolve its best-fit line. To this end, the euclidean distances of each pixel are calculated so that a criterion based on the standard deviation of these distances can be set that can adequately describe the linearity of each component. Finally, a threshold is established that is utilized to remove components that present a criterion value above it.

In greater detail, a criterion should be set for determining when two pixels will be treated as connected. In binary images, two pixels are considered connected when they are neighbors

sharing the same value. The neighborhood of a pixel is most commonly defined as all the surrounding pixels that are accessible from it by one or two orthogonal hops at most. The number of orthogonal hops required for two pixels to be perceived as connected is frequently referred to as the connectivity value and determining which value to use (1 hop or 2 hops) can greatly affect the produced results of the procedure. In the current thesis, the connectivity value for extracting the connected components of an image is selected on a per case basis as it will be discussed in the results section.

Subsequently, after all the connected components of the frame have been determined, each one should be evaluated as linear or non-linear so that it can be further decided whether it will be removed from the frame. In order to determine the linearity of each component, linear regression is utilized. Linear regression is a statistical tool for modeling the relationship between a dependent variable and one or more independent variables that utilizes a linear predictor function as the relationship model. More specifically, the model describing the relationship between the dependent and independent variables is defined as a linear combination of the independent variables and a set of unknown coefficients [59]. These coefficients are then estimated from the data.

There exist a variety of approaches for determining the coefficients of a linear regression model, depending on multiple factors such as prior knowledge on the distribution of the data. The most common approach that is bound to work well in most cases is called Ordinary least squares Linear Regression (OLR). OLR is based on the least squares principle, meaning that the model's unknown parameters are estimated by minimizing the sum of the squares of the differences between the dependent variable values as provided by the data and the ones predicted by the linear model.

If there exists only one independent variable (like in the current case), the linear regression procedure is simplified and thus called simple linear regression. In this case, the linear model can be written as follows:

$$y = \alpha + \beta x \tag{4.15}$$

where x is the independent variable, y is the dependent variable, α and β are the unknown model parameters. By minimizing the sum of squared differences between the actual

and the predicted values of the dependent variable y, the parameters α and β can be estimated. More formally, the estimates of α and β, $\hat{\alpha}$ and $\hat{\beta}$ are the solutions of the following minimization problem:

$$\min_{\alpha,\beta} \quad \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2 \tag{4.16}$$

By solving the above minimization problem, $\hat{\alpha}$ and $\hat{\beta}$ are calculated as follows:

$$\hat{\alpha} = \overline{y} - \hat{\beta}\,\overline{x} \tag{4.17}$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n} (x_i - \overline{x})^2} \tag{4.18}$$

where $\overline{x}$ and $\overline{y}$ are the average values of all x_i and y_i for $i = 1, \ldots, n$ respectively in the original data.

In order to apply the simple regression model procedure that is described above in the current setting the following approach is used: for every component of the frame under study, the coordinates of each pixel belonging to that component are utilized as the independent (x coordinate) and dependent variables (y coordinate) respectively. Subsequently, the linear model that best fits the component at hand is determined, which in the current two-dimensional problem constitutes a simple line.

Still a universal criterion remains to be defined for the characterization of each component as linear or non-linear. To this end, the euclidean distance d_i of each pixel from the best-fit line of the component it belongs to is calculated. More formally, for a pixel with coordinates (x_i, y_i) and a line described by the equation $Ax + By + C = 0$, this distance is calculated as follows:

$$d_i = \frac{|Ax_i + By_i + C|}{\sqrt{A^2 + B^2}} \tag{4.19}$$

For every connected component of the frame, the euclidean distances of its pixels from the best-fit line are calculated and then their standard deviation is computed. More specifi-

cally, for a component c consisting of N pixels that have distance d_i, for $i = 1, \ldots, N$ from the best-fit line, their standard deviation STD_c is defined as follows:

$$STD_c = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i - \overline{d})^2} \tag{4.20}$$

where $\overline{d}$ is the average of d_i, $i = 1, \ldots, N$.

It is obvious that low values of the standard deviation suggest high linearity of the component. In order to determine which components will be regarded as "almost linear" in every frame and consequently removed, a threshold for their linearity should be defined. All components presenting a standard deviation STD_c lower than that threshold value will be treated as "almost linear" and will be retained from the frame.

4.4 Evaluation mechanism

The primary aim of the current thesis is the design of a robust procedure/methodology that can detect regions where vortex flows exist when provided with a set of solar data (image cube). To this end, the final step of the implementation procedure is the creation of an evaluation mechanism that receives the processed data (that is, the data where most linear elements have been removed) as input and as an output it predicts which regions constitute the primary candidates for vortex motion.

More specifically, the set of the processed frames is provided as input to the evaluation mechanism. The first step towards evaluating the data is the aggregation of these frames in terms of time. To this end, the mean frame F_{mean} is calculated from the processed time series of frames. In this mean frame, the value of each pixel is computed as the average of the values of all the pixels that share the same spatial coordinates on each one of the original frames. In a sense, the mean frame can be perceived as the average of all frames as the time progresses, thus holding aggregated information of the entire time series.

After the mean frame F_{mean} has been determined, the regions where F_{mean} presents the higher activity should be identified. These regions will then be considered the primary candidates for the existence of vortex flows. This can be supported by the fact that the preceding processing stages removed (hopefully) most of the information not related to

vortex motion. As a result, regions with high valued pixels on the mean frame are anticipated to depict dynamic, non-linear components that most likely correspond to swirling motions.

In order to identify such regions, a rolling window methodology is applied on F_{mean}. More specifically, a window of certain shape and size is selected. This window is shifted by a step of one pixel (either horizontally or vertically) until the whole F_{mean} image is traversed. On each sub-image identified by the rolling window, an objective function is calculated. For the purposes of this evaluation procedure, the objective function selected is the summation of the pixel values contained in the aforementioned window. The higher the value of the objective function, the more probable it is that the region depicted contains swirling dynamic structures, since most relevant pixels were preserved by the pre-processing procedures. The output of this procedure is a new image, where the pixels associated with high objective function values (visualized as intensity), are likely to denote possible vortex locations.

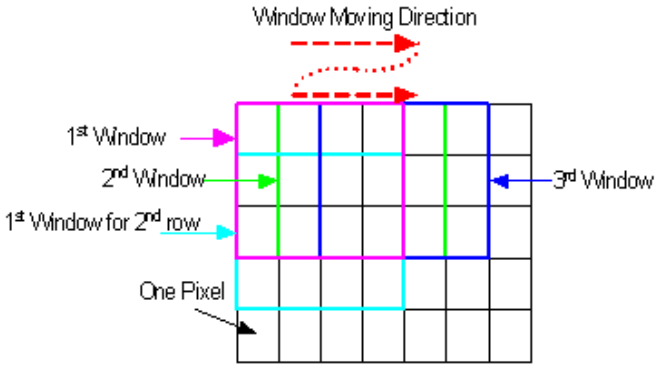


Figure 4.2: Rolling window on a 2D image. Credits: Miriam Leeser and Haiqian Yu. [6]

After the objective function has been estimated on all possible windows of the F_{mean} image, a visualization of the probability function concerning the possibility of vortex flow existence in every region is feasible. Each window value calculated by the objective function can be perceived as a pixel in an image consisting of as many pixels as the windows that were produced by traversing the F_{mean} image during the rolling window procedure. This image will be hereafter referred to as the Probability Distribution Image.

By identifying the maximum value on the Probability Distribution Image, one can correlate the coordinates of this value to the coordinates of the corresponding window in the mean

[6] https://sites.google.com/site/hsi2013logan99/computer-vision/implementation

frame. Thus, the primary candidate region for the existence of a vortex flow can be detected. Additionally, more than one candidate regions, that can even present different sizes, can be identified by utilizing the Probability Distribution Image. However, this procedure is not investigated in the present thesis and will be further analyzed in the "Future Work" section.

Two different approaches were followed for the selection of the window shape. The first one is the use of a simple square where the window size corresponds to the size of the square's sides. Additionally, a circular window was also utilized with a diameter equal to the window size. The usage of a circular rolling window was motivated by the fact that swirling motions most commonly present close to circular shapes.

The proposed evaluation mechanism is utilized on the data that have been produced after the application of the different pre-processing techniques that were described above. The results produced highly depend on the efficiency of the preceding procedures, as it will be apparent in the results section. Moreover, the selection of a suitable window size is crucial for the accurate detection of the primary vortex flow, if such a flow exists, on the provided solar data.

4.5 Toy models

Last but not least, a number of toy models were designed in an attempt to guide the implementation procedure proposed in the current thesis. Such toy models include simple linear structures as well as more complex structures like rings and vortices. The implementation of these toy models aims to approximate the formations/patterns that might be present in solar data. By applying a variety of motions on such simple structures, one can better understand the dynamics of more composite structures and thus facilitate the process of identifying them.

In addition to simple toy models, various syntheses of more than one toy components were also implemented in an attempt to perceive the effects produced when different structures are simultaneously present in the data. It is worth noting that patterns produced by the toy models were also detectable in the considered solar data set. It is clear, however, that additional analysis is required so that one can have a better insight on the original data

by the utilization of such toy models. The aforementioned results will be briefly presented in the results section.

4.6 Computational tools and characteristics

The presented methodology has been implemented using the Python programming language (version 3.6). The developed algorithms were run on a Linux machine with 16 GB RAM and a 16-core processor. The approximate running times of the individual implementations of each method described above with an input of a (699, 330, 300) matrix have as follows:

Method	Running time
RPCA	$\sim 30\ min$
Filtering	$< 5\ sec$
Linear Removal	$< 10\ sec$

Each one of the presented pre-processing methods is applied separately on the original data set and the results produced can be found in the following chapter. Subsequently, after close examination of the aforementioned results, a complete step-by-step methodology is designed consisting of the sequential application of choice of the analyzed techniques. The output of this methodology is then provided as input to the detection mechanism and the produced results depicting the detected candidate regions for vortex flows are presented in chapter 6.

Chapter 5

Evaluation results

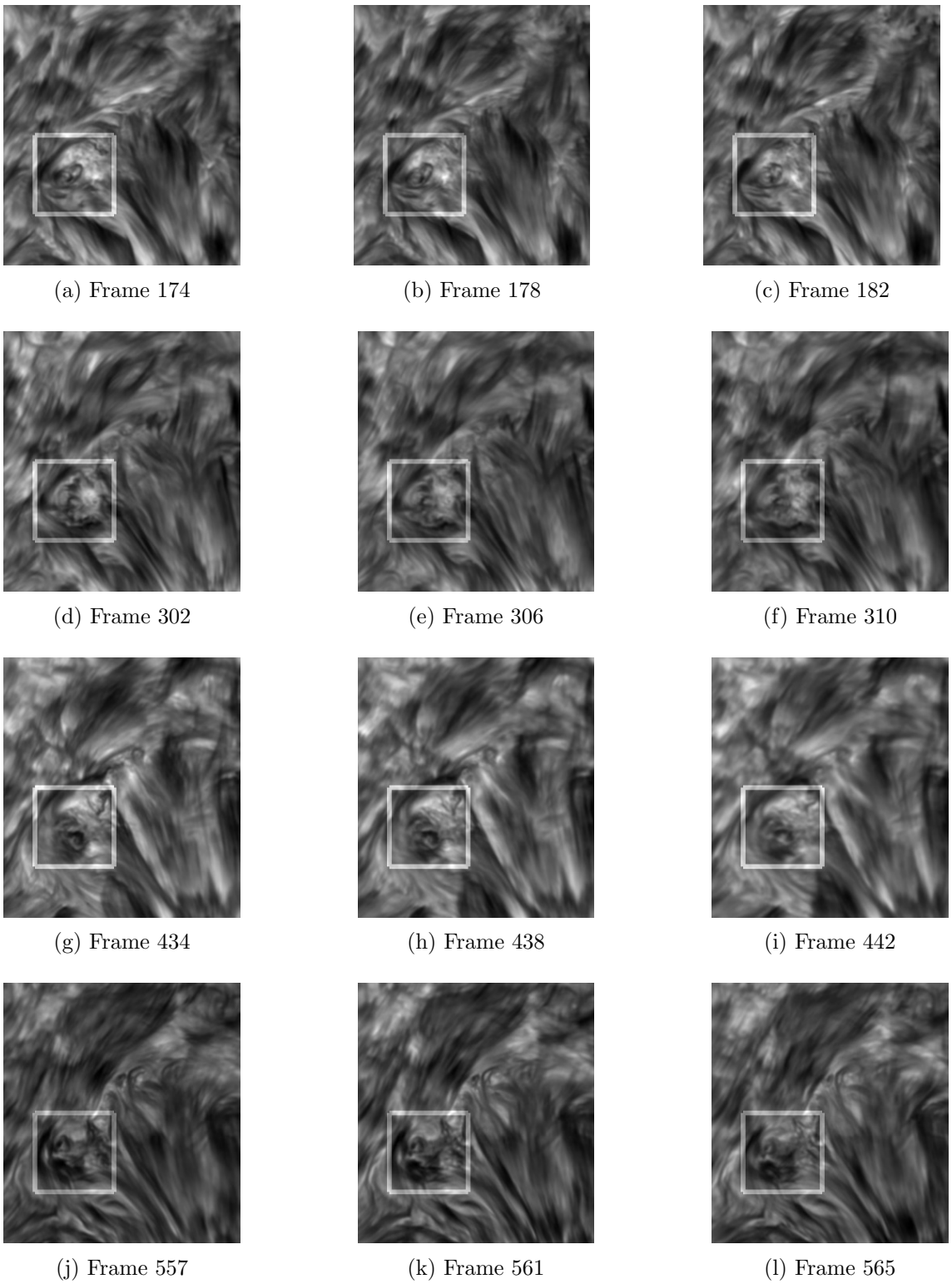
This chapter contains the collective results that were produced by applying the processing methods discussed in the previous section on the selected data set. Such results include the end-products produced by applying each methodology separately as well as results generated by the combination of different methods in a sequential manner.

The data set that has been utilized for evaluating the proposed detection methodology of solar vortex flows is a video of 699 greyscale frames (images) of size 330x300 pixels. A selection of frames from the original data set, where the conspicuous vortex flow is visible, can be found in Figure 5.1.

5.1 Robust PCA

The first method that was applied on the original data is the Robust PCA. As it was already discussed, RPCA is utilized in order to retrieve the low-rank component of a data set and the sparse one that correspond, respectively, to the static and the more dynamic parts of the frame sequence.

However, the studied solar data are characterized by complex dynamics and no static structures exist. Thus the purpose of applying RPCA on this data set is the separation of the more dynamic components from the less dynamic ones. The sparse component containing the more dynamic structures will be utilized for purposes of this thesis. This is justified from the fact that the vortex flows relate to a highly dynamical behaviour and as a result preserving the more dynamic structures will certainly maintain features related



(a) Frame 174 (b) Frame 178 (c) Frame 182

(d) Frame 302 (e) Frame 306 (f) Frame 310

(g) Frame 434 (h) Frame 438 (i) Frame 442

(j) Frame 557 (k) Frame 561 (l) Frame 565

Figure 5.1: Frames of the original data depicting the conspicuous vortex at various stages. The primary vortex flow is highlighted for easier identification.

to vorticity.

RPCA was applied on the original data as well as on processed data resulting after the application of the additional methods presented in the previous chapter, as it will be further discussed in Section 6.1. The RPCA implementation has been described in the previous chapter (Algorithm 1).

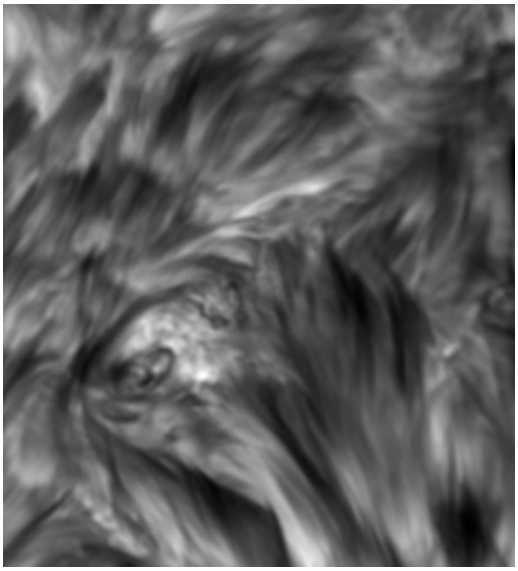
In order to provide a suitable input to the RPCA algorithm, the original data set of 699 frames, where each frame has dimensions 330x300 (a 3-D matrix), is reshaped to a 2-D tall matrix where each row contains a single pixel's values over time and each column holds the pixel values of each frame. Thus, the input matrix X has dimensions $(m, n) = (99000, 699)$. Accordingly, the values that were selected for the regularization parameter λ and the penalty parameter ρ have as follows:

$$\lambda = \frac{1}{\sqrt{max(m,n)}} \tag{5.1}$$

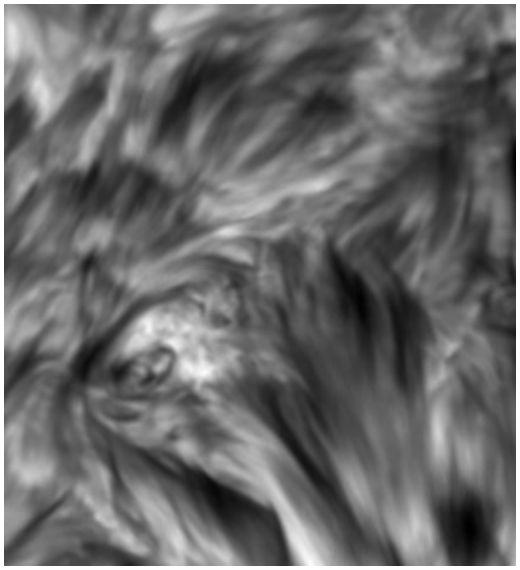
$$\rho = \frac{1}{4\overline{|X|}}, \qquad where \quad \overline{|X|} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}}{mn} \tag{5.2}$$

The procedure is terminated either when convergence is achieved or after 500 iterations. The results that were produced for the sparse S and low-rank L components of the original data for Frames 174 and 302 are shown in Figures 5.2 and 5.3 respectively. Also a set of selected frames for the sparse and the low-rank component can be found in Figures 5.4 and 5.5, respectively.

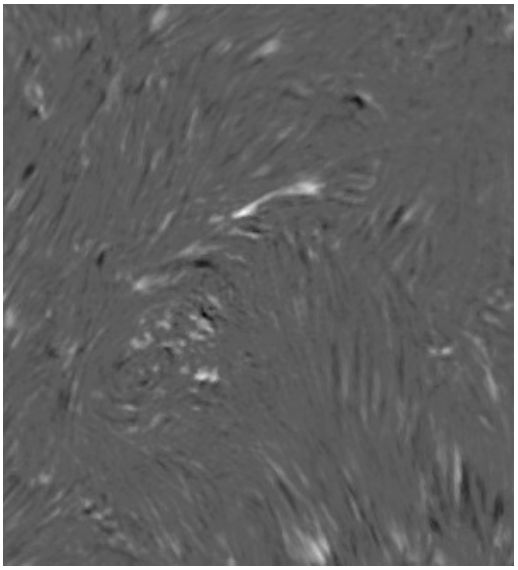
It is obvious that while the sparse component holds reduced information on each frame, the features that exhibit vorticity are preserved. Less dynamic structures are resolved as part of the Low-rank component. Consequently, RPCA is a valuable tool in the proposed procedure as it preserves more dynamic structures while discarding static or almost-static ones.



(a) Frame 174

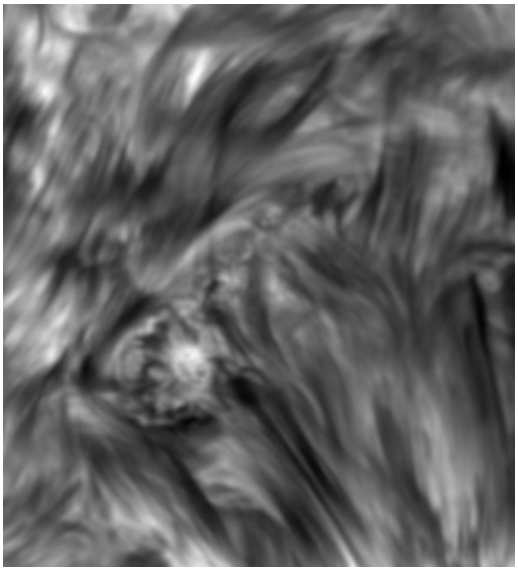


(b) Frame 174 - Low Component

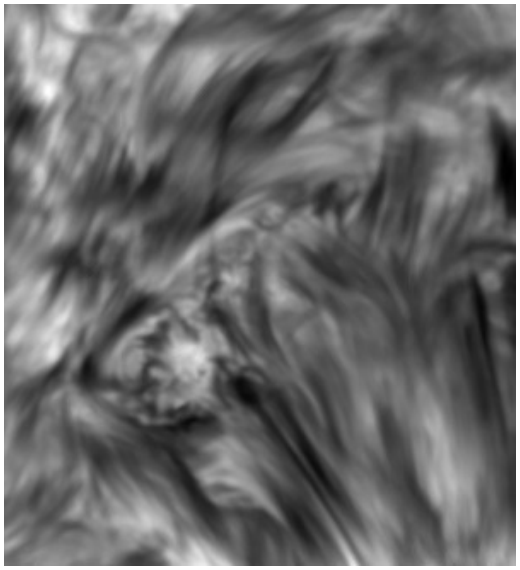


(c) Frame 174 - Sparse Component

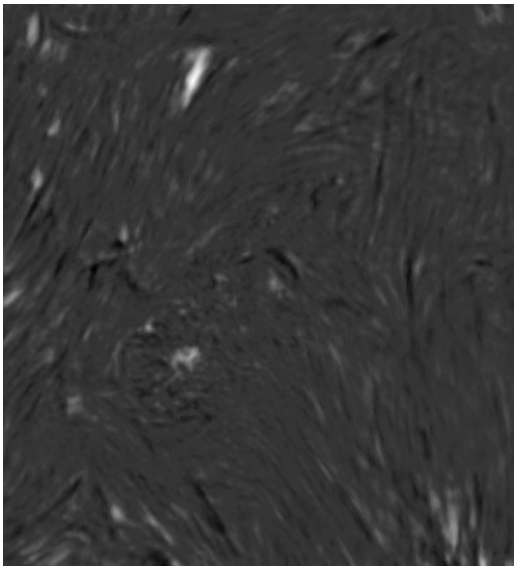
Figure 5.2: The original frame 174 as well as the Sparse and Low component that were resolved after the RPCA application.



(a) Frame 302

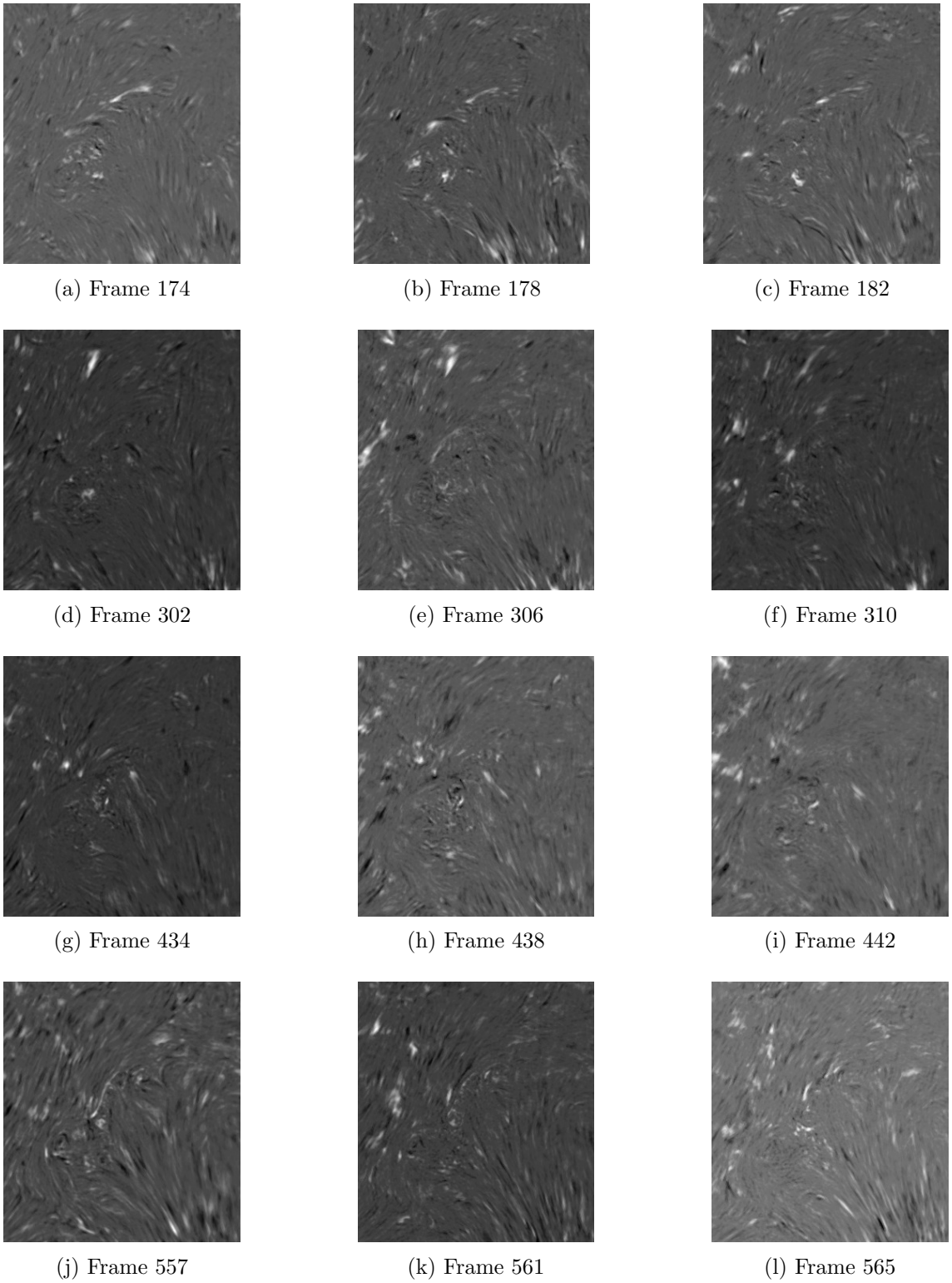


(b) Frame 302 - Low Component



(c) Frame 302 - Sparse Component

Figure 5.3: The original frame 302 as well as the Sparse and Low component that were resolved after the RPCA application.



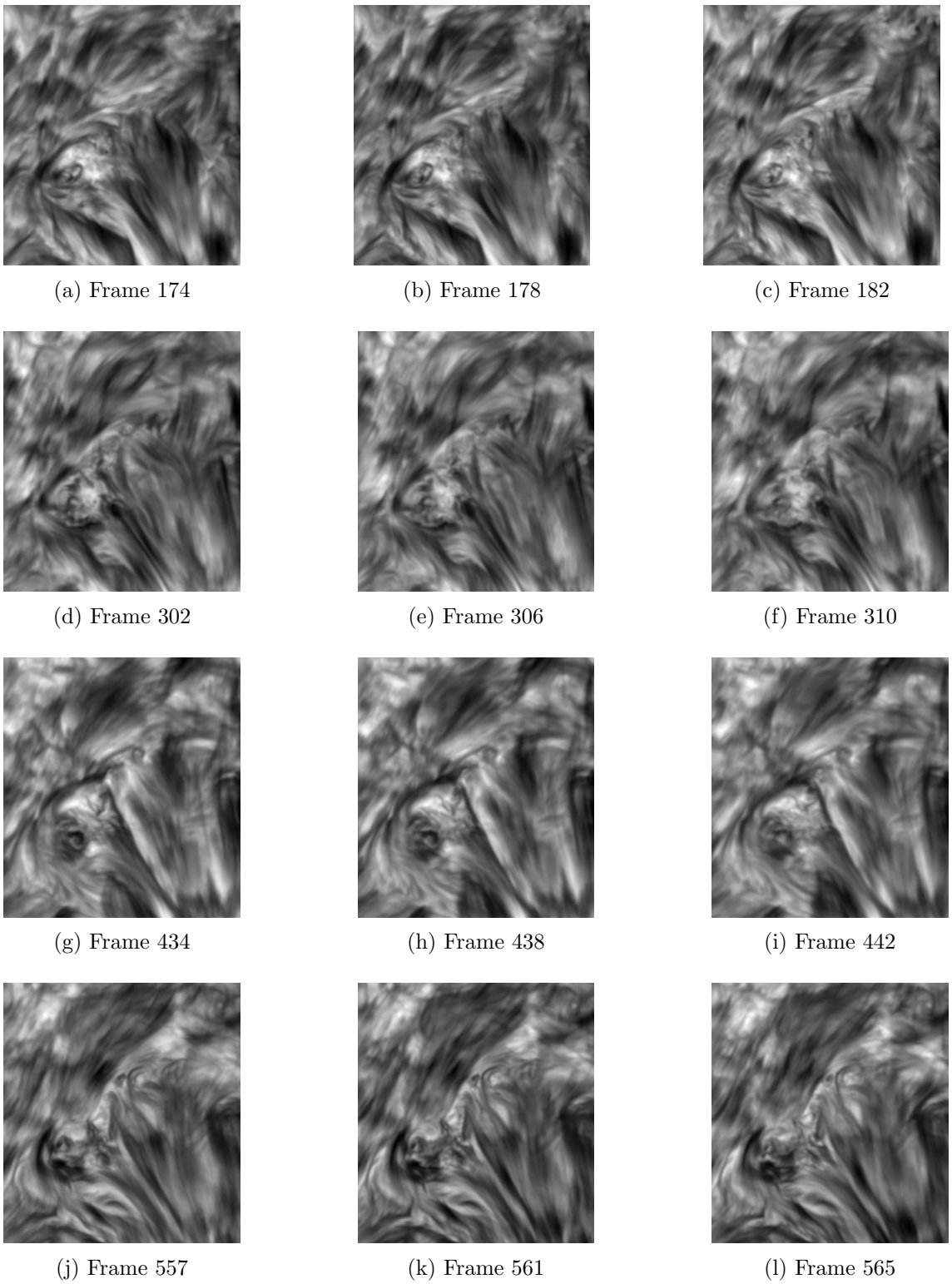
(a) Frame 174 (b) Frame 178 (c) Frame 182

(d) Frame 302 (e) Frame 306 (f) Frame 310

(g) Frame 434 (h) Frame 438 (i) Frame 442

(j) Frame 557 (k) Frame 561 (l) Frame 565

Figure 5.4: Frames depicting the Sparse component as it was resolved after the RPCA application on the original data set frames of Figure 5.1.



(a) Frame 174 (b) Frame 178 (c) Frame 182

(d) Frame 302 (e) Frame 306 (f) Frame 310

(g) Frame 434 (h) Frame 438 (i) Frame 442

(j) Frame 557 (k) Frame 561 (l) Frame 565

Figure 5.5: Frames depicting the Low-rank component as it was resolved after the RPCA application on the original data set frames of Figure 5.1..

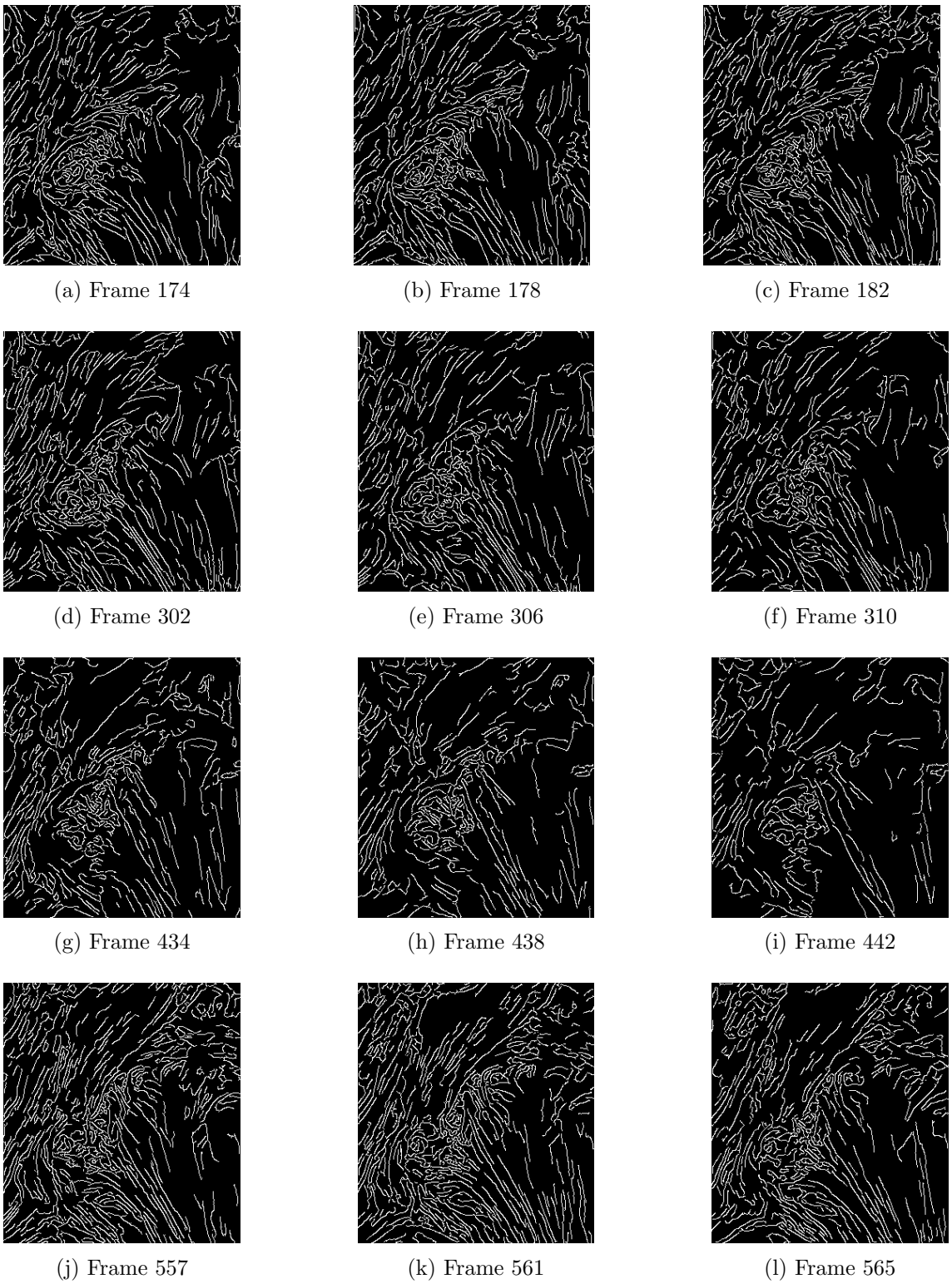
5.2 Filters

A number of filters were also applied on the original data set, as it was discussed in the previous chapter. All filters were applied on each frame separately and the frames where then merged on a frame sequence.

Specifically for the Canny Edge Detector a Gaussian kernel of size 7×7 was selected for the noise reduction step. As it was already discussed the double thresholds required were calculated by using Otsu's method.

A selection of frames produced by the application of the Canny, Sobel, DOG and Mad-max filters are provided in Figures 5.6, 5.7, 5.8 and 5.9, respectively. It is obvious that redundant information that does not contribute to the vortex detection is discarded by the application of these filters. The edge pixels detected are sufficient for the purposes of the proposed detection methodology.

Our analysis suggests that the Canny and the Sobel filter provide more satisfactory results than the other two filters and thus will provide with more reliable input if combined with other processing techniques. Subsequently, only these two filters will be further considered in our analysis.



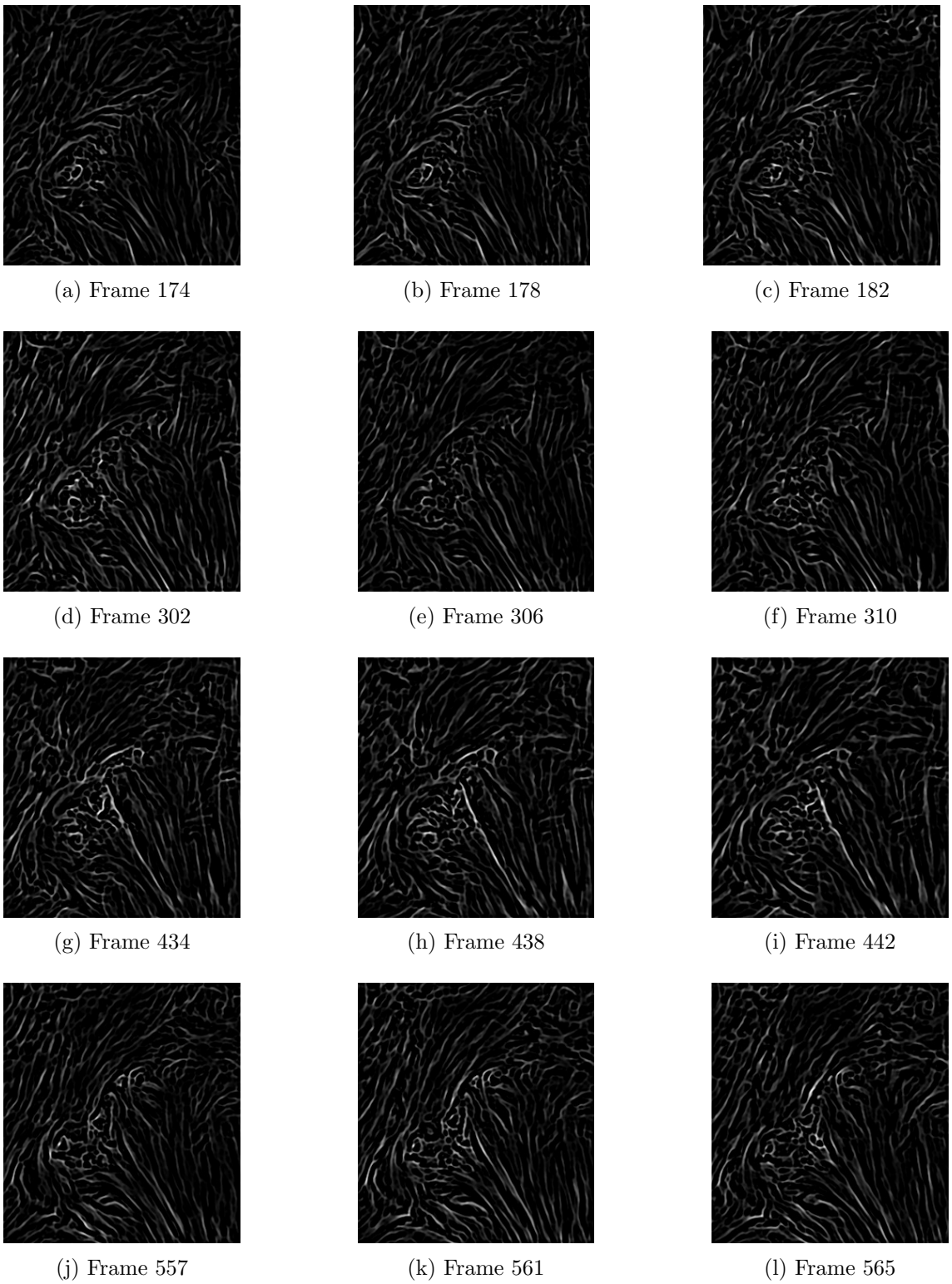
(a) Frame 174 (b) Frame 178 (c) Frame 182

(d) Frame 302 (e) Frame 306 (f) Frame 310

(g) Frame 434 (h) Frame 438 (i) Frame 442

(j) Frame 557 (k) Frame 561 (l) Frame 565

Figure 5.6: Frames of the data (similar to those of Figure 5.1) after applying the Canny filter depicting the conspicuous vortex at various evolutionary stages.



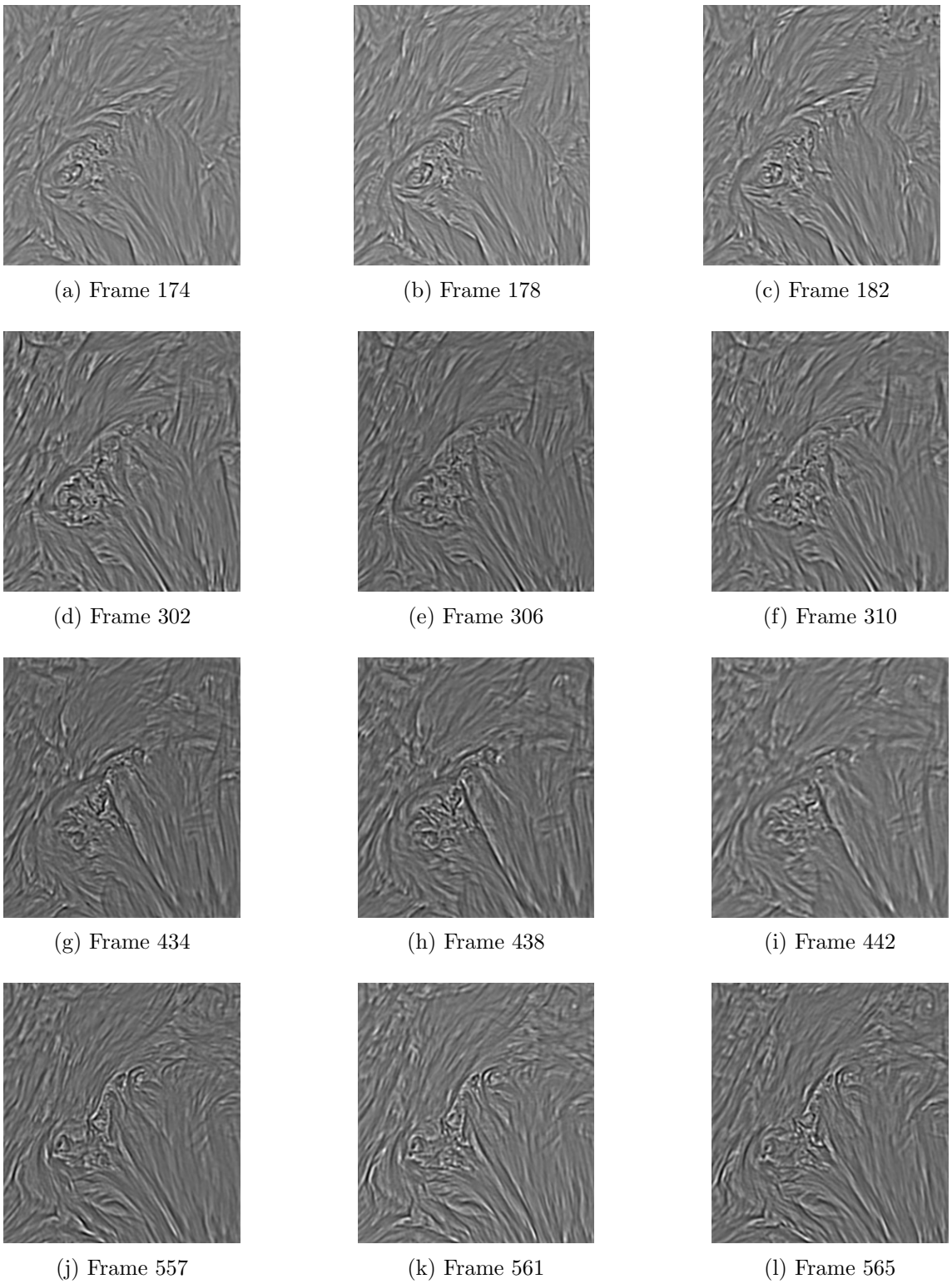
(a) Frame 174 (b) Frame 178 (c) Frame 182

(d) Frame 302 (e) Frame 306 (f) Frame 310

(g) Frame 434 (h) Frame 438 (i) Frame 442

(j) Frame 557 (k) Frame 561 (l) Frame 565

Figure 5.7: Frames of the data (similar to those of Figure 5.1) after applying the Sobel filter depicting the conspicuous vortex at various evolutionary stages.



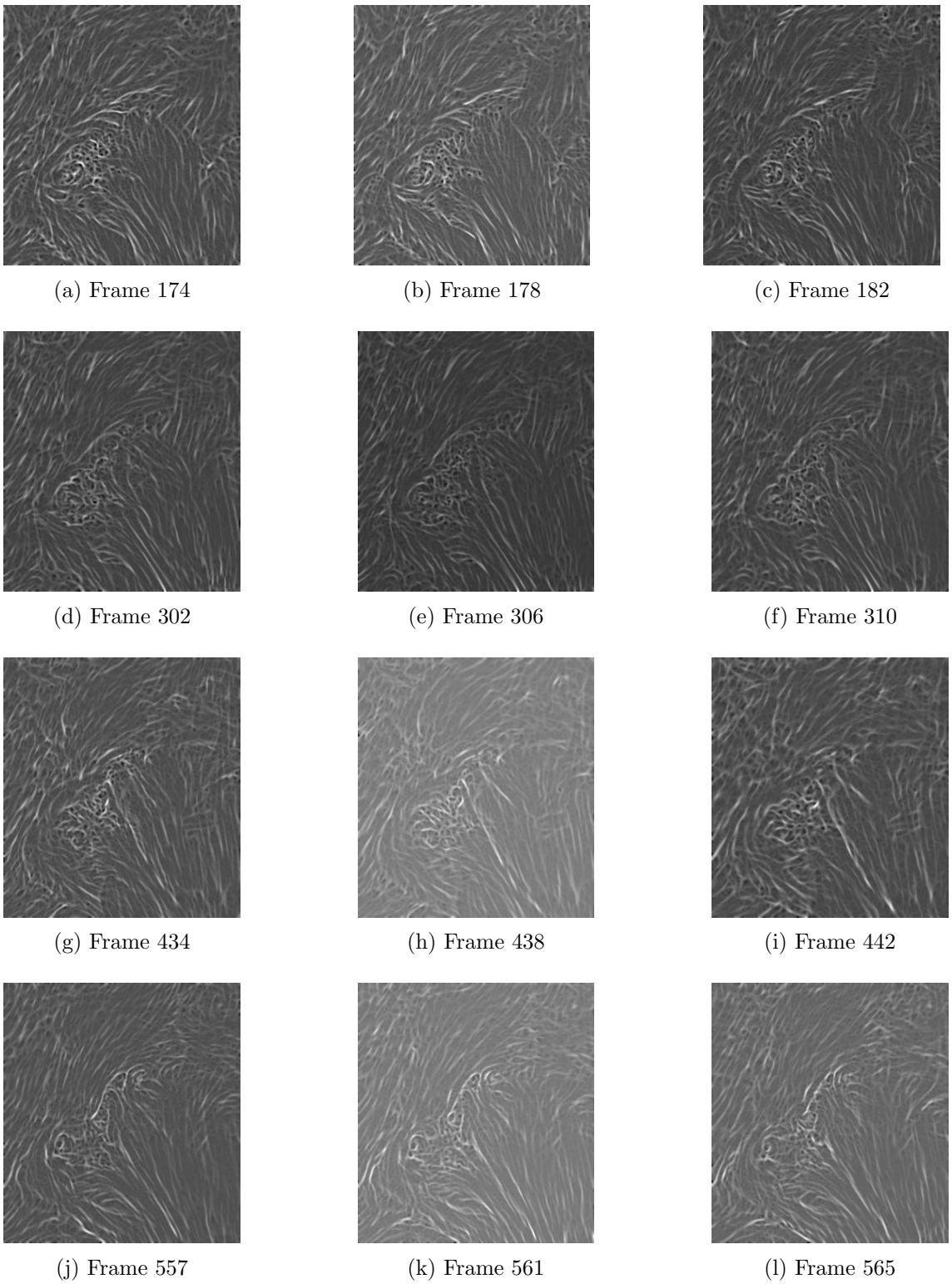
(a) Frame 174 (b) Frame 178 (c) Frame 182

(d) Frame 302 (e) Frame 306 (f) Frame 310

(g) Frame 434 (h) Frame 438 (i) Frame 442

(j) Frame 557 (k) Frame 561 (l) Frame 565

Figure 5.8: Frames of the data (similar to those of Figure 5.1) after applying the DOG filter depicting the conspicuous vortex at various evolutionary stages.



(a) Frame 174 (b) Frame 178 (c) Frame 182

(d) Frame 302 (e) Frame 306 (f) Frame 310

(g) Frame 434 (h) Frame 438 (i) Frame 442

(j) Frame 557 (k) Frame 561 (l) Frame 565

Figure 5.9: Frames of the data (similar to those of Figure 5.1) after applying the Madmax filter depicting the conspicuous vortex at various evolutionary stages.

5.3 Removal of "almost linear" structures

The removal of "almost linear" components is applied on the processed data produced after the application of the Canny and the Sobel filters. Different combinations of the input parameters were used and the produced results were visually examined in order to address which combination preserves non-linear components while discarding "almost linear" more efficiently.

Among the input parameters, the connectivity value, describing when two pixels will be considered connected, proved to highly affect the derived results. Additionally, the selection of different threshold values (for discarding components with a standard deviation STD_c below the chosen threshold value) produced a variety of outputs. High threshold values result in discarding a greater number of components which could lead to the accidental removal of structures related to vortices. On the other hand, low threshold values preserve most of the components and the "almost linear" removal procedure is inefficient.

For the Canny filter, a connectivity value of 2 and a threshold of 1.2 for STD_c produced the best results (the threshold value has been selected experimentally). Selected results of this procedure can be seen in Figure 5.10, where some frames are depicted before and after the application of the "almost linear" components removal procedure.

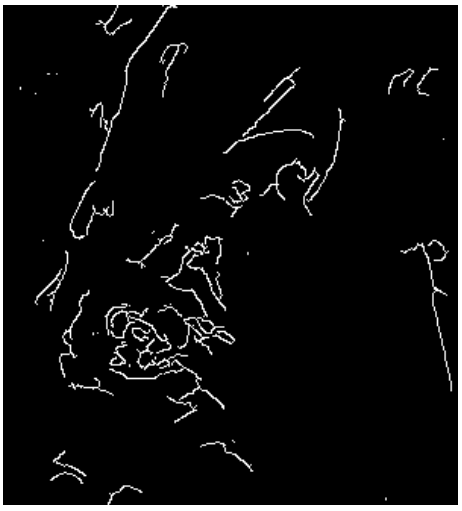
As far as the Sobel filter is concerned, a connectivity value of 2 also produced the best results. Additionally, a threshold value of 2.5 provided the most efficient removal of "almost linear" structures. Respective results are shown in Figure 5.11.

It is obvious that most structures that present a high linearity are removed from the frames. However, there still exist some components that appear to be "almost linear" and survived the procedure. Such components are most likely perceived as connected with other structures from the algorithm, even if they appear as visually separated. As a result, the larger connected components might acquire higher values than the selected STD threshold and thus they will not be discarded. Nonetheless, such a level of removal of the "almost linear" components is satisfactory for the purposes of this thesis.

(a) Frame 174

(b) Frame 174 after "almost linear" removal

(c) Frame 302

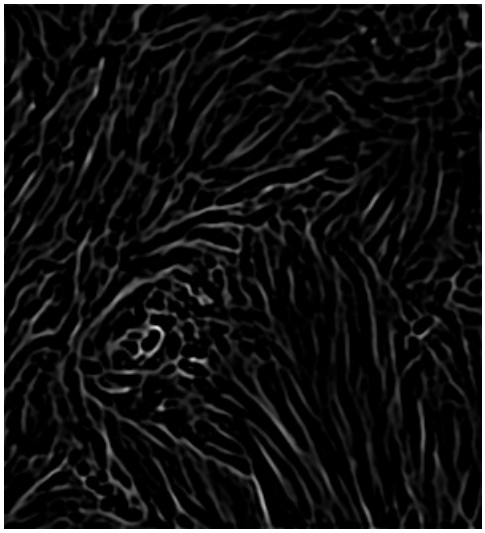


(d) Frame 302 after "almost linear" removal

(e) Frame 438

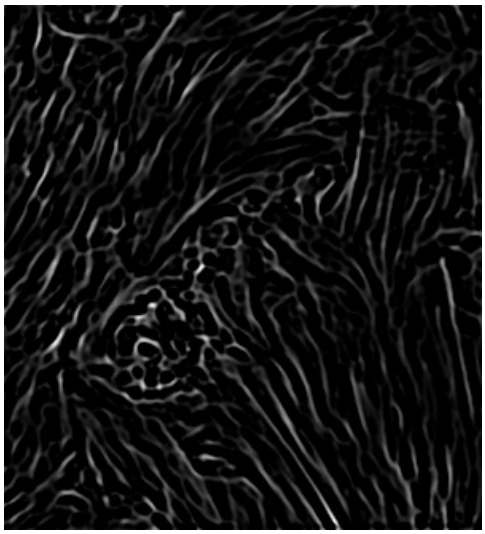
(f) Frame 438 after "almost linear" removal

Figure 5.10: Frames of the Canny processed data before (left column) and after the "almost linear" removal procedure (right column).



(a) Frame 174

(b) Frame 174 after "almost linear" removal



(c) Frame 302

(d) Frame 302 after "almost linear" removal

(e) Frame 438

(f) Frame 438 after "almost linear" removal

Figure 5.11: Frames of the Sobel processed data before (left column) and after the "almost linear" removal procedure (right column).

Chapter 6

Swirl detection methodology

The methods presented in the previous chapter are evaluated and a processing sequence of the solar data is decided, that is based on these analyzed techniques. The output of this sequence is provided as input to the proposed detection mechanism. This chapter contains the evaluation results depicting the estimated region of vortex activity produced by this procedure. Additionally, the toy models that were implemented as a guidance tool and their relevance with the actual data are also briefly presented and discussed.

6.1 Detection mechanism

The methods described and presented separately in the previous chapter should be run sequentially in order to provide a suitable input for the detection mechanism. The processing sequence selected is the following:

- Application of the Canny (or Sobel) filter on the original data set
- Removal of the "almost linear" structures
- RPCA analysis

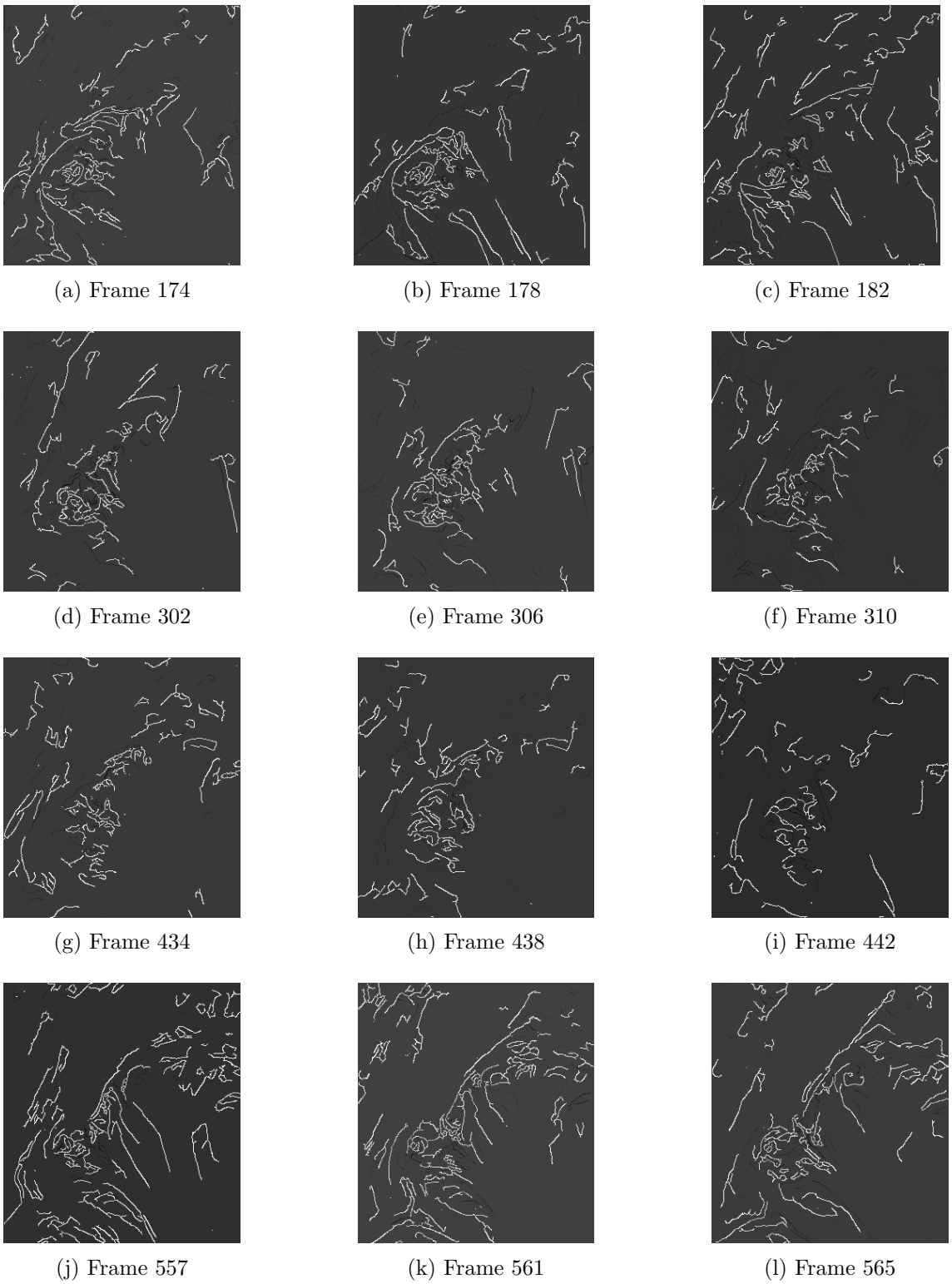
From the RPCA analysis results, the sparse component is the one preserved and evaluated. A sample set of frames produced by the methodology sequence (the sparse component of the Canny filter and the Sobel filter after the "almost linear" removal) are shown in Figures 6.1 and 6.2, respectively. These data sets will be utilized as input for the detection mechanism.

The first step of the procedure is the calculation of the mean frame F_{mean} of the input data, as it was defined in the previous chapter. The mean frames are calculated for both the Canny and the Sobel filters and can be found in Figures 6.3 and 6.4, respectively. Even from the mean frames, it is apparent that the preprocessing methods applied are efficient in removing "almost linear" and non-dynamic structures that were irrelevant to the vortex flows present. On both filters, especially the Canny one, the vortex region is clearly distinguishable in the F_{mean}.

Using the mean frame, the Probability Distribution Image (PDI) can be created as it was discussed in the previous chapter. Different PDIs calculated for a variety of window sizes and shapes (square and circular) are shown in Figure 6.5 for the Canny processed F_{mean}. Respectively, a selection of different PDIs for the Sobel F_{mean} can be found in Figure 6.6. All combinations of window sizes and shapes succeed in identifying the greater vortex region, since the values obtained by such windows close to this region are significantly higher than rest of the image. Also, it is apparent, especially with larger window sizes, that the circular-shaped windows provide greater detail in the region detection and are thus more efficient for the purposes of this thesis, as it was anticipated. We note that PDIs also depict significant substructure within this vortex flow already noticed and thoroughly described by Tziotziou etal. in [7].

Finally, the exact region corresponding to the maximum window square value can be identified, as each window can be correlated back to the region it represents. The regions of the maximum window value for the Canny F_{mean} are shown in Figure 6.7 for various square window sizes and in Figure 6.8 for various circular windows. The corresponding figures for the Sobel F_{mean} are presented in Figures 6.9 and 6.10, respectively.

It is obvious that the detection procedure successfully identifies the primary vortex flow for large window sizes. However, for smaller window sizes the main vortical region is closely missed and a proximal region is detected instead. This is attributed to the fact that more than one vortices exist on the selected data set (see [1]) and thus a more sophisticated approach for identifying all exact locations and their respective sizes is required. However, this requires additional steps that are out of the scope of the present work and it will be discussed in the "Future work" section.



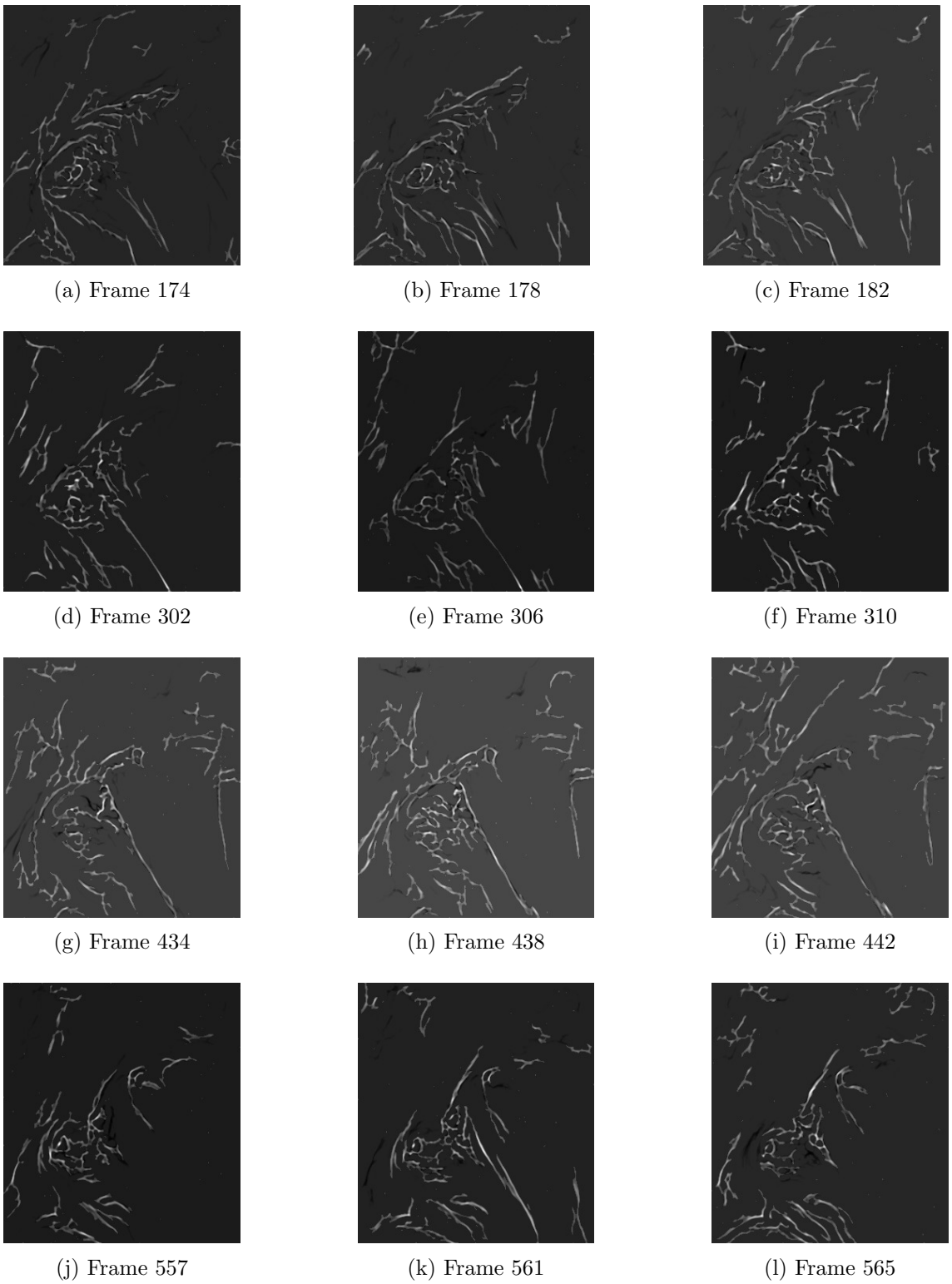
(a) Frame 174 (b) Frame 178 (c) Frame 182

(d) Frame 302 (e) Frame 306 (f) Frame 310

(g) Frame 434 (h) Frame 438 (i) Frame 442

(j) Frame 557 (k) Frame 561 (l) Frame 565

Figure 6.1: Sample frames (similar to those in Figure 5.1) of the Sparse component of the Canny processed data after the "almost linear" removal procedure.



(a) Frame 174 (b) Frame 178 (c) Frame 182

(d) Frame 302 (e) Frame 306 (f) Frame 310

(g) Frame 434 (h) Frame 438 (i) Frame 442

(j) Frame 557 (k) Frame 561 (l) Frame 565

Figure 6.2: Sample frames (similar to those in Figure 5.1) of the Sparse component of the Sobel processed data after the "almost linear" removal procedure.

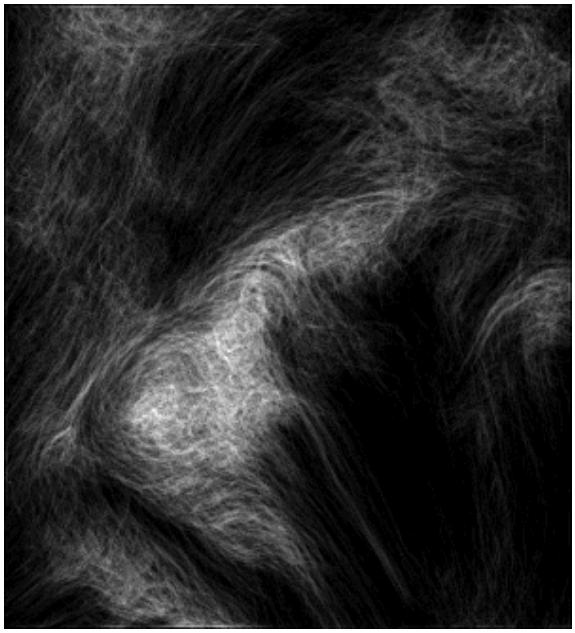


Figure 6.3: The mean frame F_{mean} of the Canny Sparse component's frames (after the "almost linear" removal).

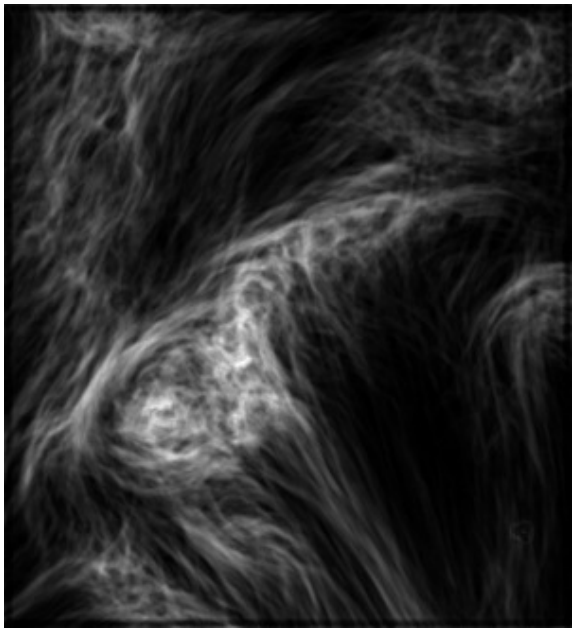
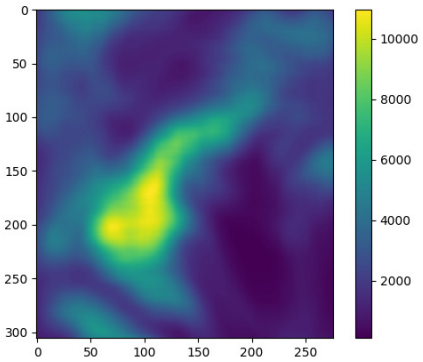
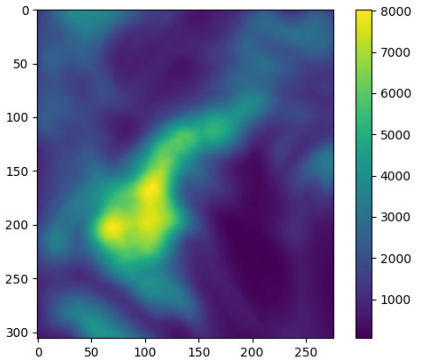


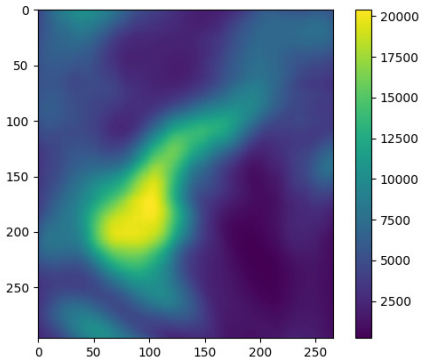
Figure 6.4: The mean frame F_{mean} of the Sobel Sparse component's frames (after the "almost linear" removal).



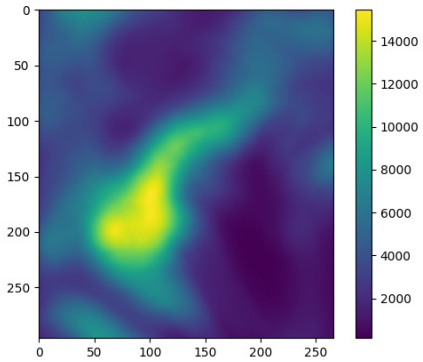
(a) Window size: 25, window shape: square



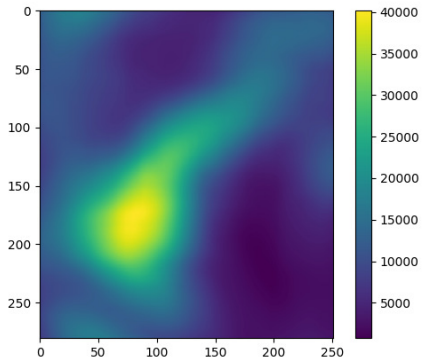
(b) Window size: 25, window shape: circle



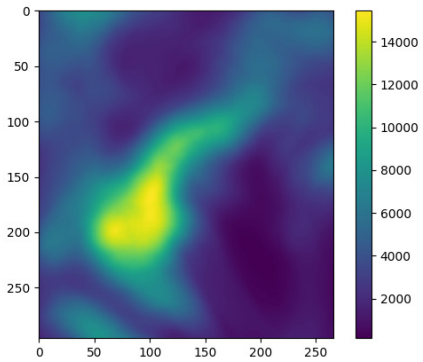
(c) Window size: 35, window shape: square



(d) Window size: 35, window shape: circle

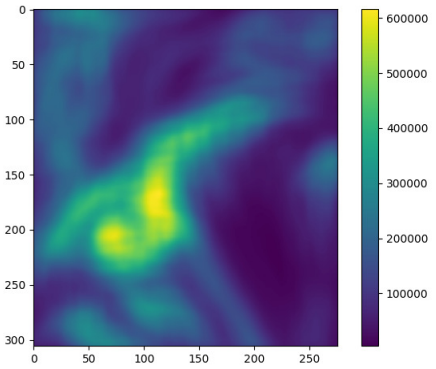


(e) Window size: 50, window shape: square

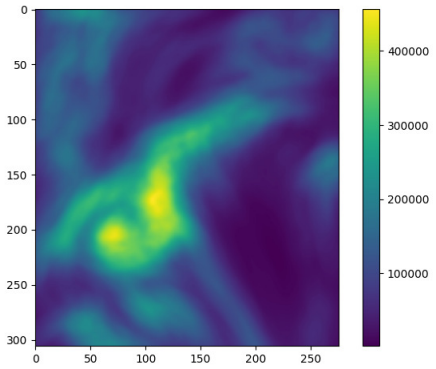


(f) Window size: 50, window shape: circle

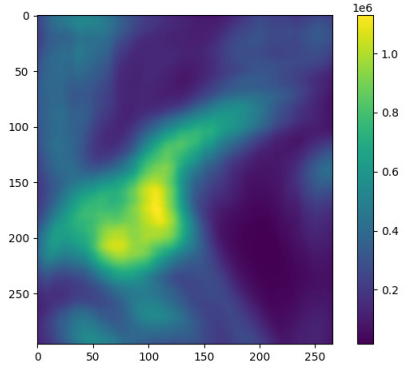
Figure 6.5: The Probability Distribution Image from the Canny mean frame for various window sizes and shapes



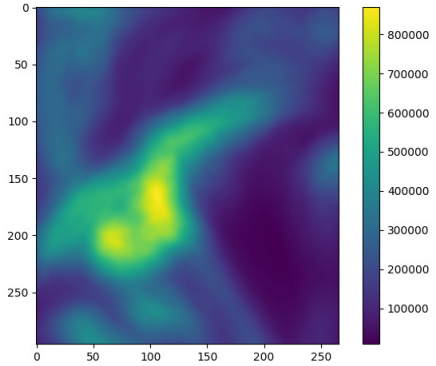
(a) Window size: 25, window shape: square



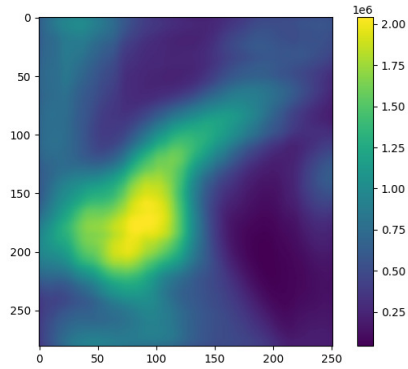
(b) Window size: 25, window shape: circle



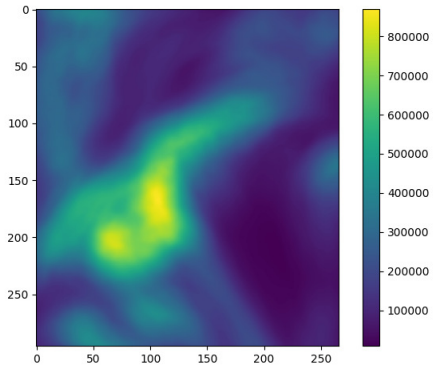
(c) Window size: 35, window shape: square



(d) Window size: 35, window shape: circle

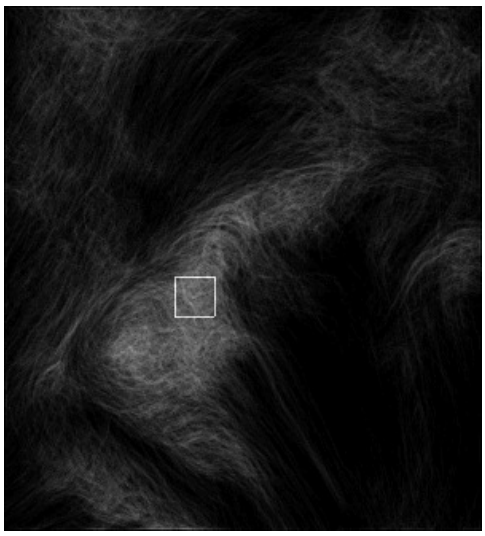


(e) Window size: 50, window shape: square

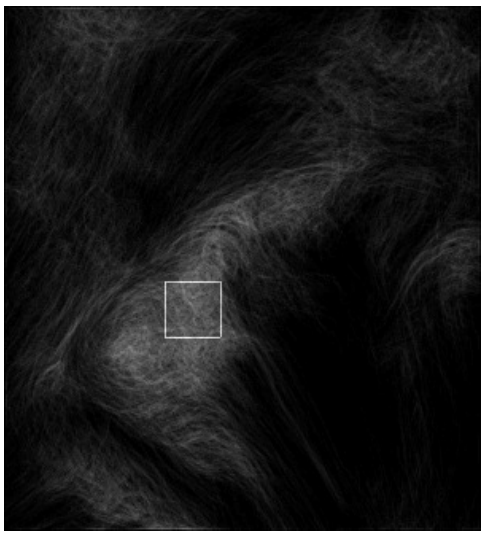


(f) Window size: 50, window shape: circle

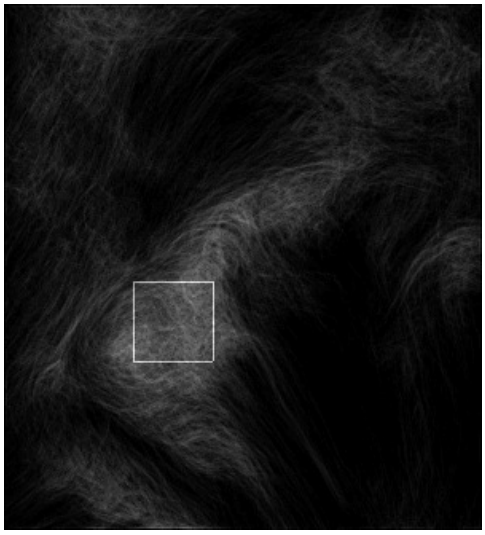
Figure 6.6: The Probability Distribution Image from the Sobel mean frame for various window sizes and shapes



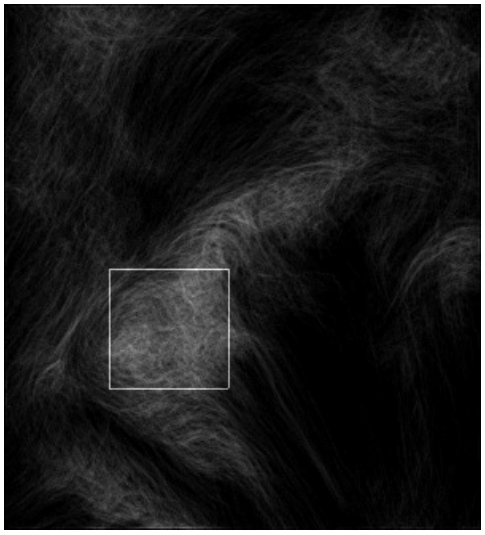
(a) Window size: 25



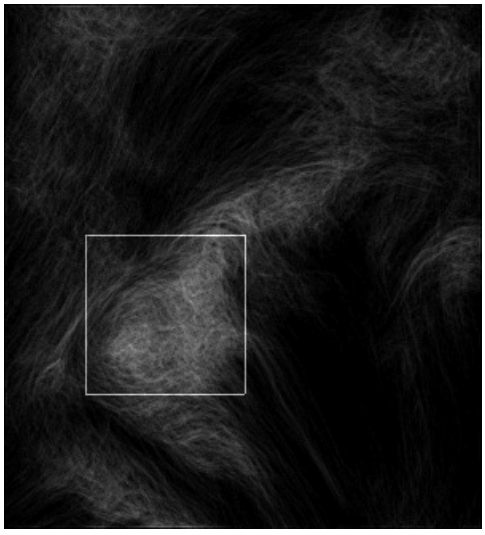
(b) Window size: 35



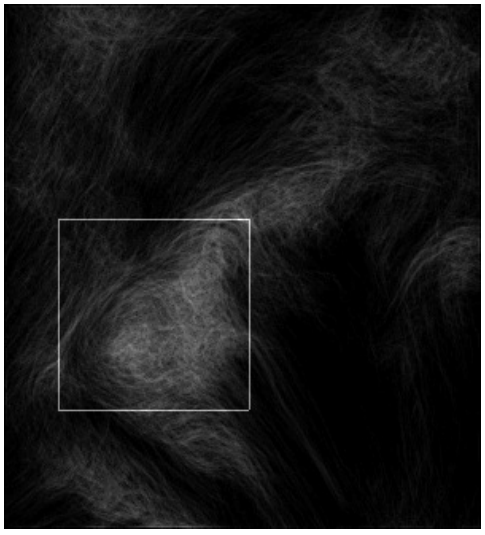
(c) Window size: 50



(d) Window size: 75

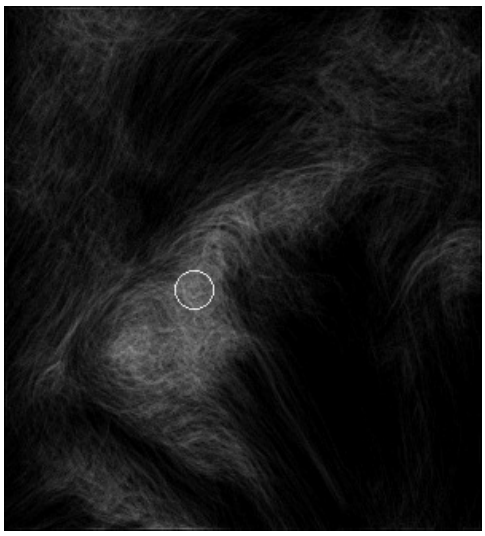


(e) Window size: 100

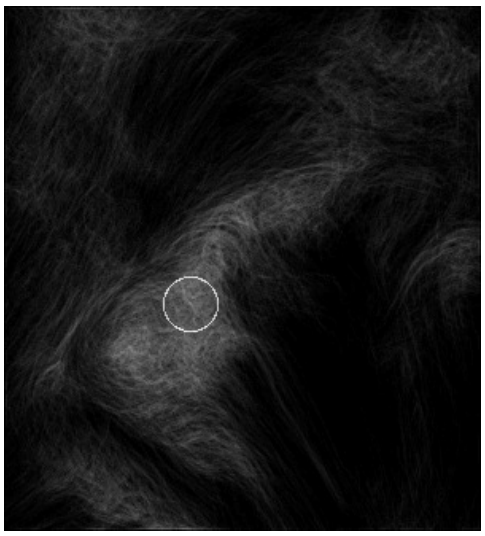


(f) Window size: 120

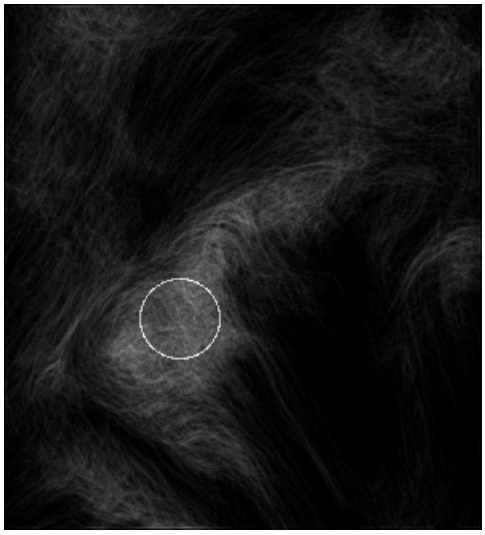
Figure 6.7: The region corresponding to the maximum square window value for the Canny F_{mean}.



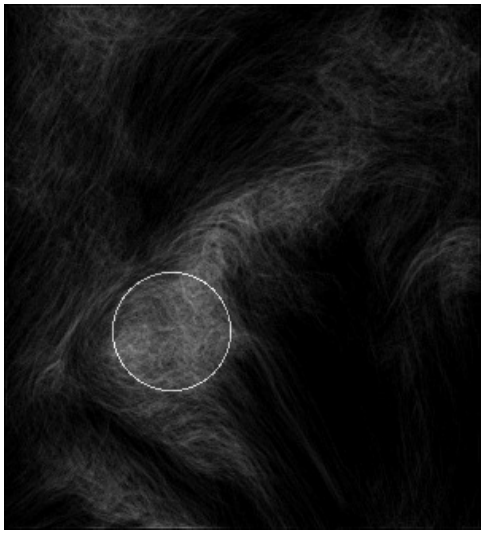
(a) Window size: 25



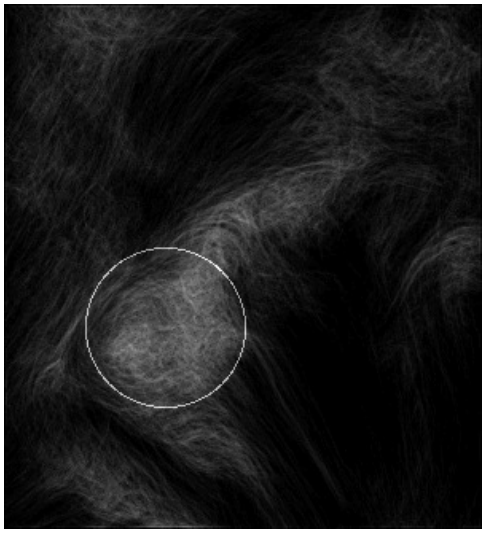
(b) Window size: 35



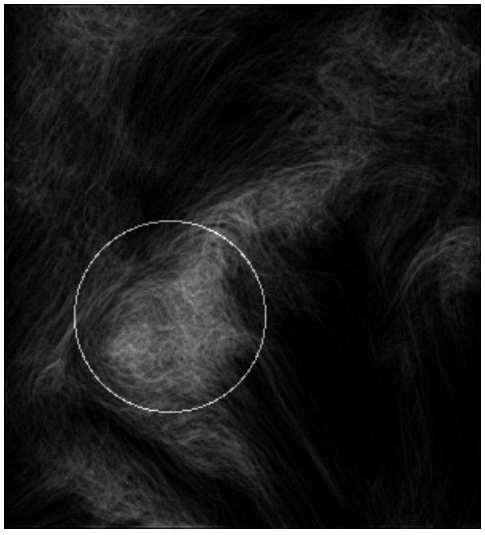
(c) Window size: 50



(d) Window size: 75

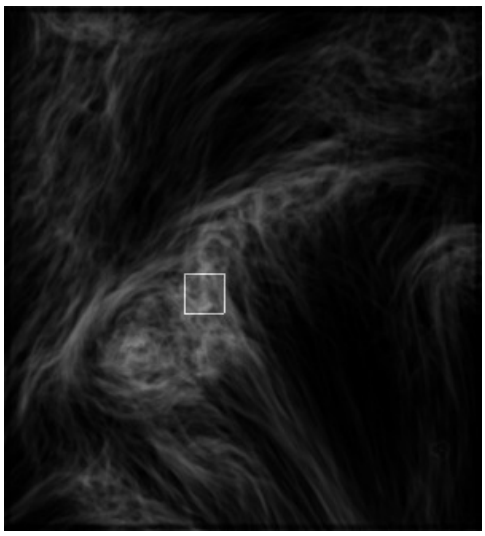


(e) Window size: 100

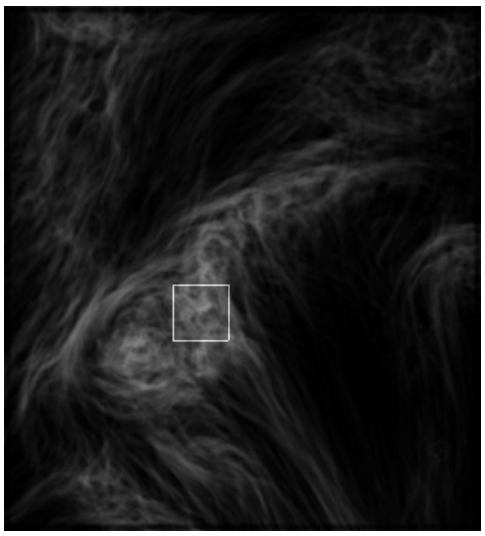


(f) Window size: 120

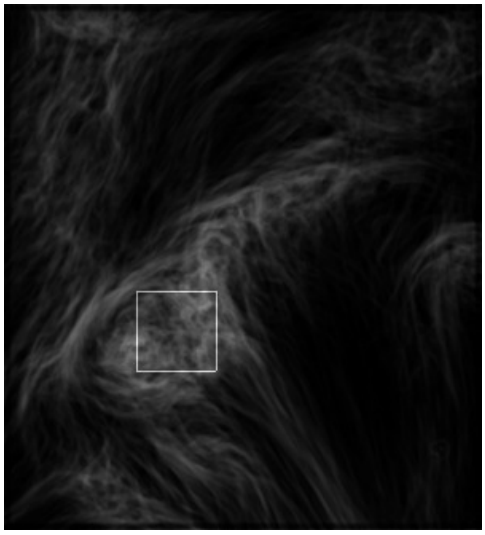
Figure 6.8: The region corresponding to the maximum circular window value for the Canny F_{mean}.



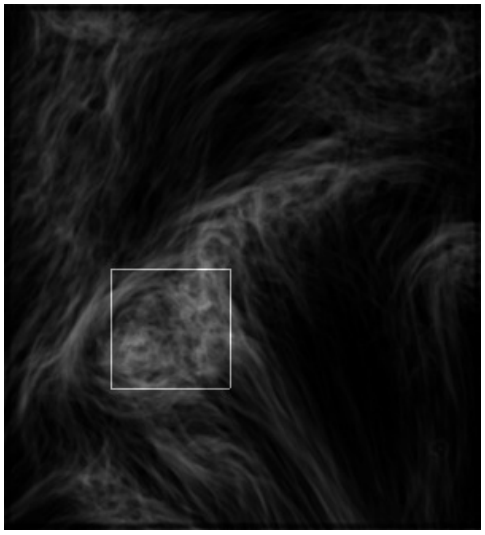
(a) Window size: 25



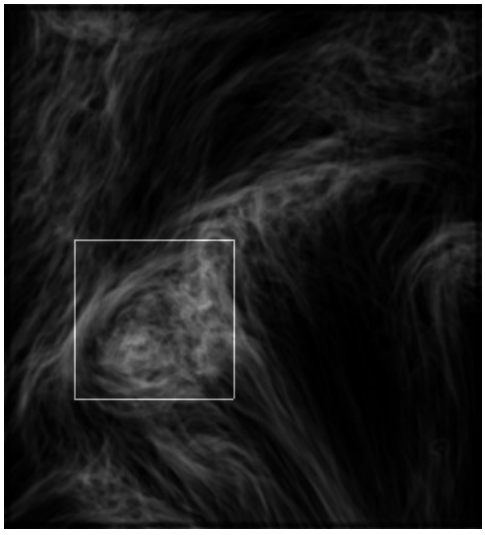
(b) Window size: 35



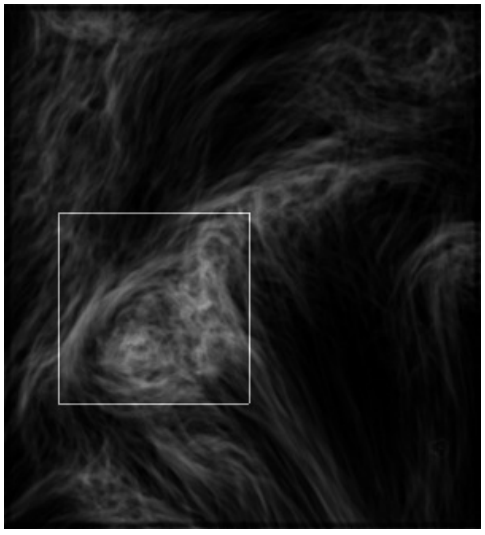
(c) Window size: 50



(d) Window size: 75

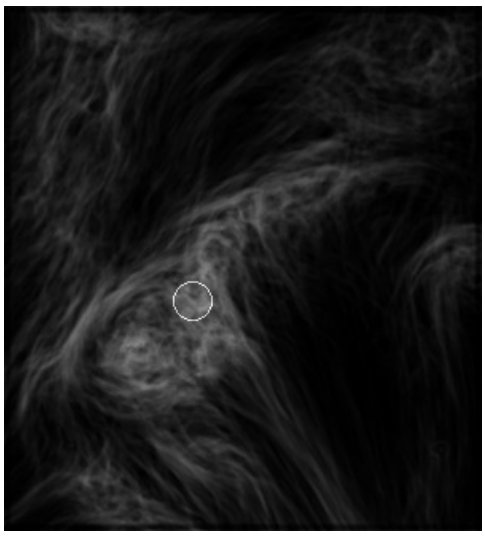


(e) Window size: 100

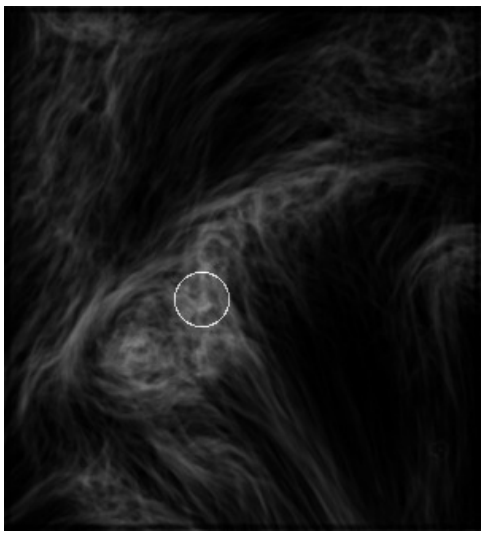


(f) Window size: 120

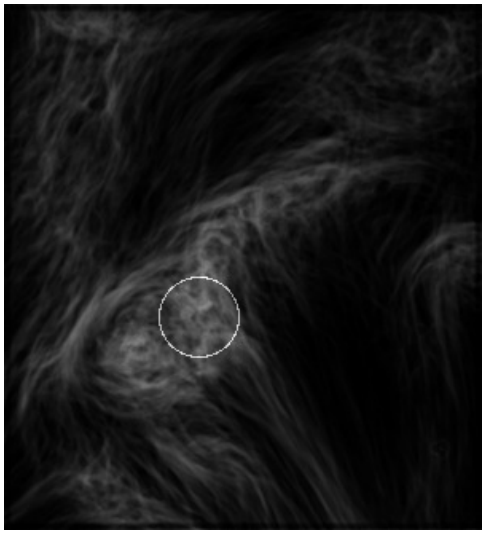
Figure 6.9: The region corresponding to the maximum square window value for the Sobel F_{mean}.



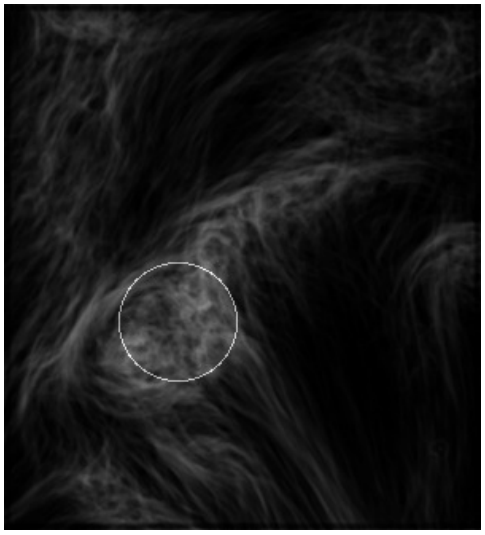
(a) Window size: 25



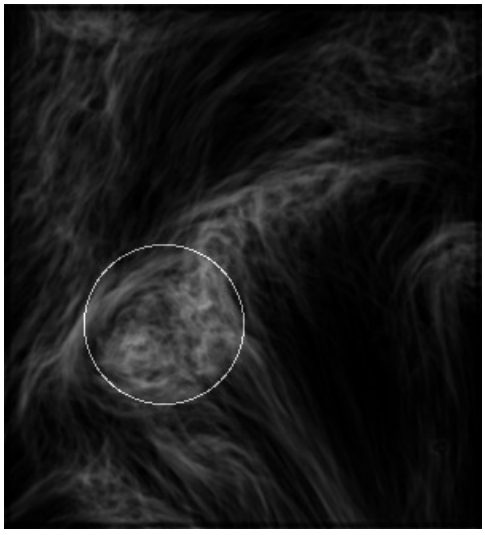
(b) Window size: 35



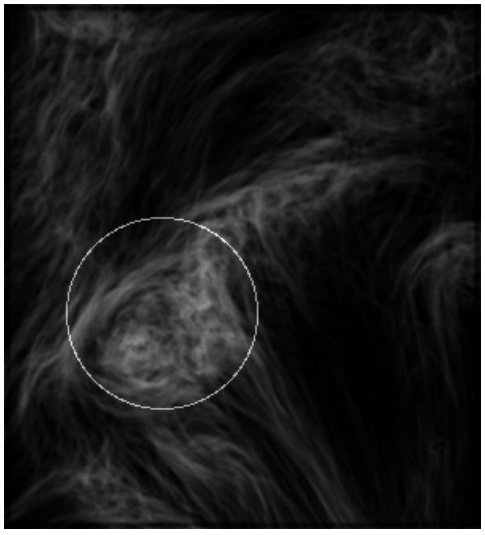
(c) Window size: 50



(d) Window size: 75



(e) Window size: 100



(f) Window size: 120

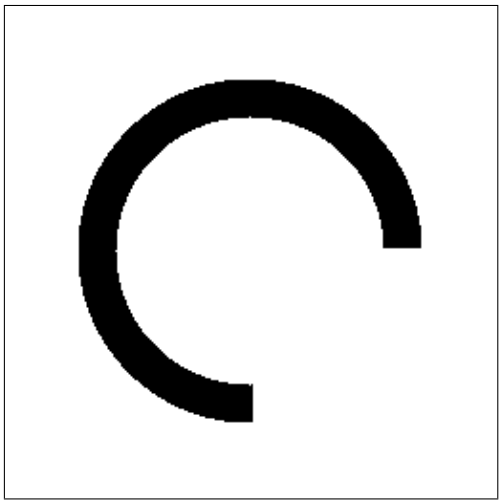
Figure 6.10: The region corresponding to the maximum circular window value for the Sobel F_{mean}.

6.2 Toy models

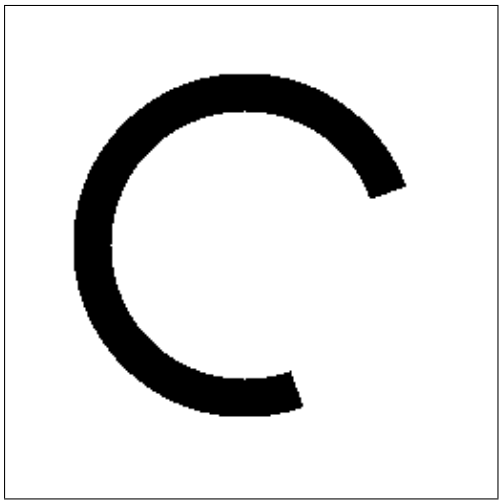
Due to the complexity of the studied solar data set, a set of simulated toy models was implemented so that the analyzed methodologies can be evaluated in a more controlled environment and a better insight on the studied structure's features can be achieved. In this section, a few of these toy models are discussed. The shapes of the toy models that were considered are approximations of the estimated shapes of actual quiet-Sun solar structures. All toy models have 650 frames and exhibit circular and/or diagonal movement.

A variety of linear, circular and elliptical shapes were implemented as well as toy examples of vortex shapes. Finally, also syntheses of two or more of the aforementioned moving shapes were used. A selection of frames from three such toy models depicting their evolution as the time progresses can be found in Figures 6.11, 6.12 and 6.13.

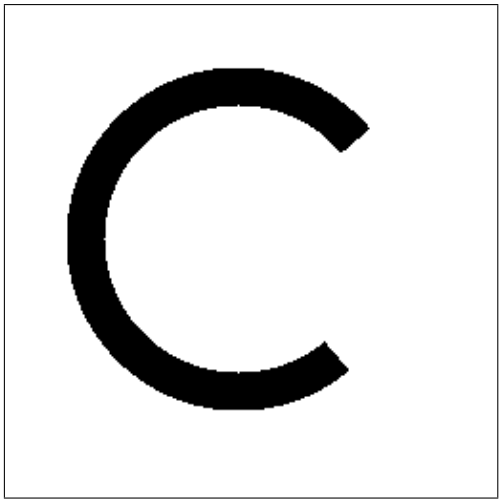
The toy models were also visualized as tall images where each column holds the pixels of a single frame and each row the progression of a single pixel's values over time. These tall images were inspected and compared with the respective tall images of actual solar data. Patterns that are created by vortex toy models were also detectable in the real solar data, as it can be seen in Figure 6.14, were small arcs, representing such vortical flows, can be found repeatedly in both representations. Thus, comparison of toy models with actual data suggests that there exists further information in the acquired RPCA images that properly combined with the present analysis and procedures could result in further improvements on the discussed automated vortex detection method.



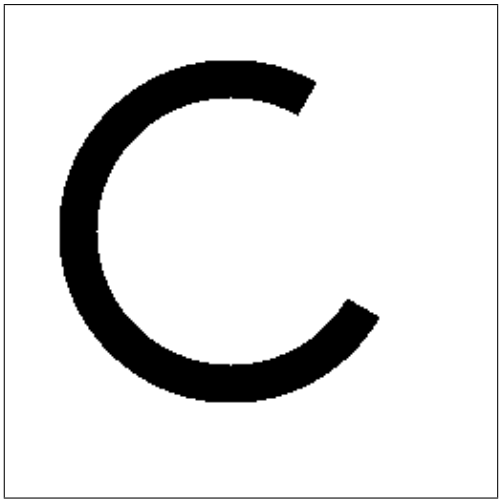
(a) Frame 1



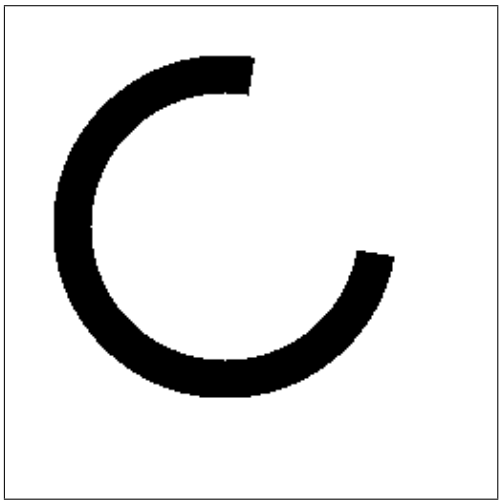
(b) Frame 20



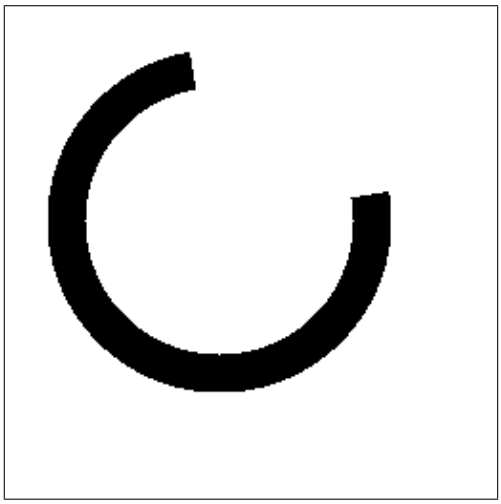
(c) Frame 40



(d) Frame 60

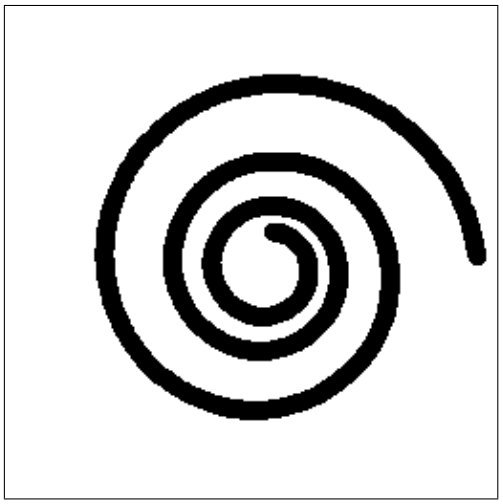


(e) Frame 80

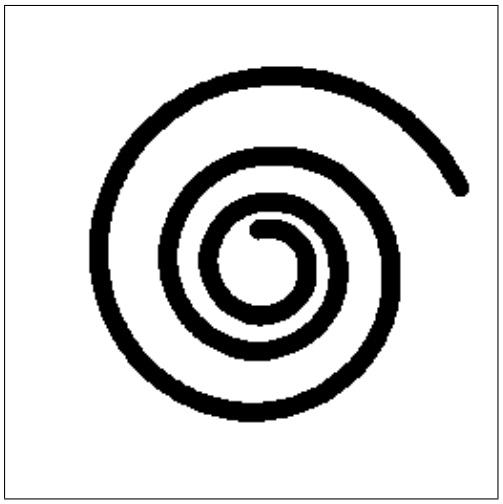


(f) Frame 100

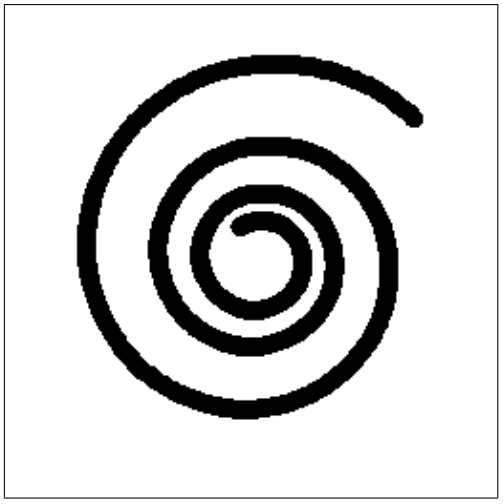
Figure 6.11: Frames of the toy model featuring a ring rotating and swinging diagonally.



(a) Frame 1

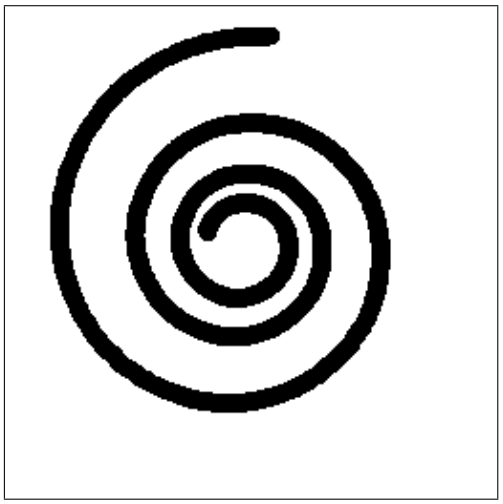


(b) Frame 20

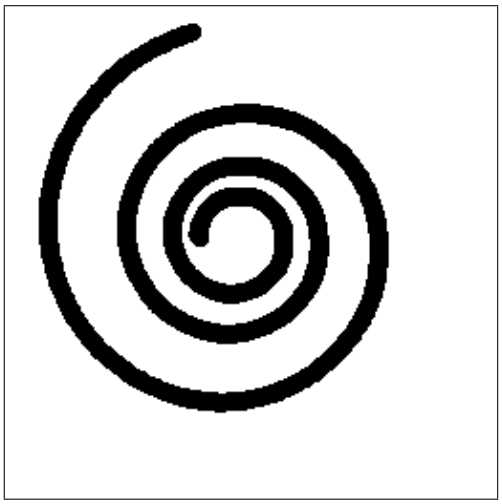


(c) Frame 40

(d) Frame 60

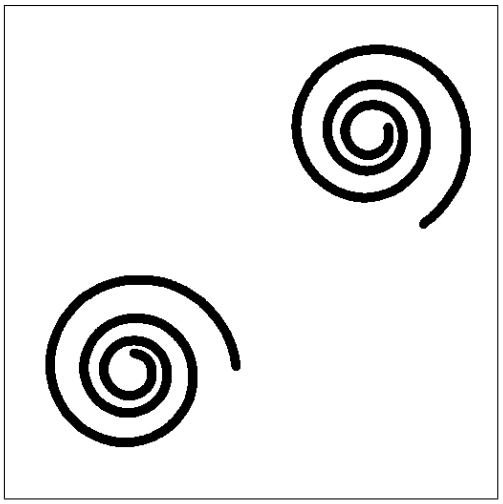


(e) Frame 80

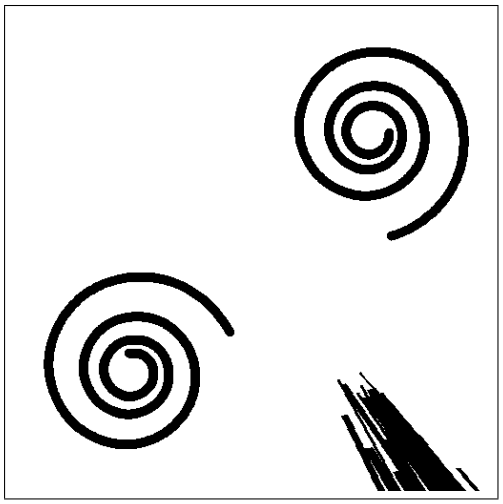


(f) Frame 100

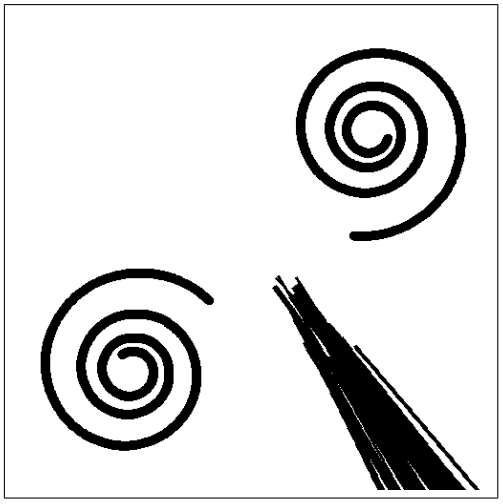
Figure 6.12: Frames of the toy model featuring a vortex rotating and swinging diagonally.



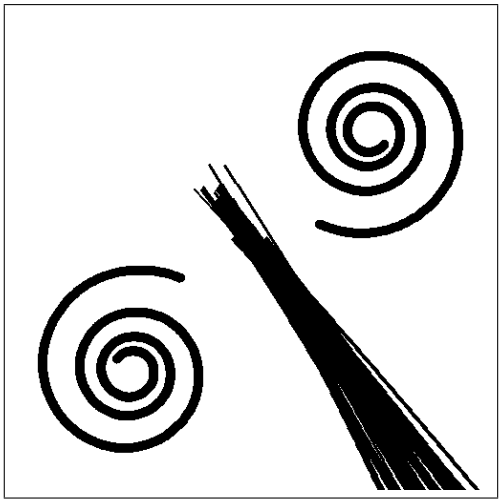
(a) Frame 1



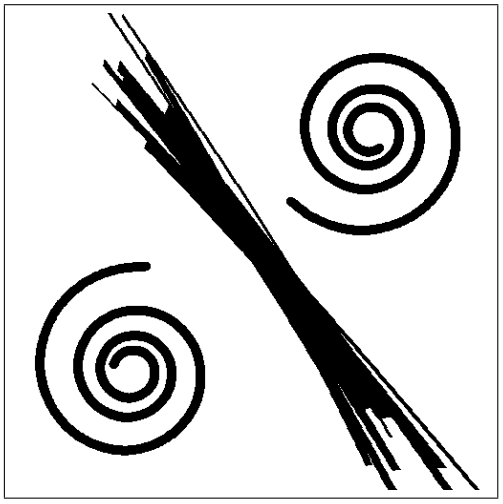
(b) Frame 20



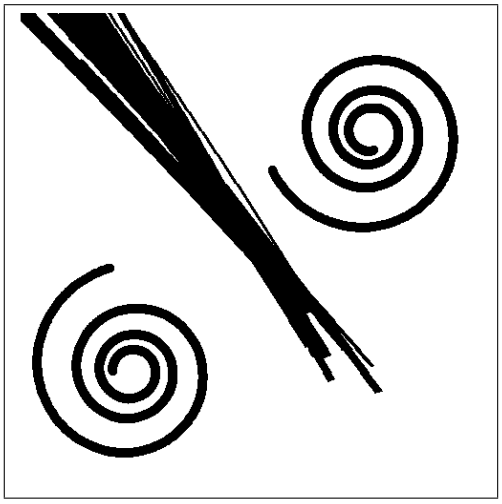
(c) Frame 40



(d) Frame 60



(e) Frame 80



(f) Frame 100

Figure 6.13: Frames of the toy model featuring 2 vortices and dynamic linear structures.

(a) Tall image of actual solar data (sparse component after RPCA)

(b) Tall image of a toy model depicting a rotating vortex.

Figure 6.14: Comparison of the solar data sparse component after RPCA and a vortex toy model. Both are visualized in tall image depictions (see text), where each column holds the pixels of a single frame. The patterns expected from the toy model are also visible in the actual solar data.

Chapter 7

Discussion and Future Work

In the current section, the produced results of the robust methodology proposed in this thesis are discussed as well as additional steps that can be implemented in the future for the improvement and further extension of the present work.

It is apparent from the results presented in the previous chapter that the proposed methodology is very efficient in predicting a major region of vorticity in solar data. The conspicuous vortex flow observed by Tziotziou et al. [7] is closely detected, especially when larger window values are utilized in the estimator mechanism. Subsequently, the presented work succeeds in identifying the primary vortex flow region when provided with a set of solar data. An expert's opinion through visual inspection or further automatic steps involving criteria based on physical characteristics of vortex flows are then required for the actual classification of the candidate region as a vortex flow or not. We note, that the proposed method seems also to detect the complex substructure (see [7] for details) that the analysed vortex flow exhibits.

As is was already underlined, the presented methodology mainly takes advantage of the morphological features that vortex flows exhibit, unlike other methodologies proposed up to this date that depend on the velocity flow derivation. The main features that the proposed methodology relies on are the dynamic nature of vortex flows and their non linear formations/structure. To this end, the pre-processing methods attempt to eliminate all structures that do not exhibit such characteristics.

However, there still exist multiple aspects of the pre-processing procedure that could

be improved or further expanded. For example, applying the pre-processing steps in a different order would potentially produce a different input for the evaluation mechanism that might lead to better results. Nonetheless, it should be noted, though, that for certain other candidate sequences additional techniques might be required.

For instance, in order to apply the "almost linear" removal procedure, a binary input is required. Subsequently, if the RPCA analysis intervened between the filters application and the "almost linear" removal procedure, the input provided would be the produced non-binary sparse component. As a result, an additional technique would be required to transform the sparse frames to binary ones. This could be achieved by the application of a Gaussian blur with a relatively small standard deviation and then the use of a mask with a selected threshold value. However, it is apparent that this procedure would introduce additional complexity and an increased investigation overhead in order to detect the optimal parameter and threshold values.

Furthermore, additional improvements to the "almost linear" removal procedure could greatly affect the produced results. The current implementation appears to be satisfactory for the selected data set but its relative crudeness could affect the evaluation of data were linear structures are denser or more interconnected with other structures. Further evaluation of different combinations of parameters and their products should take place in order to maximize the method's efficiency. Also, different techniques could be utilized in order to detect components that the current methodology evaluates as connected but should instead be perceived as separated and therefore should be possibly broken in smaller linear and non-linear parts.

Lastly, it is obvious that the application of additional pre-processing methods could greatly improve the produced results of the detection mechanism. By eliminating a greater number of structures that are unrelated to vortex flows, the proposed methodology could prove more robust and efficient in detecting vorticity. Such techniques could include the removal of other structures based on their morphology, for instance using extra knowledge provided by relevant toy models. However, it is apparent that the implementation of such techniques could prove much more complex and their utilization would require extra cautiousness, so that no vortex structures are eliminated during the process.

As far as the vortex detection is concerned, there exist a variety of aspects that could

benefit from future improvements. The current implementation only accounts for the detection of the primary vortex flow present in the input data set, if such a flow exists. However, more than one vortices can exist on such a spatial range (as it is also the case with the reference solar data set, where two or three vortices are captured). Therefore, additional effort is required for the identification of more than one windows containing vortex flow candidates. Such windows should not overlap with each other, or in case of overlap, a larger candidate window should be produced by merging them. As a result, the output of the detection mechanism would be a set of windows of variable sizes, that is bound to perform better and detect the majority of vortex flows when applied on different input data sets. A method also could be envisaged that performs detections in sub-images, resulting from an appropriate, possibly adaptive splitting of the original larger-size image where only one swirling candidate could by construct exist and therefore detected.

Additionally, the evaluation mechanism can be further improved by allowing a more flexible time window selection for the detection of vorticity. More specifically, the current proposed mechanism takes into account the entire frame sequence in order to identify the candidate vortex flow regions. However, vortex flows are quite dynamic and might be present only in a subset of sequential frames of the data set. Knowledge about the observed vortex lifetimes could be utilized in order to create a time-rolling window that would detect vortexes spatially as well as temporally. This improvement could also further facilitate the detection procedure, as irrelevant structures, that co-exist with vortex flows spatially would be easier to be detected and removed.

In possible future extensions of the current work, it is strongly suggested that the spatial window size is decided with reference to the most common radius sizes that have been derived from observed swirling motions. Additionally, the vortex flow radius estimation could also be utilized as a tool for classifying the candidate vortex region as a vortex or not. By comparing the detected window size with the most commonly observed vortex flow radius, one could argue with greater confidence on the existence of a vortex flow in the identified region. Still, an expert's opinion or additional routines would remain indispensable for the final classification of the candidate region.

Last but not least, it is of utmost importance that the proposed methodology is further evaluated using additional solar data sets as input, that might contain or not one

or more vortex structures. The results could provide valuable feedback and an important insight on the required improvements for increasing the efficiency of the presented methodology.

Chapter 8

Conclusions

The present thesis aimed to develop a robust procedure for the automatic detection of solar vortex flows. Based on the results produced by the proposed methodology, it can be concluded that this objective is achieved on a sufficient degree, as the primary vortex region is correctly identified by the presented detection mechanism. It is therefore apparent that the current work poses a very promising starting point for a fully automated and consistent detection of solar vortex flows.

The main contribution of the current thesis is the development of a novel methodology that is independent from the derivation of velocity flows from the provided solar data and it is instead based on the morphological features of vortex flows. Previous state-of-the-art approaches depended on the derivation of the velocity field, which is an error-prone procedure and can reduce efficiency as it has already been discussed in the relevant section. Consequently, the implementation of a methodology that can successfully detect vortex flows by using alternative techniques is an important asset.

Still, further future work is required in order to extend and improve the proposed methodology so that it produces more efficient results and detailed detections when presented with a variety of different data sets. The dynamic vortex nature and characteristics can be further incorporated in the detection procedure thus providing greater flexibility on the number of the vortex flows detected, as well as their spatial and temporal scale. Extended evaluation and fine-tuning of the proposed procedure on additional data sets is of course still further required.

As a final note, it should be underlined once more that the detection of solar vortex flows is a fundamental task that can lead to a greater understanding of solar dynamics and provide valuable insight on still poorly understood solar mechanisms, like the corona heating, as ubiquitous vortex structures are believed to significantly contribute to the transfer energy to higher layers of the solar atmosphere. Thus, future extension of the present work is strongly encouraged so that it can be established without doubt whether the proposed procedure can produce the promising results expected. By facilitating the easy automatic detection of solar vortex flows, the current work might allow for an in-depth study of their characteristics and dynamics that can have a substantial impact of the understanding we currently possess about the finest structures of our nearest star, the Sun.

Bibliography

[1] Kato, Y. and Wedemeyer, S. Vortex flows in the solar chromosphere - I. Automatic detection method. *A&A*, 601:A135, 2017.

[2] Ioannis Giagkiozis, Viktor Fedun, Eamon Scullion, David B. Jess, and Gary Verth. Vortex flows in the solar atmosphere: Automated identification and statistical analysis. *The Astrophysical Journal*, 869(2):169, dec 2018.

[3] J. Cuissa and O. Steiner. Vortices evolution in the solar atmosphere: A dynamical equation for the swirling strength. 05 2020.

[4] Sami K. Solanki. Sunspots: An overview. *The Astronomy and Astrophysics Review*, 11(2):153–286, April 2003.

[5] J.M. Vaquero. An early drawing of solar granulation? *Revista mexicana de física E*, 57:152–155, December 2011.

[6] Sven Wedemeyer, Eamon Scullion, Oskar Steiner, Luc Voort, Jaime Rodriguez, V. Fedun, and R. Erdélyi. Magnetic tornadoes as energy channels into the solar corona. *Nature*, 486:505–8, June 2012.

[7] Kostas Tziotziou, G. Tsiropoula, Ioannis Kontogiannis, Eamon Scullion, and J. Doyle. A persistent quiet-sun small-scale tornado. i. characteristics and dynamics. *Astronomy & Astrophysics*, 618, July 2018.

[8] J. O. Stenflo. A model of the supergranulation network and of active region plages. *Solar Physics*, 42(1):79–105, May 1975.

[9] Philippe-A Bourdin. Plasma beta stratification in the solar atmosphere: A possible explanation for the penumbra formation. *The Astrophysical Journal Letters*, 850:L29, November 2017.

[10] O. Steiner, M. Franz, N. Gonzalez, Christian Nutto, Reza Rezaei, V. Martínez-Pillet, J. Navarro, J. Iniesta, V. Domingo, M. Knölker, Wolfgang Schmidt, P. Barthol, and and Gandorfer. Detection of vortex tubes in solar granulation from observations with sunrise. *The Astrophysical Journal Letters*, 723:L180, October 2010.

[11] Sven Wedemeyer and Oskar Steiner. On the plasma flow inside magnetic tornadoes on the sun. *Publications of the Astronomical Society of Japan*, 66, June 2014.

[12] A. Nordlund. Solar Convection. *Solar Physics*, 100:209, October 1985.

[13] I. Kitiashvili, A. Kosovichev, Neveen Mansour, Sanjiva Lele, and Alan Wray. *Physica Scripta - PHYS SCR*, 86, December 2011.

[14] Chris Kitchin. *Galaxies in Turmoil: The Active and Starburst Galaxies and the Black Holes That Drive Them.* 01 2007.

[15] A. Kramida, Yu. Ralchenko, J. Reader, and and NIST ASD Team. NIST Atomic Spectra Database (ver. 5.7.1), [Online]. Available: `https://physics.nist.gov/asd` [2020, June 4]. National Institute of Standards and Technology, Gaithersburg, MD., 2019.

[16] Jorrit Leenaarts, Mats Carlsson, and Luc Voort. The formation of the Hα line in the solar chromosphere. *Astrophysical Journal*, 749, 02 2012.

[17] O. R. White and P. R. Wilson. On the Height of Formation of H-Alpha in the Solar Chromosphere. *Astrophysical Journal*, 146:250, October 1966.

[18] Jeffrey L. Linsky and Eugene H. Avrett. The Solar H and K Lines. *The Astronomical Society of the Pacific*, 82(485):169, April 1970.

[19] Cauzzi, G., Reardon, K. P., Uitenbroek, H., Cavallini, F., Falchi, A., Falciani, R., Janssen, K., Rimmele, T., Vecchio, A., and Wöger, F. The solar chromosphere at high resolution with ibis* - i. new insights from the ca854.2 nm line. *A&A*, 480(2):515–526, 2008.

[20] Manuel Vogel. Oxford dictionary of astronomy, by ian ridpath. *Contemporary Physics - CONTEMP PHYS*, 53:1–1, 09 2012.

[21] C. Beck, L. Rubio, R. Schlichenmaier, and P. Sütterlin. Magnetic properties of g-band bright points in a sunspot moat. *http://dx.doi.org/10.1051/0004-6361:20065620*, 472, 09 2007.

[22] Jorge Sanchez Almeida, J. Bonet, Bartolomeo Viticchié, and Dario Del Moro. Magnetic bright points in the quiet sun. *The Astrophysical Journal Letters*, 715, April 2010.

[23] Tino Riethmüller, Svetlana Berdyugina, M. Schüssler, V. Pillet, A. Feller, A. Gandorfer, and J. Hirzberger. Comparison of solar photospheric bright points between sunrise observations and mhd simulations. *Astronomy & Astrophysics*, 568, June 2014.

[24] Graham Kerr, Paulo Simões, Jiong Qiu, and Lyndsay Fletcher. Iris observations of the mg ii h and k lines during a solar flare. *Astronomy & Astrophysics*, 582, 08 2015.

[25] P. N. Brandt, G. B. Scharmer, S. Ferguson, R. A. Shine, T. D. Tarbell, and A. M. Title. Vortex flow in the solar photosphere. *Nature*, 335(6187):238–240, September 1988.

[26] J. A. Bonet, I. Márquez, J. Sánchez Almeida, I. Cabello, and V. Domingo. Convectively Driven Vortex Flows in the Sun. *The Astrophysical Journal Letters*, 687(2):L131, November 2008.

[27] Goran B. Scharmer, Klas Bjelksjo, Tapio K. Korhonen, Bo Lindberg, and Bertil Petterson. The 1-meter Swedish solar telescope. In Stephen L. Keil and Sergey V. Avakyan, editors, *Proceedings of the SPIE*, volume 4853 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 341–350, February 2003.

[28] J. A. Bonet, I. Márquez, J. Sánchez Almeida, J. Palacios, V. Martínez Pillet, S. K. Solanki, J. C. del Toro Iniesta, V. Domingo, T. Berkefeld, W. Schmidt, A. Gandorfer, P. Barthol, and M. Knölker. SUNRISE/IMaX Observations of Convectively Driven Vortex Flows in the Sun. *The Astrophysical Journal Letters*, 723(2):L139–L143, November 2010.

[29] R. Attie, D. E. Innes, and H. E. Potts. Evidence of photospheric vortex flows at supergranular junctions observed by FG/SOT (Hinode). *Astronomy & Astrophysics*,

493(2):L13–L16, January 2009.

[30] S. Vargas Domínguez, J. Palacios, L. Balmaceda, I. Cabello, and V. Domingo. Spatial distribution and statistical properties of small-scale convective vortex-like motions in a quiet-Sun region. *Monthly Notices of the Royal Astronomical Society*, 416(1):148–154, September 2011.

[31] S. Wedemeyer-Böhm and L. Rouppe van der Voort. Small-scale swirl events in the quiet Sun chromosphere. *Astronomy & Astrophysics*, 507(1):L9–L12, November 2009.

[32] S. H. Park, G. Tsiropoula, I. Kontogiannis, K. Tziotziou, E. Scullion, and J. G. Doyle. First simultaneous SST/CRISP and IRIS observations of a small-scale quiet Sun vortex. *Astronomy & Astrophysics*, 586:A25, February 2016.

[33] L. Jacoutot, A. G. Kosovichev, A. A. Wray, and N. N. Mansour. Numerical simulation of excitation of solar oscillation modes for different turbulent models. *The Astrophysical Journal*, 682(2):1386–1391, aug 2008.

[34] R. Moll, Robert Cameron, and M. Schüssler. *Astronomy & Astrophysics - ASTRON ASTROPHYS*, 533, August 2011.

[35] Vögler, A., Shelyag, S., Schüssler, M., Cattaneo, F., Emonet, T., and Linde, T. Simulations of magneto-convection in the solar photosphere* - equations, methods, and results of the muram code. *A&A*, 429(1):335–351, 2005.

[36] B. Freytag, M. Steffen, H.-G. Ludwig, S. Wedemeyer-Böhm, W. Schaffenberger, and O. Steiner. Simulations of stellar convection with co5bold. *Journal of Computational Physics*, 231(3):919 – 959, 2012. Special Issue: Computational Plasma Physics.

[37] Jinhee Jeong and Fazle Hussain. Hussain, f.: On the identification of a vortex. jfm 285, 69-94. *Journal of Fluid Mechanics*, 285:69 – 94, 02 1995.

[38] J. ZHOU, R. J. ADRIAN, S. BALACHANDAR, and T. M. KENDALL. Mechanisms for generating coherent packets of hairpin vortices in channel flow. *Journal of Fluid Mechanics*, 387:353–396, 1999.

[39] Laurence J. November and George W. Simon. Precise Proper-Motion Measurement of Solar Granulation. *The Astrophysical Journal*, 333:427, October 1988.

[40] Brian Cabral and Leith Casey Leedom. Imaging vector fields using line integral convolution. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, page 263–270, New York, NY, USA, 1993. Association for Computing Machinery.

[41] Laurent Graftieaux, Marc Michard, and Nathalie Grosjean. Combining PIV, POD and vortex identification algorithms for the study of unsteady turbulent swirling flows. *Measurement Science and Technology*, 12(9):1422–1429, aug 2001.

[42] Suzana de Souza e Almeida Silva, Erico Luiz Rempel, Tiago Francisco Pinheiro Gomes, Iker S. Requerey, and Abrahamm C.-L. Chian. Objective lagrangian vortex detection in the solar photosphere. *The Astrophysical Journal*, 863(1):L2, aug 2018.

[43] G. Haller, A. Hadjighasem, M. Farazmand, and F. Huhn. Defining coherent vortices objectively from the vorticity. *Journal of Fluid Mechanics*, 795:136–173, 2016.

[44] Nicolas Gillis and Stephen Vavasis. On the complexity of robust pca and ℓ_1-norm low-rank matrix approximation. 09 2015.

[45] Emmanuel Candes, Xiaodong Li, Yi Ma, and John Wright. Robust principal component analysis? *Journal of the ACM*, 58, 12 2009.

[46] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3:1–122, 01 2011.

[47] J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4(1):238–252, Dec 1962.

[48] Magnus R. Hestenes. Multiplier and gradient methods. *Journal of Optimization Theory and Applications*, 4(5):303–320, Nov 1969.

[49] Ivan Papusha. Fast automatic background extraction via robust pca. 2011.

[50] W. Xiao, X. Huang, F. He, J. Silva, S. Emrani, and A. Chaudhuri. Online robust principal component analysis with change point detection. *IEEE Transactions on Multimedia*, 2019.

[51] Neal Parikh and Stephen Boyd. Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):127–239, 2014.

[52] John Canny. A Computational Approach To Edge Detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-8:679 – 698, 12 1986.

[53] Guang Deng and L.W. Cahill. An adaptive gaussian filter for noise reduction and edge detection. volume 3, pages 1615 – 1619 vol.3, 01 1993.

[54] N. Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[55] Irwin Sobel. An Isotropic 3x3 Image Gradient Operator. *Presentation at Stanford A.I. Project 1968*, 02 2014.

[56] Olufunke Vincent and Olusegun Folorunso. A Descriptive Algorithm for Sobel Image Edge Detection. 01 2009.

[57] O. Koutchmy and S. Koutchmy. Proceedings 10th Sacramento Peak Summer Workshop, High Spatial Resolution Solar Observations. page 217, 1989.

[58] E. Tavabi, S. Koutchmy, and A. Ajabshirizadeh. A statistical analysis of the SOT-Hinode observations of solar spicules and their wave-like behavior. *New Astronomy*, 16(4):296 – 305, 2011.

[59] J. Makhoul. Linear prediction: A tutorial review. *Proceedings of the IEEE*, 63(4):561–580, 1975.