Miguel do Espírito Santo Lourenço

Licenciado em Ciências da Engenharia

Electrotécnica e de Computadores

Geo-rectification and cloud-cover correction of multi-temporal Earth observation imagery

Dissertação para obtenção do Grau de Mestre em
Engenharia Electrotécnica e de Computadores

Orientador: André Teixeira Bento Damas Mora

Professor Auxiliar, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

Júri:

Presidente: Doutor Luís Augusto Bica Gomes de Oliveira,

Professor Auxiliar com Agregação, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

Vogais: Doutor André Teixeira Bento Damas Mora,

Professor Auxiliar, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa (Orientador)

Doutor Filipe de Carvalho Moutinho,

Professor Auxiliar, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa (Arguente)

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Fevereiro, 2020

Geo-rectification and cloud-cover correction of multi-temporal Earth observation imagery

Dedicated to Djeahcademics and everyone dear.

Acknowledgements

First and foremost, I give my thanks to Professor André Mora, who has helped me all throughout this endeavour. Without him I would have never had the chance to work on this project and meet all the wonderful people at the CA3 research group. I would like to especially thank Professor Mora for the trust he put into my work, allowing me to find my own way and supporting me whenever necessary.

I would like to thank everyone at the CA3 research group for receiving me with such warmth and friendliness. For inviting me to take part in meetings pertaining to Earth observation projects, as well as workshops and seminars. I would like to particularly thank everyone involved in both the FUELMON (PTDC/CCI-COM/30344/2017) and foRESTER (PCIF/SSI/0102/2017) projects, with whom I worked in close association. Additionally, a very special thanks for the workspace I was provided with, allowing me to work in the company of my peers.

I must also thank the FCT/UNL institution for the countless opportunities provided to me as a student, ranging from academic to artistic. Throughout this academic journey I have met many people that shaped my world view and help me grow not only as a future engineer but as a person as well. I would like to give a special thanks to Professor João Paulo Pimentão, Professor Maria Helena Fino and Professor Anikó Katalin da Costa for their good spirits which have always made me feel welcome and encouraged, to Professor Robert Wiley for letting me in on the world of glass Arts and for introducing me to a number of great people such as Mr. Richard Meitner, to Professor José Manuel Fonseca and Professor Paulo Montezuma-Carvalho for always speaking their mind and sharing their honest opinions, to Professor Isabel Ventim Neves for her never ending patience and again to Professor André Mora, for all the invaluable help provided.

A great thanks to the scientific community, which has worked tirelessly to create the world we live in today. This dissertation would have not been possible if not for the open source tools created and shared, for the online help provided and overall free dissemination of knowledge.

A very heartfelt thanks to everyone that accompanied me throughout my life up to this point. To all the friends I made along the way, which filled my life with unforgettable moments and supported me when I needed most.

And to my mother, father and siblings, which have always accepted me and supported me no matter the path I chose to take, or how distant I chose to become.

Abstract

Over the past decades, improvements in remote sensing technology have led to mass proliferation of aerial imagery. This, in turn, opened vast new possibilities relating to land cover classification, cartography, and so forth.

As applications in these fields became increasingly more complex, the amount of data required also rose accordingly and so, to satisfy these new needs, automated systems had to be developed. Geometric distortions in raw imagery must be rectified, otherwise the high accuracy requirements of the newest applications will not be attained.

This dissertation proposes an automated solution for the pre-stages of multi-spectral satellite imagery classification, focusing on Fast Fourier Shift theorem based geo-rectification and multi-temporal cloud-cover correction.

By automatizing the first stages of image processing, automatic classifiers can take advantage of a larger supply of image data, eventually allowing for the creation of semi-real-time mapping applications.

Keywords: Geo-rectification; Geo-reference; Cloud cover; Temporal fusion; Image registration; Remote sensing; Multi-spectral; Multi-temporal; Sentinel-2.

Resumo

Durante as últimas décadas, avanços na tecnologia de captura remota levaram à proliferação em massa de imagens aéreas. Por sua vez, estes avanços deram lugar a novas possibilidades relacionadas com classificação de cobertura terrestre, cartografia, entre outras.

À medida que aplicações nestes campos se foram tornando mais complexas, a quantidade de informação necessária aumentou em conformidade. De modo a satisfazer estas novas necessidades, sistemas capazes de funcionar autonomamente tiveram de ser desenvolvidos. No entanto, distorções geométricas presentes nas imagens em estado primário têm de ser rectificadas de forma a garantir que a precisão exigida pelas novas aplicações seja atingida.

Esta dissertação propões uma solução capaz de automatizar os processos preliminares referentes à classificação de imagens multi-espectrais provenientes de satélites, focando-se no teorema de translação rápida de Fourier para a geo-rectificação e em métodos baseados em fusão temporal para a correcção de cobertura de nuvens.

Automatizando as primeiras tarefas referentes ao processamento de imagem, os classificadores automáticos poderão tomar proveito de uma maior quantidade de observações, permitindo a eventual construção de mapas em tempo semi-real.

Palavras-chave: Geo-rectificação; Georreferenciação; Cobertura de nuvens; Fusão temporal; Registo de imagem; Multi-espectral; Multi-temporal; Sentinel-2.

Contents

List of tables

List of images

Acronyms

ABM - Area-based matching

AOT - Aerosol Optical Thickness

BOA - Bottom of Atmosphere

BRISK - Binary Invariant Scalable Keypoints

CC - Cross-correlation

CEOS - Committee on Earth Observation Satellites

CNES - *Centre National d'Etudes Spatiales*

CNN - Convolutional neural network

CP - Control point

DEM - Digital Elevation Model

DFT - Discrete Fourier transform

EMETSAT - European Organization for the Exploration of Meteorological Satellites

ESA - European Space Agency

ETM+ - Enhanced Thematic Mapper Plus

EU - European Union

FAST - Features from Accelerated Segment Test

FBM - Feature-based matching

FFT - Fast Fourier transform

GAN - Generative Adversarial Networks

GCP - Ground control point

GIS - Geographical Information System

GMES - Global Monitoring for Environment Security

GUI - Graphical user interface

JAXA - Japan Aerospace Exploration Agency

LSS - Local self-similarity

MI - Mutual information

MR - Magnetic resonance

MRF - Markov random fields

MTCD - Multi-Temporal Cloud Detection

MVLT - Multi value logical tree

NASA - National Aeronautics and Space Administration

NOAA - National Oceanic and Atmospheric Organization

OLI - Operational Land Imager

ORB - Oriented FAST and rotated BRIEF

RAM - Random access memory

RC - Residual compensation

RFM - Rational Function Model

RM - Regression model

RMSD - Root-mean-square deviation

RS - Remote sensed

RSM - Rigorous Sensor Model

SAR - Synthetic-aperture radar

SCL - Scene-average Water Vapour map

SIFT - Scale Invariant Feature Transform

SR-SIFT - Scale Restriction SIFT

SSDA - Sequential similarity detection algorithms

SURF - Speeded Up Robust Features

SVM - Support Vector Machines

TIRS - Thermal Infrared Sensor

TOA - Top of Atmosphere

TPS - Thin-plate splines

TRMM - Tropical Rainfall Measuring Mission

UAV - Unmanned aerial vehicle

UR-SIFT - Uniform Robust SIFT

USGS - United States Geological Survey

1

Introduction

1.1 Motivation

Since the early days of humanity, cartography was used to streamline knowledge of our surroundings and, in time, this practice rose to become a major scientific cornerstone. This day and age, with the growing need for large amounts of precise and reliable information, maps have become varied and complex. This increase in complexity, however, turned cartography into an extremely arduous process and, to alleviate some of these new challenges, new technologies have been developed or adapted to act as support in the creation of new and reliable maps that can be used in numerous applications.

The near constant demand for information likely points to a future where the pressing need for real-time or near-real-time mapping of, at least, medium to small sized areas will become a reality. This data would be used to guide and manage individuals in the field, providing increased situational awareness and improving efficiency. For this to happen, however, it is necessary to develop and improve automated systems capable of observing, georeferencing and classifying terrain as needed.

Some of the technologies developed and adapted in the past decade allow now for varying degrees of automation, trivializing some the more cumbersome aspects of cartography, lowering prices and accelerating the process. With these advances, multi-temporal mapping has become a viable endeavour, opening new possibilities when it comes to managing, monitoring and analysing the geographical development of almost any given region.

Aerial imagery has been used in land cover classification applications for several decades (Pereira-pires et al. 2020)(Mora et al. 2017)(Santos et al. 2016). Nowadays,

thanks to many investments and technological advances, satellite imagery (also referred to as Earth observation or spaceborne imagery) has become more prevalent and accessible. Some platforms, like Sentinel-2 from the EU Copernicus Programme, provide global multi-spectral optical imagery at high-resolutions. By also offering a high refresh rate for their sensed imagery, coupled with an open source model, some of these missions can be used to great effect in the fields of land cover analysis and classification.

Satellite imagery, like most forms of high-altitude surveying, is greatly affected by weather conditions, resulting in unwanted and sometimes critical atmospheric noise. Clouds that form in the troposphere are often responsible for partial or total obstruction of land cover, reducing the number of useful pixels present in the image. For land classification purposes, this phenomenon can reduce the quality or even critically corrupt the samples.

Finding the perfect image samples to use throughout a study can sometimes be an arduous task, since the area of interest may be partially affected by clouds during extended periods of time. For this reason, recovering image samples that have been partially affected by cloud presence is a prospect that has generated quite some interest over the years. Under the right circumstances, by employing multi-temporal techniques, it may be possible to "merge" similar samples, generating a new image composed only of noise free fragments. On the other hand, complementary data can also be used to reconstruct target samples by means of image correlation and guided prediction algorithms.

To generate fully useful samples these methods require a considerable amount of image data, sensed at a relatively high frequency. If the images are spread too far apart from one another (this being weeks or months) the final product, while noise free, may show signs of geological discrepancies.

In addition, given the unpredictable nature of most weather conditions (namely cloud presence and spread), these methods may not always be capable of offering a solution. Without noise clear patches over extended periods of time, methods based on temporal image fusion are unable to provide the necessary results. Despite the sporadic unfavourable conditions, cloud cover removal techniques based on multi-temporal approaches have proven to be extremely useful, increasing the efficiency of many applications related to land cover classification and monitoring, among others.

It is important to note that without proper referencing, image data may become mostly, if not completely, unusable in a number of applications. Throughout the sensing process, variables such as undulation, atmospheric conditions, curvature of the Earth, projection tilt angle, and so on can negatively affect both the geometric and geo-location

components of the sensed products, misleading classifiers into targeting wrong pixels and consequently contaminating results.

Manual geo-rectification of samples can be very time consuming and sometimes even unfeasible, depending on the intended purpose. For this reason, a completely automated system that can geo-reference images for a selected area has been the focus of study for many decades. Regardless, a single universal system that manages to effectively deal with the whole spectrum of geo-rectification challenges is yet to be presented and, most likely, never will be.

Given the vast nature of remote sensed imagery, coupled with the varying degrees of distortion suffered by these images, finding an effective solution often requires a deep understanding of the problem at hand. Solutions should be developed with problem specific geometric deformations and particularities in mind, meaning that this particular area of study will remain relevant for years to come.

1.2 Dissertation objectives

This dissertation proposes the creation of an automated system capable of handling the pre-stages of land cover classification analysis of multi-spectral satellite images. The whole process will be divided into two major subsections: image geo-rectification and cloud cover removal.

The geo-rectification process must be accurate enough to ensure the constant production of reliable samples. These samples will then be used by automatic classifiers that can monitor the geographical development of a specific region. Given the nature of these kinds of monitoring methods, sample reliability is key.

Monitoring and semi-real-time analysis also benefit from a high sample rate. For this reason, cloud removal by way of multi-temporal data correlation will be applied when necessary, to assure the best possible sample frequency.

The end product should be an autonomous system capable of providing geo-rectified, noise free samples to an automatic classifier, akin to Level-3 data products.

1.3 Contributions

When dealing with problems that require some form of resource management, planning is often, if not always, one of the most vital aspects to consider when aiming for an efficient and effective solution. A well-planned operation will always help save time

and reduce costs but, in the case of life-threatening situations, it can also act as a lifeline for personnel working in the field.

In some circumstances, planning is heavily dependent on existing geographical information. However, detailed mapping is an arduous and expensive process, so it becomes impractical, and sometimes unfeasible, to map the same area consecutively in a short time span. While some geographical characteristics tend to remain relatively unchanged for longs periods of time, others can change drastically, increasing the rate at which some maps become outdated. Vegetation is often one of the features that may become hard to track, given that, for example, a single fire can completely change a previously known setting.

Firefighters must often work under extremely hazardous conditions, sometimes relying only on previously existing cartography which may not accurately reflect the current situation. These maps are then updated manually in order to increase their accuracy and reliability. Under these circumstances, a tool that could reliably gather, geo-rectify and classify vegetation in a semi-real-time manner would prove to be extremely useful. Technology such as this could also be used to monitor vegetation growth within firebreaks, predicting their effectiveness and help plan any necessary maintenance.

Regarding the aforementioned problem, this dissertation will contribute towards a better solution by proposing a system capable of:

- Accurate geo-rectification of multi-spectral Sentinel-2 satellite imagery;
- Cloud removal by way of multi-temporal complementary data correlation.

1.4 Dissertation outline

This section aims to provide a general sense of how this dissertation will progress and how it will be organized. Excluding the current chapter, this dissertation is composed by seven more chapters, all of which will be described ahead in a succinct manner.

- **Chapter two:** An overview of the underlying theme of this dissertation will be presented. By glancing over the topic of remote sensed imagery (mainly satellite provided) and some of the inherent difficulties associated this kind of data processing, this chapter aims to ease the reader into the more specific topics that will be further discussed and analysed throughout the course of this dissertation.
- **Chapter three**: Responsible for encompassing all that is related to the State of the Art. As was mentioned previously, this dissertation will tackle two major challenges: image geo-referencing and cloud cover correction.

In similar fashion, chapter three will also be focused on these two topics. Firstly, various georeferencing methodologies will be presented, analysed, compared and then discussed. The remainder of this chapter will follow the same approach, only now regarding cloud cover correction techniques.

- **Chapter four:** Documentation of the work performed over the course of the dissertation. All stages of development will be examined, including all of the relevant achievements and shortcomings that acted as a foundation for the final implementation. This chapter will be focused on image preparation, geo-rectification algorithms and the cloud cover correction technique ultimately developed. It will then culminate in a detailed explanation of the plugin created to run all designed algorithms.
- **Chapter five:** Validation of the proposed geo-rectification and cloud cover correction algorithms. Opening with various tests on simulated and real-case scenarios followed by an accuracy and performance evaluation, including a final overall analysis of results.
- **Chapter six:** As the final chapter, it will be dedicated to a comprehensive conclusion, followed by a roadmap of possible future improvements.

2

Theoretical framework

2.1 Remote sensing

Remote sensing, as an ideal, refers to the acquisition of information without direct physical contact with the target. As a result, other detectable variables must be collected and analysed by means of data-based computation models, resulting in a link between the remote sensed (RS) data and the physical object. This kind of approach is based on the *inverse problem* principle.

This process hinges on the use of remote sensors that can be further classified into two distinct categories:

- **Passive:** These types of sensors respond to external stimuli by recording the energy emissions of the target. In most cases, the energy received comes in the form of radiation caused by the reflection of sunlight;
- **Active:** Contrary to passive sensors, active sensors utilize internal stimuli in order to collect the required data. An example of this method can be seen in synthetic-aperture radars (SARs), that send out electromagnetic waves in sequence and subsequently collect the echoes to create a three-dimensional reconstruction of the target.

These techniques promptly found their way into the fields of land cover and ocean monitoring, usually with the aid of an aircraft or satellite platform in more recent years. High altitude sensed data grew to become an integral part of the Information Age, with some of its aplications ranging from coastal and ocean monitoring to hazard assessment and natural resource management, naming only a few.

Aerial photography was first practiced over Paris in 1858, by the French balloonist Gaspard-Félix Tournachon. Since then, its practice has been refined and it is still widely used for mapping purposes all around the globe. However, despite the apearance of new platforms such as drones or unmanned aerial vehicles (UAVs), subsequent scans over long periods of time can be still expensive or even unfeasable in some circumstances.

In October of 1957, the Soviet Union was able to successfully launch and place a Man-made satellite in orbit, which sparked the beginning of the Space Age. In the subsequent decades, technological advances allowed for the creation of satellites with remote sensing capabilities, pioneering the wide network of optical and radar satellites we have today. Despite the higher initial costs, space imagery can be used to effectively collect land cover data over long periods of time, assuming the orbit was carefully plotted for that purpose. A single satellite platform can harbour a great variety of powerful sensors and can work in combination with other satellites in the same orbit, mitigating some of the shortcomings that a longer orbital period might bring.

2.2 Satellite imagery

In recent years, vast investments and technological advances in the space industry were responsible for a widespread availability of satellite provided data and RS visual information (Shaohong, Xiaocong, and Qian 2014). These platforms have become especially important when working with mapping or land cover classification applications (Silverio and Jaquet 2005)(Takahashi et al. 2013)(Hashimoto et al. 2013) and the open source nature and high sensing frequency provided by some missions allow for easy collection of land cover image data. This data can then be integrated in automated systems capable of various forms of mapping and classification.

One problem that can arise comes from the fact that satellite land coverage is still not quite homogeneous around the globe. While some areas may have higher availability of high-resolution satellite imagery, with very high geometric precision, some regions may still lack these image data resources. Lately, with the increase in technological developments, areas affected by this problem are decreasing in number (Long et al. 2016) so it can be predicted that, in a not so distant future, these occurrences will be greatly reduced.

Despite many of the great benefits of spaceborn imagery, some elements may turn to be disadvantageous, depending on the circumstances. High altitude observation is a great way to monitor large scale phenomena such as global atmospheric developments

however, the same atmospheric conditions may hamper land cover observation, something that could be mitigated or outright avoided with a lower altitude sensing platform.

Nevertheless, despite some of its shortcomings it has become irrevocable that satellite imagery, whether acquired by passive or active means, has become a key part of the Information and Space Ages, creating an unbreakable synergy between the two. The missions launched by numerous countries and federations have marked a time where the connection between space and the civilian populace is at an all-time high, showing no signs of stagnation. Projects like Copernicus by the European Commission in partnership the European Space Agency (ESA) and Landsat by the National Aeronautics and Space Administration (NASA) provide helpful information that can be used for a number of applications, ranging from curiosity driven and didactic use to innovative research and technological development.

2.2.1 Copernicus Programme

The Copernicus Programme was established in 2014, succeeding the European Union's previous Earth observation project known as Global Monitoring for Environment Security (GMES).

Composed of seven *Sentinel* missions, each geared towards a specific function, the Copernicus initiative aims to provide a high-quality autonomous system capable of continuous wide range Earth observation:

- **Sentinel-1:** A constellation of two polar-orbiting satellites performing C-band synthetic aperture radar imaging over land and sea, regardless of weather conditions and time of day;
- **Sentinel-2:** A constellation of two polar-orbiting satellites placed under a sun-synchronous orbit that capture high-resolution optical imagery;
- **Sentinel-3:** A satellite responsible for measurements of sea surface topography, sea and land surface temperature and colour with high accuracy;
- **Sentinel-4:** Geostationary satellite that will provide hourly data on tropospheric constituents over Europe;
- **Sentinel-5P:** The first satellite from the Copernicus mission to be fully dedicated to atmosphere monitoring. Performs atmospheric measurements with high spatio-temporal resolutions;
- **Sentinel-5:** Satellite that will be focussed on air quality composition with daily global coverage;
- **Sentinel-6:** Satellite that will perform altimetry measurements.

As a whole, this network of satellites can be used to generate large amounts of data to help improve environmental management and understanding. The services provided by the Copernicus Programme reflect this nature by focussing on the monitoring and forecasting of the Earth's subsystems. The six main services can be categorised under the following:

- Land Monitoring;
- Marine Environment Monitoring;
- Atmosphere Monitoring;
- Climate Change;
- Emergency Management Service;
- Security.

2.2.1.1 Sentinel-2

According to the data sheet provided by ESA (https://sentinel.esa.int), Sentinel-2, from Copernicus Programme, is a wide-swath, high-resolution, multi-spectral imaging mission. The project started in June of 2015 with the launch of the first satellite (Sentinel-2A), designed with a 7-year lifespan. Later, in March of 2017, the second satellite (Sentinel-2B) was launched on a polar-orbit to that of Sentinel-2A, finalizing the constellation.

Composed by the two twin satellites (Sentinel-2A and Sentinel-2B) flying at the same orbit (786 km mean altitude) and phased by 180 degrees, this mission provides high quality imagery coupled with several interesting tools and features:

Mission profile

- Temporal resolution of 5 days between the two satellites. However, some regions may be seen twice or more in the same timespan (under different viewing angles). This welcoming phenomenon is caused by the presence of swath overlap;
- Land coverage ranging from -56º to +86º degrees latitude (including the whole of the Mediterranean Sea);
- On-board radiometric calibration.

Multi spectral instrument (MSI)

- A total of 13 bands ranging in the visible, infrared, near infrared and shortwave infrared part of the spectrum (443 nm – 2190 nm);
- Spectral resolutions ranging from 15 nm to 180 nm;
- Spatial resolution of 10, 20 and 60 meters;
- Radiometric resolution / accuracy of 12 bit / < 5%;
- A swath of 290 km.

These specifications, coupled together with the open source nature of the data provided, make of this platform an excellent data source candidate for land cover classification studies and applications. The vast geographical cover and fast refresh rate allows for a multitude of uses, such as agricultural monitoring, emergency mapping, and so on.

In terms of geometric distortion, Sentinel-2 is able to provide relatively high-quality imagery, as can be seen in the specifications provided by ESA:

Geometric performance

- A *priory* absolute geo-location uncertainty: 2 km 3σ;
- Absolute geolocation uncertainty: 20 m 2σ without GCPs and 12.5 m 2σ with GCPs;
- Multi-temporal registration: 0.3-pixel 2σ including compensation for the effects of terrain height variation.

Without further geo-rectification applications requiring high accuracy and high precision imagery may find Sentinel-2 geometric data to be somewhat lacking.

2.2.2 Landsat Program

The Earth Resources Technology Satellite was launched in 1972, as a joint effort between NASA and the United States Geological Survey (USGS), initiating what was later known as the Landsat program. Between the launch of the first satellite, which was later renamed to Landsat 1, and the time of writing, the Landsat program has been composed of 8 separate missions named Landsat 1 through 8, respectively (https://landsat.gsfc.nasa.gov). The mission's chronology can be seen in Figure 2.1.

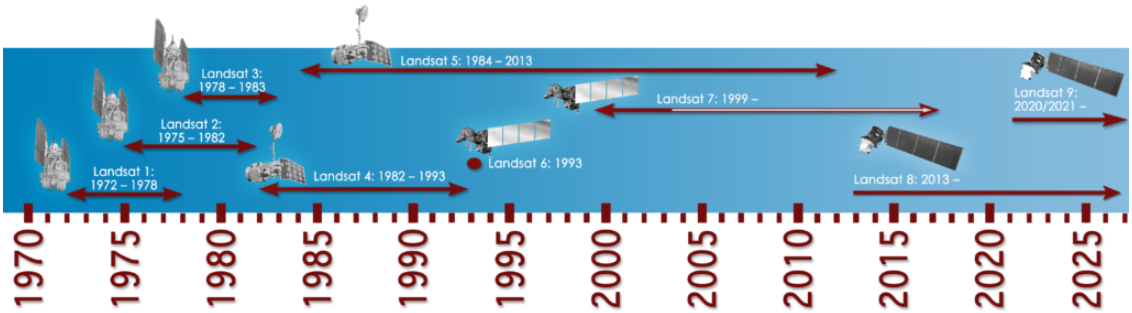


Figure 2.1: Landsat Program chronology, taken from (https://landsat.gsfc.nasa.gov).

The main focus of this project was to create a platform capable of space-based, continuous capture and recording of the Earth's land cover. These images can be freely accessed through the USGS "Earth Explorer" website and have been extensively used over the years in the fields of:

- Cartography;
- Agriculture;
- Geology;
- Foresting;
- Regional planning;
- Surveillance;
- Education.

Currently only two of the missions remain active, these being Landsat 7 and 8, with Landsat 9's predicted launch in 2021. Together they are equipped with an Enhanced Thematic Mapper Plus (ETM+), that replicates the capabilities of the past Thematic Mappers of Landsat missions 4 and 5, an Operational Land Imager (OLI) and a Thermal Infrared Sensor (TIRS), that provide seasonal coverage of the global landmass.

2.2.2.1 Landsat 8

The eighth satellite in the Landsat Project was launched on February 11, 2013 as a coordinated effort between NASA and the USGS. It was developed to have a 5-year life span and carries a two-sensor payload composed of an OLI and a TIRS.

Mission objectives

- Collecting and archiving multi-spectral imagery of global landmasses while providing seasonal coverage of 16 days;
- Ensuring an output of data compatible with previous Landsat missions to facilitate the study of landcover and land-use changes over time;
- Distribution of Landsat 8 data products on a free, non-discriminatory basis to the general public.

Specifications

- Two longwave thermal bands (10.30 μm – 12.50 μm) with a 41-meter circular error, 90-percent confidence, and a designed 3-year life span;

- A total of 9 shortwave bands (0.433 µm – 1.390 µm) with a 12-meter circular error, 90-percent confidence;
- Spatial resolution of 15, 30 and 100-meters for panchromatic, multi-spectral and thermal imagery, respectively;
- Temporal resolution of 16 days and a capture size of 185x180 km².

Despite a slower revisit time when compared to Copernicus's Sentinel-2 mission, this platform manages to hold on to its relevancy by continuously providing data to many projects worldwide. However, in applications that require high data refresh rate, coupled with the highest possible spatial resolution, Landsat 8 may not be the best choice for a data acquisition platform.

2.2.3 Other Earth observation satellites

Besides the previously mentioned Programs, many other endeavors have been devoted to the creation of Earth observation platforms. Missions like the Jason-3, which started as a partnership between NASA, the European Organization for the Exploration of Meteorological Satellites (EMETSAT), the National Oceanic and Atmospheric Organization (NOAA) and the *Centre National d'Etudes Spatiales* (CNES), provide detailed information regarding global sea surface height, a critical factor in understanding the climate change phenomenon. The Tropical Rainfall Measuring Mission (TRMM), started as a cooperation between the Japan Aerospace Exploration Agency (JAXA) and NASA, was designed to study and monitor tropical rainfall, providing additional insight in the fields of global energy and water cycles. CloudSat, from NASA's CloudSat mission, focusses on the use of radar to measure the altitude and properties of clouds, providing additional information about the Earth's climate system.

CloudSat was part of the "A-train" satellite constellation composed now of four multi-national Earth observation satellites, these being Aura, Aqua and OCO-2 from NASA and GCOM-W1 from JAXA. Collectively, the information gathered by this constellation can create high definition three-dimensional imagery of the Earth atmosphere and surface, once again providing invaluable data on the subject of climate change.

Many more Earth observation satellites exist in orbit or are planned to be launch in the near future however, most will go without mention during the course of this dissertation for the sake of brevity.

2.2.4 Earth observation products

Most Earth observation missions offer a variety of different product levels based on the amount of processing applied over the sensed data. In an effort to help standardize the different classes of products, the Committee on Earth Observation Satellites (CEOS), responsible for coordinating and providing easy access to Earth observation data, proposed the creation of five distinct product level categories:

- **Level 0:** Represents products with the least amount of processing. Composed of unprocessed sensor information, reconstructed at full resolution, with all communication artefacts and the metadata required for further processing;
- **Level 1:** Based on unpacked and reformatted Level 0 information, this category of data is further enhanced by the inclusion of time-referencing and may optionally incorporate additional radiometric and geometric calibration coefficients, along with georeferencing parameters to allow representation in physical units. Metadata is appended for subsequent data processing;
- **Level 2:** Addition of retrieved environmental variables such as ocean height, soil moisture and so forth, at source data resolution. Pixel processing is tied to its nature, resulting in different treatment depending on whether the pixel represents water or clouds, as an example;
- **Level 3:** Products in this category are created from multi-temporal resampled data, obtained from Level 1 or Level 2 imagery. The resampling process may include averaging and compositing;
- **Level 4:** The highest product level defined by CEOS is a model generated by the analysis of lower level data, being solely composed of indirect measurements.

2.2.4.1 Sentinel-2 data products

Like most Earth observation missions, Sentinel-2 provides a variety of different products. By referring to the previously presented standard, ESA categorized the products provided under the following classes:

- **Level-1A:** Obtained by decompressing raw data from Level 0 imagery. Includes a geometric model which allows any pixel in the image to be located;

- **Level-1B:** Provides in sensor geometry and radiometrically corrected imagery in Top Of Atmosphere (TOA) radiance values, compressed by the JPEG2000 (.jp2) algorithm. Includes a refined geometric model to be used in subsequent processing operations;
- **Level-1C:** Composed of 100x100 km² tiles, this product includes geometric and radiometric corrections which include ortho-rectification in a UTM/WGS84 projection. Radiometric measurements per-pixel are provided in TOA reflectances, including the parameters required to transform them into radiances with sub-pixel precision. Depending on the native resolution of the spectral band these products can either have a spatial resolution of 10, 20 or 60 meters. In addition, cloud, water and land masks are included;
- **Level-2A:** Composed of 100x100 km² tiles, this product offers Bottom Of Atmosphere (BOA) reflectance derived from Level-1C imagery by correcting some of the atmospheric noise. In addition, it provides Aerosol Optical Thickness (AOT) and Scene-average Water Vapour map (SCL) bands;
- **Level-3A:** Monthly compound derived of cloudless Level-2A imagery.

2.2.4.2 Landsat 8 data products

The eighth mission of the Landsat project also follows the norm by producing a diverse range of data products, classified by NASA under the following groups:

- **Level-1TP:** Representing the highest quality Level-1 data products provided by the mission, it provides radiometric and geometric accuracy by incorporating GCPs and employing a DEM for topographic displacement;
- **Level-1GT:** Following the previous product in terms of geometric quality, it provides systematic, radiometric and geometric accuracy by employing a DEM to correct relief displacement;
- **Level-1GS:** Representing the lowest Level-1 data product, it provides systematic radiometric and geometric correction derived from data collected by the platform and sensors;
- **Surface Reflectance:** Akin to Level-2 data products, provides information concerning the solar radiation reflected by the surface of the Earth;
- **Provisional Surface Temperature:** Also homologous to Level-2 data products, provides information regarding surface temperature of the Earth;

- **Dynamic Surface Water Extent:** Provides a raster layer representing surface water inundation per-pixel, akin to Level-3 data products;
- **Fractional Snow Covered Area:** Informs about the percentage of snow present in any given area represented by a single pixel. Comparable to Level-3 data products, with a 30-meter spatial resolution;
- **Burned Area:** Designed to identify burned areas in all possible ecosystems. Contains two rasters representing burn classification and burn probability. Akin to Level-3 data products, with a 30-meter spatial resolution.

2.3 Geographical Information Systems

Data analysis is a critical element in every scientific field, with multiple facets and a vast spectrum of approaches. Visualization is often employed as a strategy to accelerate the process of manual data analytics, commonly manifested as graphs, maps and other well-known visual interpretations of data.

A framework created to gather, manage, analyse and present spatial or geographical data is known as a Geographical Information System (GIS). The geospatial perspective obtained by merging data and geographical elements allows for easier interpretation and manipulation of information by human users, turning more abstract concepts such as a pixel or coordinate table into raster or vector representation. Raster data is represented in bitmap form, consisting of an image created by individual pixels with a set value, usually representative of a physical concept such as reflectance or height. Vector data is either represented by points, lines or polygons and is usually used to store discrete data such as country borders, parcels or roads.

The use of GIS software can help expedite many spatial analysis applications in environmental, military and governmental fields, among others. Climate change, natural catastrophes, regional planning and population dynamics are only a few of the areas that can take advantage of the relation between data and spatial attributes provided by GIS software.

2.3.1 ArcGIS

Designed by the Environmental Systems Research Institute, ArcGIS is a market leading professional GIS software for governmental, industrial and commercial use.

The software is divided into multiple applications, with ArcMap being considered the central component. This core application was primarily designed as mapping and

analytics platform, providing users with multiple options in terms of analysis and visualization of spatial data. Besides ArcMap, the ArcCatalog component provides a catalog window with a file explorer function capable of managing and editing GIS data under various formats. ArcGlobe and ArcScene instruments grant this software a distinct advantage over the vast majority of competitors by allowing a three-dimensional representation of the Earth.

The ArcToolbox also contains a vast collection of geoprocessing tools which can be used in a variety of tasks such as:

- **3D analysis:** Surface modelling, management and analysis utilities based on three-dimensional data;
- **Geocoding:** Incorporation of places and addresses into GIS data by conversion into a variety of coordinate systems;
- **Cartography:** Analysis and construction of maps.

Licensing for ArcGIS comes in three different tiers, these being basic, standard and advanced, with only the latter providing access to all of the GIS functionalities.

2.3.2 QGIS

Built around the core principal of Free and Open Source Software (FOSS), QGIS is a professional cross-platform GIS application, started by Alaskan Sherman in 2002.

This software supports both raster and vector layers, allowing the user to create, import and export multiple raster and vector formats such as .jpg, .png, .shp and .tif, naming only a few. An explorer functionality is also present allowing users to easily acquire and manage any GIS data they might require. At a processing level, hundreds of tools are immediately available allowing the user to perform a multitude of tasks including:

- **Raster:** Creating, managing and manipulating raster graphics;
- **Vector:** Creating and editing vectors, points, lines and polygons;
- **Raster-vector:** Converting data types;
- **Images:** Manipulating, calibrating, analyzing and segmenting images.

Furthermore, the focus on user created content led to the proliferation of numerous plugins that can be downloaded and incorporated into the main software, adding to QGIS's inherent capabilities.

3

State of the Art

This section will present the literature that served as a base for this dissertation. As stated in the first chapter, the proposed system aims to automatize the first stages of land cover classification, these being the geometric rectification (geo-reference) and cloud-cover correction aspects of satellite image data processing.

Throughout the investigative work, existing solutions were gathered, analysed and understood, fleshing out their pros and cons in relation to the problem proposed by this dissertation.

3.1 Geo-rectification

Working with RS imagery often demands some level of pre-processing. This is required due to the fact that, in their initial state, images tend to suffer from the presence of geo-rectification errors, preventing them from becoming fully useful samples.

Satellite imagery is no different. Despite the many technological advancements seen in the last decade, as of now, even the most advanced state-of-the-art satellite geo-positioning equipment is prone to errors, which may lead to inaccurate geo-location (spanning from several to hundreds of meters) on the ground (Long et al. 2016). In addition, given the innate conditions associated with satellite RS imagery, these images may also become distorted. The most common causes for this phenomenon can be traced back to projection, tilt angle, curvature of the Earth, atmospheric conditions, undulation, and so on, at the time of sensing (Nguyen 2017).

Using these RS images for geospatial applications such as generation of orthoimages, creation of Digital Elevation Models (DEMs), image fusion, change detection, and so on requires them to be geo-rectified. To put it simply, the geo-rectification process

can be described as sort of coordinate transformation between the image and the ground coordinate system (Topan and Kutoglu 2009), as exemplified in Figure 3.1.

It is important to keep in mind that, even if well classified, poorly geo-rectified image samples will cause misleading results in automated monitoring systems, since two different geographical locations may be targeted for comparison.

(a) **(b)**

Figure 3.1: Aerial photography (a) and digital orthophoto (b), taken from (http://gsp.humboldt.edu).

Image acquisition technology has seen a technological boom in the past decades, resulting in multiple, new and diverse, sources of image data. With this growth, together with vast new amounts of data requiring attention, came the need to research and develop new systems for automatic image registration.

Geo-rectification has been a topic of study for many years and it manages to remain a relevant field up to this day. However, despite this near constant attentiveness, no universal method capable dealing with the full range of geo-rectification problems has been presented.

Problems related to geo-rectification, or image registration, cover a very large spectrum. This spectrum includes a vast diversity of sensors and RS imagery, coupled together with many different degrees of image degradation. Because of this, the prospect of creating an effective universal system, capable of dealing with all forms of image geo-rectification, is borderline unrealistic (Flusser and Zitova 2003).

Each method should consider not only the nature of the RS imagery, coupled with the predicted type of geometric distortion, but also the amount of radiometric distortion

and noise corruption present between the images. The accuracy requirement is also a measure worth accounting for, given that certain projects require different levels of accuracy. Such planning may lead to the use of less taxing methods, improving the speed and overall efficiency of the solution.

As previously stated, each individual solution is generally balanced for a single purpose and designed in a way that suits the problem at hand. In this section, the most common geo-rectification, or image registration, methods will be presented and studied.

Lastly, these methods will be analysed side by side with the problem proposed by this dissertation and go through a selection process. By the end of this analysis, methods that remain relevant will be adapted to fit the requirements of this dissertation or be used as corner stones for the development of a brand-new methodology.

3.1.1 Ground control points

A ground control point (GCP) can be described as a physical location marked with a highly accurate GPS coordinate, Figure 3.2.

Often, exterior orientation parameters lack the accuracy to produce raw imagery capable of meeting the required level of geometric precision. Because of this, GCPs are widely used when it comes to the correction of the geometric biases found in raw RS imagery (Long et al. 2016)(Nguyen 2017)(Oniga, Breaban, and Statescu 2018).

It is important to note that, given their purpose, GCPs tend to be quite distinct in nature, making them easily identifiable by human or automated checks (Pan, Tao, and Zou 2016). When working with multi-temporal land cover analysis, these locations should also remain relatively unchanged to provide good reference points throughout the entirety of the study.

The efficiency of this method relies heavily on the quality, number and distribution of the GCPs. To accurately geo-rectify an image, a minimum of three GCPs are required, but additional GCPs will increase the overall accuracy of the endeavour, assuming they are properly placed and utilized. It is important to note that, despite the increase in accuracy with an increase in the number of GCPs, there is an optimal number of GCPs that should be used for each specific problem. Exceeding this number of GCPs is prone to reduce efficiency, both in the field and computationally (Oniga, Breaban, and Statescu 2018).

Proper selection of GCPs becomes an essential step in assuring positive results in image rectification. It is important to keep in mind that this selection must always occur before any coordinate transformation functions can be applied.

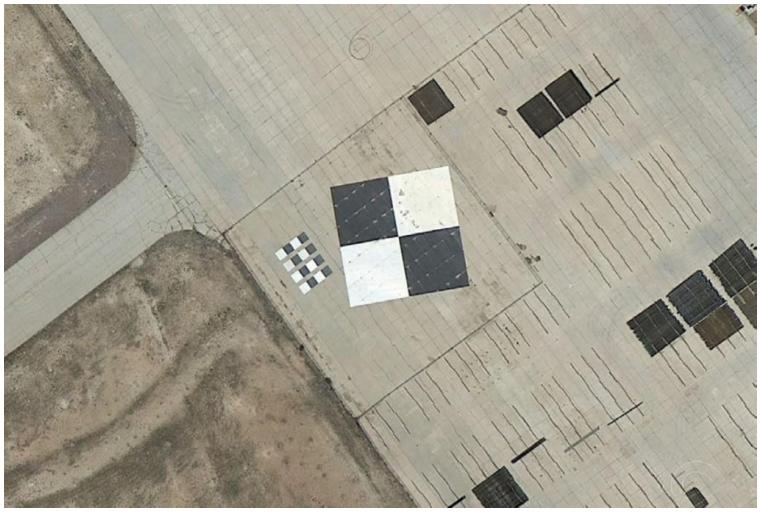


Figure 3.2: Ground control point in Big Spring, Texas, taken from (http://apogeospatial.com).

As stated above, for this method to be successful and viable, GCPs must be numerous and well spread. However, this becomes a problem when considering that the acquisition of plausible and effective GCPs may be a time-consuming and taxing effort. Some places may also be inhospitable to the acquisition of GCPs, adding to the number of problems that GCP selection already presents (Nguyen 2017).

Despite these difficulties, this method remains one of the most widely used when it comes to the geo-rectification of RS imagery. Usually, GCP studies focus mainly on the effects of the disposition and number of GCPs, on the development of coordinate transformation functions and on automatic gathering and selection of GCPs.

In this section, the most common methods for GCP acquisition and exploitation will be presented and analysed, culminating in a side by side analysis with the problem proposed by this dissertation.

3.1.1.1 Manual GCP selection

Despite being useful in some situations, more often than not, the amount of data required by most applications has become so large that these methods end up being, in a best-case scenario, ineffective and in a worst-case scenario, borderline unfeasible.

Manual GCP selection requires the use of staff or operators to collect geographical data from obvious and distinct features in the landscape. GCPs can be collected in the field, with the help of precise GPS equipment, by taking advantage of already existing and marked GCPs or by analysing existing RS imagery or GIS data. Thanks to the availability of open source online platforms, it is also possible to manually write down coordinates by accessing some of these applications. Such is the case with Landsat 8 Ground Control Point Search, a platform that provides tools to help with the selection of GCPs for Landsat 8 Level-1 data, using an online map service provided by Google (P. Li et al. 2017).

Given the nature of this dissertation, manual methods will not be studied further, for they do not provide the necessary level of automation required.

3.1.1.2 Automatic GPC acquisition

As previously stated, technological development has led to a considerable proliferation of RS imagery. For this reason, automated systems have been studied and developed to cope with the vast amounts of data provided by these remote sensing platforms.

Automatic GCP acquisition has been thoroughly researched with great success, showing very positive and reassuring results. However, it is important to keep in mind that due to the sheer scale and variety of RS imagery, coupled together with the noise and distortion associated with said imagery, a single universal solution that can encompass and solve the whole spectrum of problems related to GCP acquisition is very unlikely.

Automatic GCP selection is heavily dependent on image matching. Succinctly, image matching can be summed up as a process that establishes a correspondence between two homologous images, usually by using a set of features common to both samples. Detection of these features is a process that usually follows two major approaches: Area-based matching (ABM) and feature-based matching (FBM).

The structure of the following subchapters relating to ABM and FMB was borrowed from the survey (Flusser and Zitova 2003).

3.1.1.2.1 Area-Based matching

Methods that rely on ABM process image samples without any attempt to detect discernible objects or features. Correspondence estimations are generally carried out by using windows of predefined size or by analysing the image in its entirety, Figure 3.3. This idea, however, acts as an intrinsic limitation for ABM methods. The frequent use of

rectangular windows means this method if often better suited to register images that suffer only from a local translation.

If the imagery is affected by more complex transformations, rectangular windows may prove ineffective, since they might be unable to cover the same areas in the reference and sample image, as the rectangle can be transformed into other shapes. The use of circular windows has been proposed by some authors for images with mutual rotation. However, if the geometric deformations are revealed to be of a higher complexity, as can be seen with perspective transformations, for example, the circular window will also suffer in terms of comparability.

One other disadvantage of ABM methods is related to the possible 'remarkableness' of the sample selected by the window. When areas without any discernible details are encompassed by a window, there is a high chance that they will be matched incorrectly with other similar areas found in the reference image, due to their non-saliency.

Generally, classical ABM methods, like cross-correlation (CC), match images directly by taking their intensity into account, making no effort to analyse any structures whatsoever. Therefore, these methods end up being sensitive to changes in image intensity. Such changes can be the result of noise or variations in lighting. The use of different sensor may also create discrepancies in terms of sample intensity (Flusser and Zitova 2003).

Cross-correlation. The classical representatives of ABM methods are the normalized CC and its many modifications. These methods work by taking window pairs from the reference and sample images and then computing a similarity measure. In cases where subpixel accuracy is a requirement for registration, the interpolation of CC measure values is necessary.

Even though registration based in CC is generally only effective at dealing with mutually translated images, it can also be applied with some success when scaling and slight rotation are present.

Some generalized forms of CC exist to deal with images affected by greater distortions. These methods compute the CC for all assumed geometric transformations in the sample window. Because of this, these methods can deal with even more complex geometric deformations in comparison to the usual translation similarity transform. It is worth to note that, in terms of computing, these methods can become very taxing with slight increases in transformation complexity.

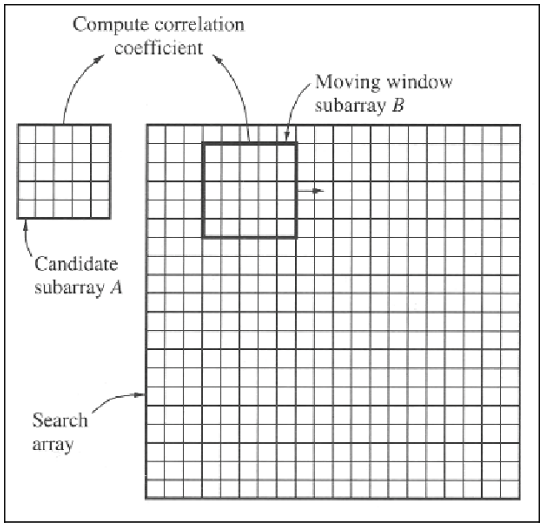


Figure 3.3: General methodology for ABM, taken from (Wolf and DeWitt 2000).

Sequential similarity detection algorithms (SSDA) are similar to CC methods. Employing a sequential search approach and a simpler distance (in terms of computation), they apply the threshold criterion after accumulating the sum of absolute differences in the image. If this accumulated sum exceeds the defined threshold, the pair of windows under analysis is rejected and the next pair is then tested. Given the way this method works, it is likely to be less accurate than the previously mentioned CC, but it should prove to be somewhat faster (Flusser and Zitova 2003).

Huttenlocher et al. (Huttenlocher, Klanderman, and Rucklidge 1993) proposed a method that registered binary images (transformed by rotation and translation or simply translation) by using the Hausdorff distance. This algorithm was compared with CC and proved to outperform CC in images with perturbed pixel locations, something that tends to be problematic when employing CC methods.

The main drawback of CC methods is the high computational complexity for some of the transformations. There is also the fact that the maxima of the similarity method have a tendency to be rather flat, due to the intrinsic similarity of the images. To solve this issue, pre-processing can be applied to sharpen the maxima. This can also be achieved by using edge or vector correlation. A method that applied image filtering before the registration process was use by Pratt (Pratt 1974) and managed to improve CC performance in images with noise or a high correlation.

Despite having the previously stated limitations, thanks to their easy hardware implementation CC registration methods are still widely used, especially in real-time applications (Flusser and Zitova 2003).

Fourier methods. These types of methods handle images by representing them in the Fourier frequency domain. It is advantageous to process images corrupted by frequency-dependant noise with Fourier based methods, in opposition to CC based methods. Regarding computational speed there is also an advantage that becomes more apparent as the sample image size increases.

Phase correlation is based on the Fourier Shift Theorem (Bracewell 1965) and it was originally intended to register translated images. It works by computing the cross-power spectrum of both images and by searching for the location of the peak in its inverse. This method reveals itself to be robust against frequency dependent and correlated noise as well as non-uniform time varying lighting discrepancies.

For cases where image rotation is also present, extensions of this method have been developed, as can be seen by the methodology proposed by De Castro and Morandi (Castro and Morandi 1987). When differences in image scale are also a factor to consider, images can be registered by means of a combination between polar-log mapping of the spectral magnitude, corresponding to the Fourier-Mellin transform, and the phase correlation (Srinivasa and Chatterji 1996) (Chen, Defrise, and Deconinck 1994). Applications of this extended algorithm have been seen in RS, used with SPOT images and in medical imaging, using MR images.

Other applications that take advantage of the Fourier transform have also been proposed. One computed the correlation in the frequency domain and allowed for the handling of multimodal images when applied to edge representations, as opposed to the original grey level images (Anuta 1970). Foroosh et al. (Shekarforoush et al. 2002) proposed the extension of phase correlation to subpixel registration by applying, on down sampled images, the use of the analytic expression of the pulse correlation. To reduce the computational cost, a technique focused on the use of a local discrete Fourier Transform was proposed in (Guizar-sicairos, Thurman, and Fienup 2008)

Mutual information methods. These methods represent the leading technique in multimodal registration and were originated from information theory. Images with different modalities are particularly suited for this kind of registration.

One of the first implementations of this methodology can be seen in Viola and Wells (Viola and Wells 1997). Thevenin and Unser used various approaches at every step to try and enhance mutual information (MI) registration. For the joint probability computation, they employed the Parzen window and, to maximize the MI, they employed the

Jeeves or the Marquardt-Levenberg methods. To increase the computational speed, spline pyramids were used (Unser 1996)(Thkvenaz and Unser 1997)(Thtvenaz and Unser 1998). To search for the maximum MI, Ritter et al. (Ritter et al. 1999) developed a method that used a hierarchical search strategy, coupled with simulated annealing. It is important to note that all the previously presented methods work directly with image intensities and use the entirety of the image data (Flusser and Zitova 2003).

Optimization methods. Searching for the maximum or minimum similarity measures is an optimization problem that is heavily dependent on the expected geometrical transformation. The only way to obtain a global extreme solution is to perform an exhaustive search over the entirety of the image. Even if computationally demanding, this method is often used when only translations are to be expected.

In cases where geometrical transformations are more complex, algorithm optimization needs to become more sophisticated. One of the solutions is to apply a Gauss-Newton numerical minimization algorithm to minimize the sum of squares difference. Such a method is described in (Sharma and Pavel 1997). A system using the gradient descent optimization method was used in (Viola and Wells 1997) and (Sawhney and Kumar 1999). To minimize the variance in corresponding pixels, the Levenberg-Marquardt optimization method was applied.

One more method was introduced in (Jenkinson and Smith 2001). It had its applicability proven by MI and correlation ratio. Same as before, the optimization methods can see an increase in computational speed by employing the pyramidal approach (Flusser and Zitova 2003).

Summary. To summarize, ABM methods should be applied when the images show a lack of prominent features and details. In other words, cases where the images distinctive features can be found in the colours or graylevels, instead of in the formation of local structures and shapes. These methods are limited by the need for the sample and reference images to share close intensity functions. These must either be identical, in which case CC methods can be applied, or at least be statistically dependant, something that generally occurs in multimodal registration.

In terms of geometric distortion, it is recommended to keep the use of ABM methods restricted to situations where just translations or small rotations are to be expected. It is possible to use ABM methods to deal with full rotation and scaling, however, these

algorithms become extremely taxing in terms of computation. For this reason, such methods tend to remain unused.

Generally, to speed up the searching process, ABM methods employ pyramidal image representations, coupled with complex optimization algorithms to find the maximum of the similarity matrix (Flusser and Zitova 2003).

3.1.1.2.2 Feature-based matching

Like the name implies, FBM relies heavily on the extraction of noticeable structures, or features, present in the imagery. In this context, features can be described as any distinguishable regions (forest, lakes, buildings), lines (rivers, roads, coastlines) or points (region corners, intersections). As previously mentioned, these features should be quite distinct and detectable in nature. They should also be present and spread evenly across the samples and be expected to remain relatively unchanged throughout the entire duration of the study.

In opposition to ABM, methods that employ a feature-based approach have the advantage of not working directly with image intensity values. Since the features used in FBM are a higher-level interpretation of information, these methods become fitting for situations that demand multi-sensor analysis or are expected to have changes in illumination, like multi-temporal land cover classification.

Despite the nuances found in each individual solution, FBM consists of four main steps: The first step is defined by the extraction of features from the sample image. In the second step, feature matching takes place. The third step is where the transformation model is estimated. Finally, in the last step, the sample image is transformed (Huang and Li 2010). A depiction of this rectification process can be seen in Figure 3.4.

Region features. These features are generally composed of high-contrast close-boundary regions that bear a significant size. Often, they are represented by their centres of gravity. This proved to be an effective method, since centres of gravity have shown to be invariant regarding image rotation, scaling and skewing. They also managed to remain stable under random noise and grey level variation.

Detecting region features requires the use of segmentation methods. As a result, the effectiveness of the image registration process is tightly bound to the accuracy of these operations.

Refinements to the segmentation method have been proposed, namely by Goshtasby et al. (Goshtasby and Stockman 1986), to increase the quality of image registration. By iterating the segmentation of the image together with the registration, a rough estimate of the object correspondence can be used to tune the parameters at every iteration. This method claims that sub-pixel accuracy for the registration can be achieved (Flusser and Zitova 2003).

Line features. Composed of line segments, these features are often extracted form coastal lines, object contours, roads and so on. Generally, line correspondence is obtained by pairs of line ends or middle points. Detecting line features is usually done by employing standard edge detection methods. These can be classified into two categories: Gradient-based methods and Laplacian-based methods.

In gradient-based methods, edge detection is done by taking the first order derivative of the sample. The gradient magnitude is then utilized to calculate a measure capable of defining edge strength. Laplacian-based methods utilize the second order derivative expression, computed from the sample. The edges are then detected by searching for the zero crossings of a non-linear differential expression. As a refinement for this kind of edge detection, having a Gaussian smoothing pre-processing step is not unusual (Syst and Mutneja 2015).

Many methods have been developed and have shown great practical results. Such is the case with the Canny method (Canny 1986) and Laplacian of Gaussian (Marr and Hildreth 1980). Fuzzy logic can also be utilized and has shown very positive results, as can be seen by Shashank Mathur and Anil Ahlawat (Mathur and Ahlawat 2008). These results can also be extended to colour format images, as stated in (Gonzalez and Melin 2017). Edge detection based on fuzzy logic has the added advantage of having fuzzy rules and parameters that can be added, changed and tuned to affect edge thickness, creating a flexible structure using relatively low complexity (Syst and Mutneja 2015).

Point features. These features are obtained by methods that work with corners, road crossings, high variance points and so on. Generally, the definition of a point in the most prevalent algorithms can be summed up as: an intersection between two lines, the geometric centre of a closed-boundary region or local modulus maxima of the wavelet transform. Corners are easily comprehended by human observers but are much harder to define mathematically. For this reason, corners are usually considered specific class features. A multitude of methodologies have been developed over the years in order to solve

the problem of corner detection. This comes as no surprise, seeing that corners, when used as control points, show a promising level of invariance regarding imaging geometry. Some of the most prominent approaches are, Dreschler and Nagel (Dreschler and Nagel 1981), Kitchen and Rosenfeld (H. Wang and Shen, n.d.), Förstner (Forstner and Gulch 1986), Harris (Noble 1988).

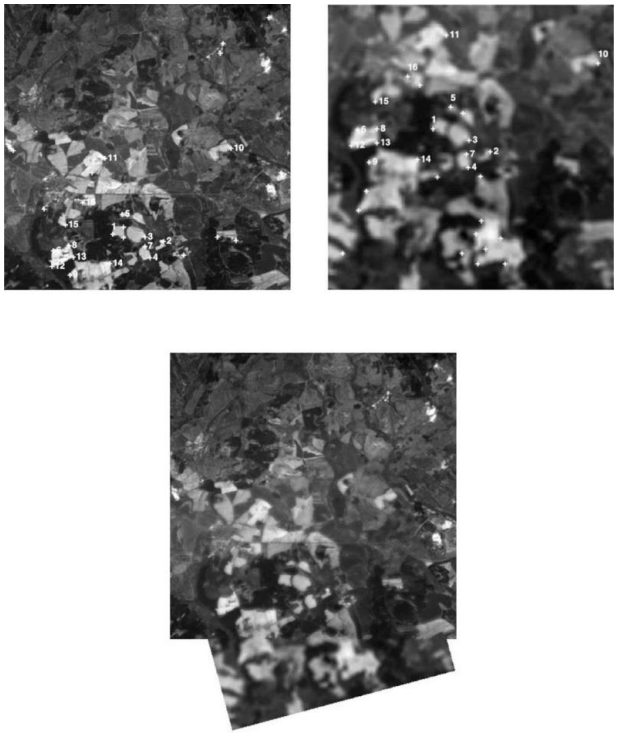


Figure 3.4: GCP acquisition based on FBM and subsequent image rectification, take from (Flusser and Zitova 2003).

Kitchen and Rosenfeld take advantage of the second-order partial derivatives of the image function to achieve corner detection while Dreschler and Nagel use the Gaussian curvature extreme. It is important to keep in mind that corner detectors based on the second-order derivatives of the image functions have shown to be susceptible to noise interference.

The method developed by Förstner uses only first-order derivatives and proves to be more robust, however, this comes at a price, since this brand of corner detection reveals to be more time consuming in relation to the previous methods. The Harrys detector, also known as Plessey detector, is also widely used and has been a cornerstone in the development new methods. It has also been the target of many improvements over the years.

In (Systeme, Informatik, and Hamburg 1994), most of the previously mentioned approaches are tested and compared.

Smith and Brady (Smith and Brady 1995), developed the SUSAN method that proposed using the colour of the central pixel to designate an area size. Trajkovic and Hedley created a method based on the notion that the intensity seen at the corners should be high in all given directions (Hedley 1998). Some methods proposed refrain from using any derivatives and are designed to process blurred and noisy data. Such is the case with Recently (Zitov et al. 1999).

It is important to note that the number of detected points can become quite large, resulting in increased computational work that, in turn, culminates in an overall slower registration process (Flusser and Zitova 2003).

Spatial relations methods. Such methods are based on the spatial relations between features and are generally employed when these suffer from ambiguity or if their neighbours are locally distorted. In other words, these methods take advantage of the distance between CPs, exploiting their spatial distribution.

Goshtasby in (Goshtasby and Stockman 1985) described a registration method based on a graph matching algorithm. After a particular transformation, the number of features in the sample image that fell within a certain range of the features in the reference image were evaluated. Valid estimates were defined by the highest transformation parameter scores.

A clustering technique was proposed by Stockman et al. (Stockman, Kopstein, and Benett 1982), where points are matched by use of abstract line segments or edges. By using this method, local errors end up not having an influence in the global registration process. Chamfer matching for image registration was introduced by Barrow et al. (Barrow et al. 1928). By minimizing the generalized distance between line features detected in images, it is possible to match them. An improved version was proposed by Borgefors (Borgefors 1998), that also used the pyramidal speed up method referred previously in the ABM section (Flusser and Zitova 2003).

Invariant descriptors. Instead of taking advantage of spatial relations, like the previous methods, invariant description-based methods use, as the name implies, the description of features to correspond them.

A good description must follow several conditions. Of these conditions, the most critical ones are invariance, meaning the descriptions of corresponding features in both the sample and the reference images must be the same. Uniqueness, meaning each feature should have a unique description. Stability, meaning if a feature is slightly deformed in an unexpected manner that the description should still resemble the one from the original feature, and finally independence, meaning that if the description feature is a vector, that its elements must be functionally independent. It is not unusual, however, to come across situations where some of these conditions are not met simultaneously. It is not strictly necessary that they do, but in these circumstances, a trade-off must be achieved.

These methods work by pairing features according to their invariant description similarity. After all features in the sample and reference images have been analysed, the pairs with the biggest invariant description similarities are paired as the corresponding ones. Choosing the invariant description, however, is dependent on feature characteristics as well as the expected geometric deformation. Usually, when searching for the ideal feature match in the space of feature descriptors, the minimum distance rule with thresholding is applied. In situations where an algorithm of higher robustness is required, the matching likelihood coefficients method can prove be an adequate solution. In (Fidrich et al. 2016), a method that selects features according to the reliability of their matches is proposed.

The most straightforward way to define a feature description is with the image intensity function, limited to the close neighbourhood of the feature. To obtain feature correspondence, some authors compute the CC on these neighbourhoods. Other methods for measure similarity can be applied too. The method proposed by Zheng and Chellapa (Zheng and Chellappa 1993) utilizes correlation coefficients. In (Zitova, Flusser, and Sroubek 2002), MI methods were employed to improve feature correspondence.

The following methods can be classified as intuitive descriptors, since they generally do not fulfil some of the previously mentioned criteria for invariant descriptors. In Sester et al. (Sester, Hild, and Fritsch 1992) region features were composed of forest and these were proposed to be described by elongation parameters, number of openings, density, and several other characteristics of the minimum bounding rectangle. Vujovic and Brzakovic, in (Vujovic, Member, and Brzakovic 1997), defined every detected feature, composed of elongated structure intersections, through the signature formed from the longest structure, as well as angles, found between every other structure that took part in the intersection. In (Montesinos et al. 2000), Montesinos et al. proposed the use of differential descriptors of the image function in the neighbouring areas of detected control points.

Closed-boundary regions are often employed as features by many authors. In theory, if a shape descriptor is sufficiently invariant and discriminative, it can be used in region matching. Peli (Peli 1981) proposed a straightforward description by means of radial shape vector. It showed a fast description process, but the use of this method is restricted to star-shape regions exclusively. Shape matrices can also be used to, for example, deal with rotated and scaled satellite imagery, as can be seen in (Goshtasby and Stockman 1986). In (H. Li, Manjunath, and Mitra 1995), chain code representations of contours were proposed as invariant descriptors, while the chain correlation-like measure was used as a means to find correspondence.

Many methods have exploited moment-based invariants for the description of features defined by closed-boundaries. Taking the generally assumed deformations, Hu (Hu 1962), proposed the introduction of moment invariants to the similarity transform. Flusser and Suk derived the affine transform invariants and were able to use them for the successful registration of SPOT and Landsat images (Flusser and Suk 1993)(Flusser and Suk 1994).

Some methods relied on the invariant combination of geometric properties of features to form geometrically leaning descriptors. Wang and Chen (W. Wang and Chen 1997) performed the computation of the line-length ratios histogram and the histogram of angle differences between any two-line segments in the sampled and sensed images. Ventura et al. (Ventura, Rampini, and Schettini 1990) defined image features by utilizing various descriptors such as angle, thickness, ellipticity and so on. The relations were then represented by a multi value logical tree (MVLT). The MLVT of both images, sample and reference, were then compared to find feature correspondence (Flusser and Zitova 2003).

Point-based methods. Since detection based on region and line features is generally more complex and less accurate, point-based methods are highly utilized. Of these methods, one that stands out due to its importance, is the scale-invariant feature transform (SIFT) (Lowe 2004). This method is highly insensitive to image scale and rotation and proves to be robust across a wide variety of affine distortion, however, it is rather computationally taxing.

In recent years, some point-detectors were developed to increase computational speed, such as Speeded Up Robust Features (SURF) (Bay, Tuytelaars, and Gool 2008), Features from Accelerated Segment Test (FAST) (Rosten, Porter, and Drummond 2010), Binary Robust Invariant Scalable Keypoints (BRISK) (Leutenegger, Chli, and Siegwart, n.d.), as well as Oriented FAST and rotated BRIEF (ORB) (Rublee, Rabaud, and

Konolige 2011). All these methods can be used as a fast and efficient alternative to SIFT however, they lack the robustness provided by SIFT.

Despite its robustness, SIFT needs to overcome the following challenges when dealing with RS imagery: large sample size, large scenes, multi-source images, outliers, accuracy, and so on. Because of this, and as previously mentioned, many improvements to SIFT have been proposed in order to deal with these drawbacks. Some of these improvements relate to efficiency, distribution control, multi-source imaging and outlier estimation.

Regarding efficiency, SURF remains one of the most relevant sped up versions of SIFT, but the computational cost is only slightly reduced. Some implementations were developed with GPU acceleration in mind. Such is the case with SiftGPU (Wu, 2007) and CudaSift (Bjorkman, Bergstrom, and Kragic 2014). On the downside, these implementations require the use of particular hardware, that is not commonly found in personal computers, adding to the fact that they lack the robustness to deal with large satellite imagery.

Multi-source image solutions have refined the SIFT descriptor, allowing it to deal with different main orientations of corresponding intensity points. In (Abraham and Leiss 2006), an improvement was proposed that allowed the registration between SAR and optical images. To determine the correspondence between multi-source images, a similarity metric based on local self-similarity (LSS) descriptor was proposed in (Ye and Shan 2014).

Distribution control was improved by methods such as Uniform Robust SIFT (UR-SIFT) (Sedaghat, Mokhtarzade, and Ebadi 2011) and the outlier elimination problem was mended by Scale Restriction SIFT (SR-SIFT) (Yi, Zhiguo, and Yang 2008), eliminating the obvious scale, rotation and translation differences between sample and reference image (Long et al. 2016).

Summary. In summary FBM methods are best used with samples that show a variety of easily detectable, distinct and well spread objects. RS samples tends to fall under this category of imagery. The aerial nature of these images often means that rivers, forest edges, urban areas, lakes, roads and other easily discernible features will generally be present and well spread throughout the whole image. On the other hand, images that lack distinctive features may prove to be problematic when working with methods relying on FBM. A common drawback of FBM methods is the fact that respective features may be difficult to identify or might be temporally unstable.

For FBM methods to work properly, it is indispensable to have discriminative and robust features descriptors that prove to be insensitive to all assumed discrepancies between the images (Flusser and Zitova 2003).

3.1.1.3 Transformation model estimation

This stage takes place right after the establishment of feature correspondence is complete. In this part of the process, a mapping function will transform the sample image with the intent of overlaying it with the reference image. A well design mapping function should assure that, after the transformation, the corresponding pairs of CPs from the sample and reference images are as close together as possible.

Choosing the type of mapping function, as well as its parameter estimation, is key when working on this stage of the image registration. Like in previous stages, the selection of the method is heavily dependent on three major aspects: the type of deformation that is to be expected, the acquisition method for the images and in the amount of accuracy required for the registration.

In some special cases, where the geometric deformations are partially known (mainly due to the existence of a model for the distortion caused by the sensing platform), a pre-correction can be applied, based on the inverse of the deformation.

Function mapping models can be broadly divided into two distinct categories. These models are classified in relation to the volume of image data they use as support and it goes as follows: Global mapping models and local mapping models.

Global mapping models use the total amount of CPs to estimate a single set of the mapping function parameters for the whole image. On the opposite side, local mapping models process the image as if it was composed of patches. The function parameters are then dependant on the location of their support in the image. This usually means the function parameter will have to be defined individually for each patch.

One other way of classifying mapping function models is by their general accuracy when it comes to overlaying the CPs used for the computation of the parameters. This distinction is usually found in literature under the following names: Interpolation functions and approximating functions.

Interpolation functions perform an exact mapping of the of the sample image CPs onto the reference image CPs. On the other hand, approximating functions work by finding the best trade-off between the final mapping accuracy and other requirements of the

mapping function. In most situations, the accuracy of the CPs coordinates is not required to be absolute, so the approximation model is widely used.

Global mapping models. One of the most widely used global mapping models uses bivariate polynomials of low degrees. The similarity transform is the most straight-forward method, consisting only in rotation, scaling and translation, Equation 3.1 and 3.2. Due to the preservation of curvatures and angles, coupled with the fact that this method is unambiguously determined by two CPs, it is referred to as shape-preserving mapping.

$$u = s(x\cos(\varphi) - y\sin(\varphi)) + t_x \quad (3.1)$$

$$v = s(x\cos(\varphi) + y\sin(\varphi)) + t_y \quad (3.2)$$

As an alternative, the affine transform can be used, and it proves to be slightly more general than the previously stated method, Equation 3.3 and 3.4. It remains a linear model, however, it is capable of mapping a parallelogram onto a square.

$$u = a_0 + a_1x + a_2y \quad (3.3)$$

$$v = b_0 + b_1x + b_2y \quad (3.4)$$

This model preserves straight lines and straight-line parallelism and it requires three non-colinear CPs to be defined. This method can be used in multiview registration applications, in situations where the scene is flat and there are no local factors to the present geometric distortion. The camera should also be pin-holed and its distance to the scene should be high in comparison to the expanse of the scanned area. In situation where this last condition could not be satisfied, the perspective projection model should be used, Equation 3.5 and 3.6.

$$u = \frac{a_0 + a_1x + a_2y}{1 + c_1x + c_2y} \quad (3.5)$$

$$v = \frac{b_0 + b_1x + b_2y}{1 + c_1x + c_2y} \quad (3.6)$$

If moderate violations of these conditions occur, it may be necessary to use models that take advantage of second, third or even higher order polynomials. In practical applications, however, high order polynomials see very little use. This is due to the fact that, when aligning, such methods may cause unnecessary warping within the sample image in areas that are located away from the CPs.

Usually, the number of CPs utilized in this process is superior to the minimum amount required to determine the mapping function. To minimize the sum of squared

errors at the CPs, the mapping function parameters are computed by means of least-square fit. Mapping functions such as these do not map CPs exactly to their counterparts and have proven to be very effective and accurate when dealing with satellite imagery, as an example (Flusser and Zitova 2003).

Local mapping models. Contrary to global mapping models, local mapping models can properly handle images that suffer from local geometric distortion. Such a local image deformity can be seen, for example, in airborne and medical imagery.

Least-square technique leads to the averaging of the local geometric distortion equally throughout the whole image, and such an effect is almost always undesirable. To fix this problem, registration of images that suffer from local geometric distortion should rely on information relating to each distortion in particular. Under these circumstances, many authors have proven the advantage of locally sensitive registration over the global mapping methods. Among others, Goshtasby (Goshtasby 1988a), Ehlers and Fogel (Ehlers and Fogel 1994), Wiemker (Wiemker et al. 1971), and Flusser (Flusser 1992) have discussed the topic.

By introducing some variations and adapting the least-squares method, the weighted least square and weighted mean methods were able to register images locally (Goshtasby 1988a). Local methods by the name of piecewise linear mapping (Goshtasby 1986) and piecewise cubic mapping (Goshtasby 1987), in addition to Akima's quintic approach (Wiemker et al. 1971), belong to the group of interpolating methods. They work by applying the combination of triangulation based on CPs and the collection of local mapping functions, each valid in the bounds of one triangle.

Radial based function mapping. These functions belong to global mapping methods. However, they are capable of handling images with locally varying distortion. These mapping functions were originally intended for the interpolation of irregular surfaces.

The most interesting property of these functions is the fact that, the value at each point, depends only on the distance of the points to the CPs, and not on their particular positioning. Hence the name "radial".

Multiquadric, reciprocal multiquadric, Gaussians, Wendland's, and thin-plate splines are some examples of radial based functions often used in the image registration process. In (Ehlers and Fogel 1994), multi quadratic functions are used to handle aerial RS imagery. It is compared to Akima's method in (Wiemker et al. 1971).

Thin-plate splines (TPS) shows very good result, however, in terms of computation it proves to be very time consuming, especially when a high number of CPs is present. Flusser (Flusser 1992) adapted the method and proposed simpler functions to adapt square or triangle regions. TPS, polynomials and multiquadric functions are used to reference aerial imagery and then compared in (Wiemker et al. 1971)(Goshtasby 1988b).

3.1.1.4 Resampling and transformation

Functions discussed in the previous section are used to transform image samples, completing the registration process. This transformation can be done in two distinct manners: forward and backward.

The forward method transforms each pixel from the sample image directly, using the estimated map function. This, however, can be quite complicated to implement, due to discretization and rounding, causing some holes and overlaps to occur in the output image.

The backward method sees a much higher use in practical applications. Registering information from the sensed image is determined by using the coordinates of the target pixel, together with the inverse of the estimated mapping function. Image interpolation takes place in the sample image on the output grid, meaning that no holes or overlap will be seen the output image.

Interpolation is generally done by convolution of the image with an interpolation kernel. However, optimal interpolants are hard to implement in practice. Because of this, some literature investigates the use of simpler interpolants of bounded support. To reduce the computational cost, separate interpolants are often considered.

Nearest-neighbour function, bilinear and bicubic functions, quadratic splines (Dodgson 1997), cubic B-splines (Hou and Andrews 1978), higher-order B-splines (Lehmann, Gonner, and Spitzer 2001), Catmull–Rom cardinal splines (Keys 1981), Gaussians (Appledorn 1996) represent some of the most used interpolants.

Some authors like Meijering et al. (Meijering, Zuiderveld, and Viergever 1999) proposed the use of higher-order polynomial kernels, however, only slight improvements were seen, at the price of a higher computational cost. For 2D images, a very detailed investigation was made in (Parker, Kenyon, and Troxel 1983), comparing the previous methods.

In terms of accuracy, bilinear interpolation is outperformed by methods of higher orders. However, it provides a better trade-off between accuracy and computational tax.

Cubic interpolation is recommended for situations when the geometric transformations cause a significant enlargement of the sample image. Nearest-neighbour is often avoided, except in cases where the sample image contains a small number of intensities and the introduction of graylevels (introduced by higher order interpolations) is not desired (Flusser and Zitova 2003).

3.1.2 Rigorous sensor model and rational function model

The Rigorous Sensor Model (RSM) is exceptionally useful when no GCPs can be acquired, usually due to small image contrast, high cloud cover, and so on. This method also benefits from being completely automated, requiring no human interaction.

By using collinearity functions, a relation between image points and homologous ground points is established. This relation is achieved through the use of parameters with known physical meanings.

It is important to note that RSM is only capable of modelling systematic and known distortions by using complicated platform movement and positioning models. This also means that this method is heavily dependent on the sensing platform, greatly reducing its flexibility.

One other problem is the fact that sometimes, parameters may not be readily available, causing difficulties. Some methods have been developed to help solve these issues, including the Rational Function Model (RFM), that has been thoroughly studied for photogrammetric processing in (Tao and Hu 2000). This method is well known for its sensor independence and for its similar accuracy when comparing with the RSM. However, RFM can sometimes have a hard time dealing with parameter interpretation and possible correlation between parameters.

Tong and Liu (Liu and Tong 2008) proposed a transformation between RSM and RFM to try and take advantage of both their strengths. It was concluded that such a method could be used successfully in practical geo-referencing applications. Other proposed examples that take advantage of RSM and RFM can be seen in (Pan, Tao, and Zou 2016), where they were applied to geo-rectify high-resolution imagery from the ZiYuan-3 earth observation satellite.

3.1.3 Summary

Geo-rectification of RS imagery, as a whole, involves an extensive array of problems as numerous as the combination of different sensors, RS imagery types and variety

of geometric distortions that can affect said image, leading to a vast amount of proposed solutions. After analysing a considerable number of methodologies relating to geo-rectification, it became clear that no single method is capable of solving the whole spectrum of challenges that can arise from this endeavour.

One of the major concerns of this dissertation relates to the geo-rectification of multi-spectral satellite imagery. As implied by the previous chapter, the vast majority of image samples will be provided by Sentinel-2A and Sentinel-2B, from the Copernicus Programme. Taking into consideration the nature of the applications that this dissertation aims to improve, it is safe to assume that images will contain a variety of easily discernible features and geometries. Furthermore, the data products provided by the Copernicus Programme offer a relatively high level of geometric reliability. However, despite the geometric quality of these images, the accuracy provided is still not enough to satisfy the needs of the target automatic classifiers.

As stated previously, given the nature of the expected RS imagery, usage of GCPs will be greatly explored. After analysing the data collected, regarding GCP acquisition, great results are to be expected form ABM methods due to the simple nature of the geometric distortions affecting the samples. These methods range from simple to complex (but highly reliable) algorithms and, by virtue of the high geometric quality of Earth observation imagery used, as well as major concerns relating to computational speed, the initial focus will be set on the medium to low complexity end of the algorithm spectrum, favouring techniques with lower computational requirements.

Despite the prediction that ABM methods will be capable of achieving sufficiently adequate results, FBM methods will also be explored due to the potential for highly discernible features within the samples, such as roads, rivers, buildings and so on.

Methods that do not rely on GCPs were also presented and have proven to obtain reliable results when correctly applied. Use of accurate error estimation models in the geo-rectification of data samples is common practice for most Earth observation missions, usually being the only technique applied when GCPs are unavailable. However, in the context of this dissertation, creation of a model capable of further geo-rectifying the data products provided is extremely unlikely, given the intended accuracy requirement.

3.2 Cloud cover correction

High-altitude RS imagery is of critical importance to many applications, namely ones related to land cover classification, landscape change detection, and so on. However, given the high-altitude nature, weather effects can sometimes prove to be problematic.

In particular, clouds that form in the troposphere are often responsible for blocking (partially or completely) the target ground zone, reducing visibility in the area of interest. This noise, depending on its severity, may critically compromise image samples rendering them unusable for many applications (Zhang et al. 2017).

Clouds themselves are generally described in literature as either *thick* or *thin*. Thick clouds consist of opaque obstacles that fully block visibility whereas thin clouds, while still corrupting visibility, maintain a level of transparency. In regions where cloud activity is especially high, like in the case of humid tropical areas, or in situations where sample imagery is innately hard to acquire, solving this problem may greatly increase the performance of the target application. (D. Tseng, Tseng, and Chien 2008).

Cloud cover removal is, in its essence, a data reconstruction process. By considering the origin of the complementary data used, these methods can be grouped into three main categories: non-complementation, multi-spectral and multi-temporal based approaches. Selecting the best approach is heavily dependent on the amount of data available, together with cloud characteristics.

Cloud detection is, in itself, a key aspect of the process. Ideally, cloud removal algorithms should only work on the area of interest (namely cloud contaminated areas). Therefore, without proper cloud detection, these applications will suffer greatly in terms of efficiency and accuracy.

Throughout this section, the most common methods pertaining to each of the previously mentioned categories will be explored and analysed, culminating in a summary section that will flesh out the advantages and disadvantages of each method, regarding the problem proposed by this dissertation.

3.2.1 Cloud detection

Usually, the process of detection is based on the assumption that clouds are colder and brighter than the ground, meaning they emit less infrared radiation and reflect more solar radiation to space, respectively (Melgani 2006). Some authors took this concept further and also proposed the exploitation of their spatio-temporal characteristic to help with detection, as can be seen in (Rossow et al. 1993).

As previously stated, clouds should also be classified according to their particular nature. This is especially important given the fact that *thick* and *thin* clouds produce different levels of noise, something that should be taken into consideration when looking for cloud removal solutions, Figure 3.5.

One other possible by-product of cloud presence is cloud shadows. This problem is mostly related to detection, since many solutions do not differentiate cloud presence from cloud shadows. This is particularly true for multi-temporal based approaches (Melgani 2006).

Classification is a very broad and far reaching topic, and numerous methodologies have been developed for the effect. Given the vast amount of problems encompassed by the topic of classification, it is hard for any given method to serve as a universally ideal solution. When it comes to cloud classification, some studies point towards artificial neural networks as the most effective method (Christodoulou, Michaelides, and Pattichis 2003)(Saitwal, Azimi-Sadjadi, and Reinke 2003)(Visa and Iivarinen 1997)(Tian, Shaikh, and Azimi-Sadjadi 1999). Another cloud detection method based on Multi-Temporal Cloud Detection (MTCD) is proposed in (Hagolle et al. 2010) and claims it can achieve good results on Sentinel-2 satellite imagery.

A detailed analysis of various cloud detection methods in satellite imagery was compiled in the form of a survey in (Chandran and Jojy 2015). It is important to keep in mind that, thanks to the nature of classifiers, most studies are prone to subconscious data manipulations, sometimes leading to contradictory conclusions between authors.

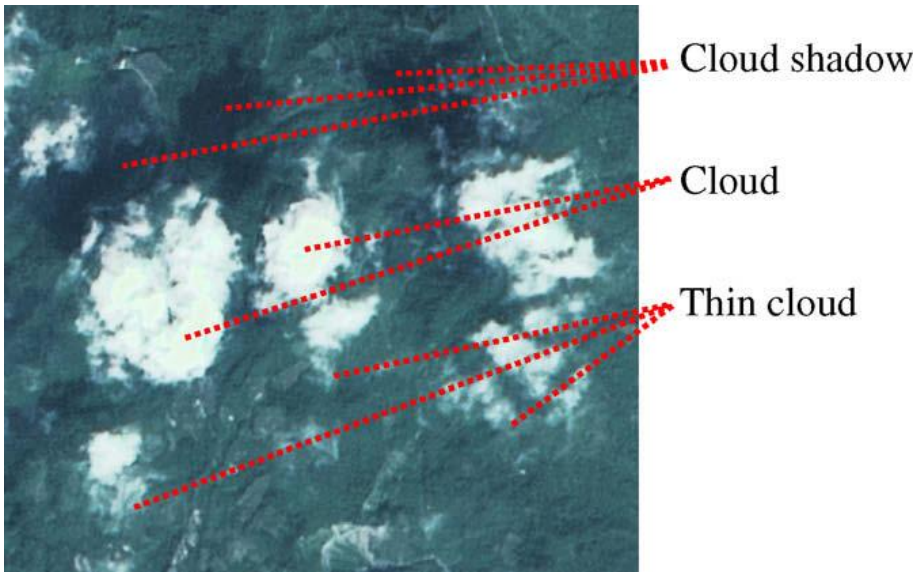
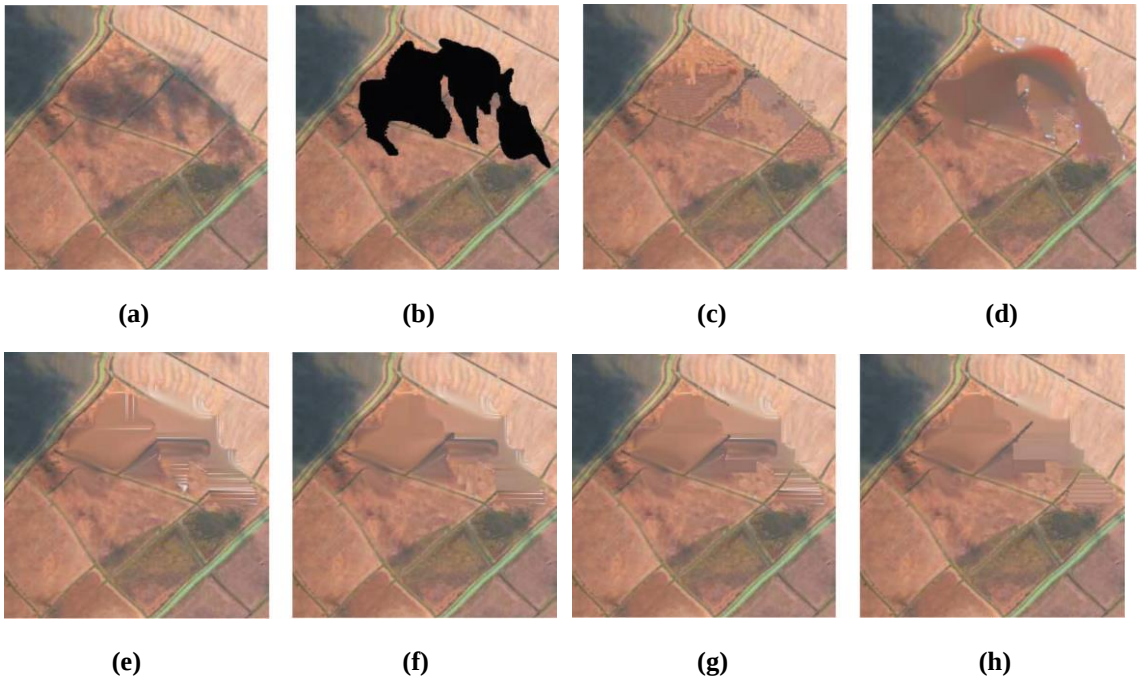


Figure 3.5: Depiction of three distinct varieties of atmospheric noise, taken from (D. C. Tseng, Tseng, and Chien 2008).

3.2.2 Non-complementation-based

Approaches based on non-complementation can be implemented without the need for complementary information. In other words, problems related to multi-temporal image processing do not apply, since all the data required by the algorithm is found on the image sample being worked upon.

Traditional inpainting and synthesis techniques can produce quite good results when applied to small sized defects, like scratches, and so on. However, these methods fail to produce satisfactory results when larger scaled flaws are present in the imagery. Despite this disadvantage, many improvements have been made over the past decades, allowing for longer region connection and texture recovering.



(a) **(b)** **(c)** **(d)**

(e) **(f)** **(g)** **(h)**

Figure 3.6: Image contaminated by cloud shadows (a) and mask of target area for reconstruction (b). Synthesized image using bandelet-based inpainting technique (Maalouf et al. 2009) (c), Liew et al. method (Liew, Li, and Kwoh 2004) (d), Tschumperlé and Deriche technique (Tschumperlé and Deriche 2005) (e), Peng et al. method (Peng et al. 2005) (f), Zhou et al. method (Zhou, Wang, and Qi 2006) (g) and Hsu et at. method (J.-F. Wang, Hsu, and Liao 2007) (h).

By taking advantage of uncontaminated pixels to correct cloud presence (based on a prediction), these methods are able to reconstruct imagery. Despite the visual plausibility achieved, more often than not the accuracy provided is not enough to satisfy the requirements for most applications (Zhang et al. 2017).

Despite these major shortcomings, methods based on non-complementation can obtain satisfactory results if cloud contamination consists mainly of thin clouds with a high-level transparency. A practical implementation of a non-complementation based method, as well as a comparison with other similar approaches, can be seen in (Maalouf et al. 2009), represented in Figure 3.6.

Considering the somewhat poor results usually obtained, coupled together with the amount of accuracy required to solve the problem proposed by this dissertation, these methods will not be studied further.

3.2.3 Multi-spectral based

Nowadays, multi-spectral imagery is a well establish commodity. Some RS platforms provide a plethora of multi-spectral bands that may be used independently or act as support data for many image processing applications.

When clouds are thin, meaning they retain some level of transparency, they are generally considered a form of low-frequency noise. Channels with shorter wavelengths are more susceptible to this kind of noise presence. Such is the case with the visible spectre, for example. In contrast, channels with a higher wavelength prove to be more resistant, as is the case with infrared and lower ends of the spectrum. By taking advantage of this phenomenon, multi-spectral data is able provide great results when it comes to cloud removal, in the aforementioned situation.

It is worth noticing however, that with an increase in wavelength comes a decrease in spatial resolution. Moreover, images originated from the sensing of longer wavelengths differ significantly from those that originated from the visible spectrum.

Since thin clouds are represented on the low-frequency domain, they can be extracted through the use of low-pass filters. Both Wavelet analysis (Du, Guindon, and Cihlar 2002) and Homomorphic filters (Stockham 1972)(Modestino 1979) have been utilized in thin cloud removal applications.

Wang et al. (Maramatsu and Fujiwara 1999) was able to restore imagery using only a few multi-spectral images of the scene. First, primary noise was removed by the use of a Wiener filter. This was followed by the exact segmentation of the area covered by infrared cloud, using Poisson Matting technique, and the use of wavelet analysis to restore the area that was originally occupied by infrared cloud. Residual holes were then restored by the use of a B-spline model.

In (Shen et al. 2014), a locally adaptive method to remove thin clouds in visibly sensed data is proposed. It makes use of an adaptive homomorphic filter and, despite its simplicity, provides good results. The method proposed by (Enomoto et al. 2017) takes advantage of Multi-spectral conditional Generative Adversarial Networks (McGANs) based on conditional Generative Adversarial Networks (GANs). An improvement to the homomorphic filtering method is proposed in (Feng et al. 2004).

Methods based on multi-spectral complementing data prove to be quite capable of effectively dealing with thin cloud removal. However, when thick clouds are present, these methods prove to be unable to properly reconstruct image samples. This is mainly due to the high opacity of thick clouds, that manage to somewhat corrupt lower wavelength originated imagery and completely block imagery in the visible spectre.

Furthermore, the use of filters to remove thin cloud noise may also inadvertently filter out useful image data, since bright and large land cover masses may share similar spectral characteristics with clouds.

3.2.4 Multi-temporal based

When the sensing frequency is sufficiently high, using multi-temporal complementary data becomes a possibility. By mosaicking samples and fusing cloud free data it is possible to create a cloud free, high-quality composite image, Figure 3.7. This assumes, however, that no relevant geographical changes have occurred between image samples.

Generally, the multi-temporal imagery used in these techniques is original to the same sensors and sensing platforms, which helps retain core characteristics such as spatial resolution, bandwidth, and so on. Some authors proposed the use of multi-temporal, multi-sensor imagery and, despite the increased difficulty, positive results have been achieved.

The multi-temporal approach is highly used when it comes to cloud cover rectification. This is mainly due to its effectiveness in dealing with thick clouds, something that non-complementation and multi-spectral approaches have trouble with. Currently, multi-temporal based methods can be sorted into three distinct categories. These are often described in literature as: integration-prediction, temporal-replacement and self-replacement methods (Zhang et al. 2017).

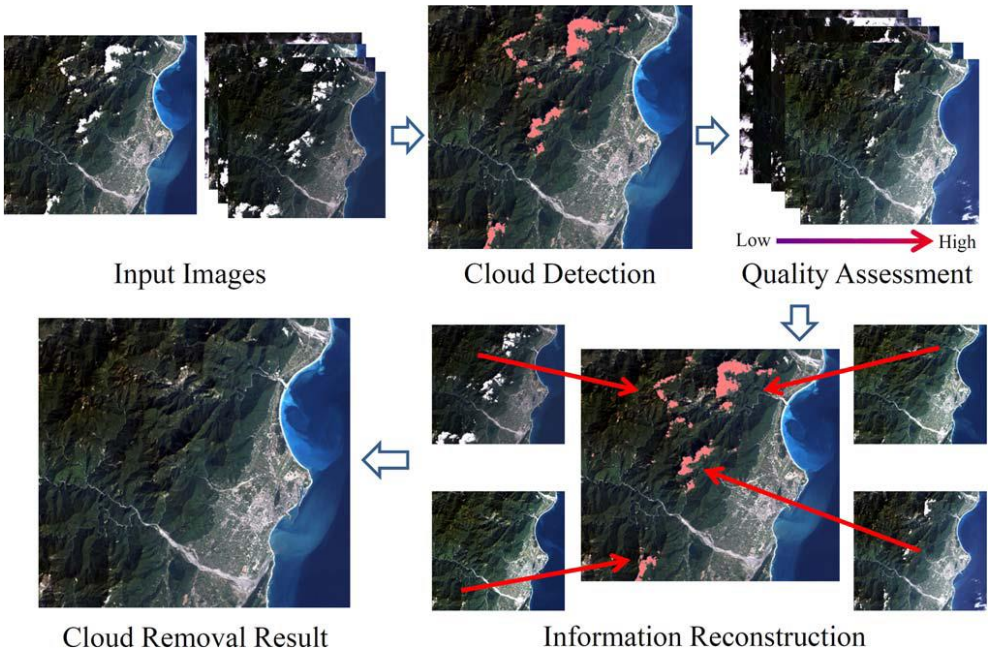


Figure 3.7: Process of cloud removal by means of multi-temporal image fusion, taken from (Lin et al. 2013).

3.2.4.1 Integration-prediction

Like the previously presented non-complementation-based methods, integration-prediction methods work, as the name implies, by predicting corrupted or missing image data. However, unlike nom-complementation, these methods take advantage of reference imagery to help guide the algorithms. Such methods usually deal with images on a pixel-by-pixel or patch-by-patch basis (Zhang et al. 2017).

In (Melgani 2006), two novel cloud removal methodologies have been proposed. In the first method, a contextual prediction process is implemented by a group of linear predictors. The second method takes advantage of local spectral-temporal relationships, that are reproduced by a single non-linear predictor based on the Support Vector Machines (SVM) approach.

A method proposed by (Zeng, Shen, and Zhang 2013) uses a multi-temporal approach to correct most of the defects present in the imagery. When auxiliary multi-temporal data is unable to recover the target pixels, a non-reference regularization algorithm is applied to fully restore the sample image.

The use of regression tree models to predict the band values of cloud pixels in a reference scene, by using other scene dates, is proposed by (Helmer and Ruefenacht 2005). This method improves histogram matching to adjacent scenes by using maximum likelihood classification and spectral data.

3.2.4.2 Temporal-replacement

Methods based on temporal-replacement are generally straightforward, consisting only in a few steps. Same as all cloud removing techniques, these methods start by identifying cloud masses on the imagery. These regions of interest are then replaced by cloud free sections found in reference images, followed by a brightness adjustment.

Given the right circumstances, these methods are able to provide outstanding results, however, they are highly dependent on a multitude of factors. First, the availability of relevant cloud free patches in the sample imagery is crucial. If the cloud areas overlap, no relevant complementary data will be available, and a restoration will not be possible. Second, sample and reference images must be accurately calibrated, meaning that no geometric discrepancies should exist between the two. Finally, sensing frequency must be relatively high, or at least high enough to ensure that no relevant geographical changes occurred between sample and reference images (Shen et al. 2014).

All these requirements mean that temporal-replacement methods are sometimes limited in their application. Despite this fact, many propositions based on this method have been made throughout the years.

In (Lin et al. 2013), a patch based reconstruction, mathematically formulated as a Poisson equation, is proposed. This approach avoids the use of a global optimization process, meaning that good results in terms of consistency and radiometric accuracy are to be expected. A great variety of novelty tools, also based on the Poisson equation, are introduced in (Pérez, Gangnet, and Blake 2016).

Tseng et al. (D. Tseng, Tseng, and Chien 2008), proposed a multidisciplinary method to detect and remove clouds, outputting mosaic cloud free images from SPOT imagery in three steps. The first step consists in the enhancing the brightness and chromaticity of image samples. The second step involves choosing the image with less cloud presence to set as a base for the transformation and then dividing it into grid zones (with thin clouds being detected in the neighbourhood of thick clouds). The final step is the replacement of cloud zones with cloud free zones from the auxiliary multi-temporal data. This method shows positive results, outputting high-quality imagery with varying degrees of brightness.

3.2.4.3 Self-replacement

Much like integration-prediction methods, self-replacement methods are guided by a reference image. Cloud contaminated data is then filled by using the remaining information found in the target image, hence the name self-replacement.

Cheng et al. (Cheng et al. 2014) demonstrates this method well. To recover missing pixels, an algorithm locates a similar pixel within the remaining image. In order to maximize the effectiveness of this method, multi-temporal complementary data is used as guidance to locate the similar pixel. A spatio-temporal Markov random fields (MRF) global function, based on pixel-offset, is employed to further help select the most suitable similar pixel. The study stated that this method is effective and highly accurate when it comes to dealing with obvious seasonal and atmospheric differences.

3.2.5 Other methods

Nowadays, it is not uncommon for cloud removal methods to be multidisciplinary in nature, combining two or more of the previously presented techniques. Opting for a multidisciplinary approach generally increases the robustness of the application but may also prove to be computationally taxing.

Throughout this subsection, methods that do not quite fit previous subchapters will be presented and analysed.

In (Q. Wang and Atkinson 2018), Wang et al. claims that a novel method called Fit-FC can be used to obtain nearly daily Sentinel-2 imagery by way of spatio-temporal fusion between Sentinel-2 and Sentinel-3 data. This method includes three techniques: a regression model fitting (RM fitting), special filtering (SF) and residual compensation (RC). RM fitting is used to relate images acquired at two different times, minimizing their differences. The SF is employed to remove artefacts present in the RM fitting prediction. Finally, RC compensates for residuals caused by RM fitting in order to preserve the spectral information. The study concluded that Fit-FC can obtain good results and that its use is particularly relevant in situations where correlation between coarse images is small.

Zhang et al. (Zhang et al. 2017) summarized and compared some of the previously presented multi-temporal methodologies and proposed a multi-temporal Poisson based method of its own. This method proved to be effective at dealing with significant multi-temporal changes in imagery.

He et al. (He and Yokoya 2018) proposed the combined use of Sentinel-1 SAR data and Setinel-2 data to produce cloud free Sentinel-2 image samples. This method took advantage of convolutional neural networks (CNN) to capture image characteristics from SAR data. For training, a generative adversarial network proved to be useful and effective.

3.2.6 Summary

When it comes to image pre-processing, geo-rectification and cloud cover correction are two key concerns. However, literature regarding the latter is scarce in comparison to the vast studies performed on the topic of geo-rectification. Despite this discrepancy, many great efforts have been made to find solutions capable of detecting and removing cloud presence in Earth observation imagery. Proposed methods range from prediction and synthetization algorithms to multi-spectral and multi-temporal approaches.

Cloud detection and classification plays a major role in the overall process. Without proper detection, algorithms are unable to rectify the relevant sections of the image, resulting in unsatisfactory results. Cloud classification by type is also considered a relevant subject, given the fact that clouds may share different characteristics. In literature, clouds are usually classified as *thick* or *thin*, with both having distinct characteristics.

Thin clouds can be dealt with in numerous ways. Because of their innate transparency, prediction algorithms can take advantage of uncorrupted pixels in the neighbourhood of cloud pixels, restoring the image and achieving very positive results in some situations. Some methods work with multi-temporal complementary data to help guide these algorithms while some work without the need for complementary information. The latter methods obtain poorer results but may act as an alternative when complementary data is scarce.

From a frequency perspective, thin clouds exist in the low-frequency domain, which lead some authors to propose the use of filters to remove thin cloud presence. However, some land mass features may share similar spectral characteristics, meaning they would be filtered as well, resulting in the destruction of possibly useful pixels. Some other methods utilize multi-spectral (and sometimes also multi-temporal) complementary images to bypass cloud noise through use of the longer wavelength imagery. Despite major differences between data types, it is possible to correlate these images to produce a new, cloud free sample.

Thick clouds, on the other hand, pose a much bigger problem to the previously mention methodologies. Given their large size and opaque nature, prediction algorithms lack the necessary information to accurately recreate covered land masses. Despite their increased resistance to cloud noise, under these conditions, even longer wavelength bands also start suffering from varying degrees of corruption.

Multi-temporal methods, namely ones based in image fusion, are capable of solving all manner of problem related to cloud cover removal. However, these methods are highly dependent on the quality and availability of reference imagery. Despite having a

simple methodology behind them (consisting mainly in the replacement of cloud affected areas, followed by a brightness adjustment), these methods require a special set of circumstances to work properly. In particular, if cloud overlap is present between sample and reference images, these methods have no way of correcting the image. Furthermore, if geological changes occurred between images, these discrepancies will be visible in the final product, greatly reducing quality.

In recent years, some methods have started to use a combination of some, or all, of the previously presented methodologies. These approaches garner some interest thanks to their increased robustness, applicability and accuracy. However, these methods may be too computationally taxing for some applications, namely ones that are reliant on speed or are meant to be ran on lower end hardware.

Given the context of this dissertation, methods based around inpainting and synthesis will not be able to provide the required pixel accuracy. On the other hand, methods centred around image fusion are able to synergise well with the high temporal-resolution of Sentinel-2 data products, providing a computationally light and accurate solution for all manners of atmospheric noise.

4

Prototype development

This chapter is responsible for documenting the work developed throughout the course of this dissertation. It will open with a short introduction on the development platform, followed by an in-depth study of all relevant algorithms designed pertaining to both geo-rectification and cloud-cover correction. Lastly, the QGIS plugin designed to run the required operations will be showcased and explained.

4.1 Development platform

Selecting a suitable development platform to create and test the automated system proposed by this dissertation was the first requisite to be filled during the initial stages of prototype construction. Due to the multi-temporal focus, this project required handling of large amounts of geographical data and so, to facilitate information logistics, GIS software had to be acquired. A stable source of high-quality sample data would also be required to provide the means to test and validate the developed algorithms.

4.1.1 Data sets

Prototype development was expected to be heavily influenced by the nature of the data samples available for testing, accentuating the importance of a reliable source of Earth observation image data products.

Primarily, this dissertation envisioned a system capable of pre-processing satellite imagery at high temporal resolutions, providing a larger and richer database for land cover classification applications. To reflect this philosophy, an Earth observation mission that

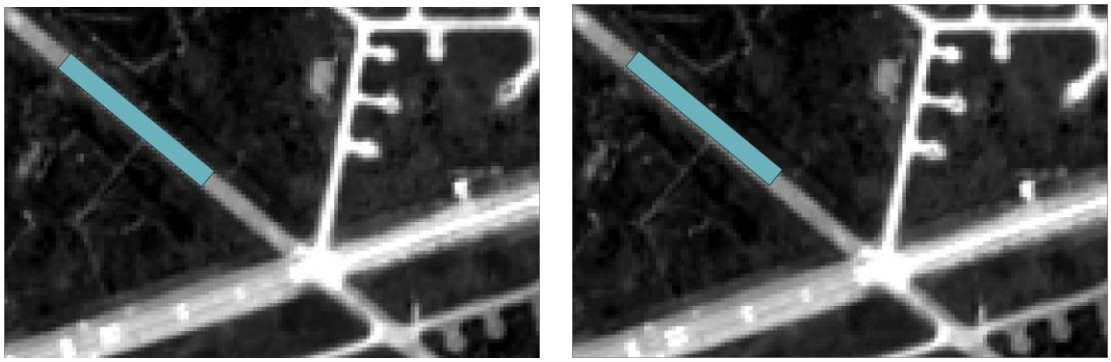
could offer multi-spectral, high temporal and spatial resolution imagery of ocean and land cover had to be incorporated into the overarching design.

Sentinel-2, from the ESA's Copernicus Programme was chosen to be the primary source of data samples for the development of both geo-rectification and cloud-cover correction algorithms. This dissertation was focussed around the use of Level-1C and Level-2A imagery, as classified by ESA.

As previously stated, both Level-1C and Level-2A data products provided by Sentinel-2 include geometric and radiometric corrections, tethering pixel values to geographic coordinates.

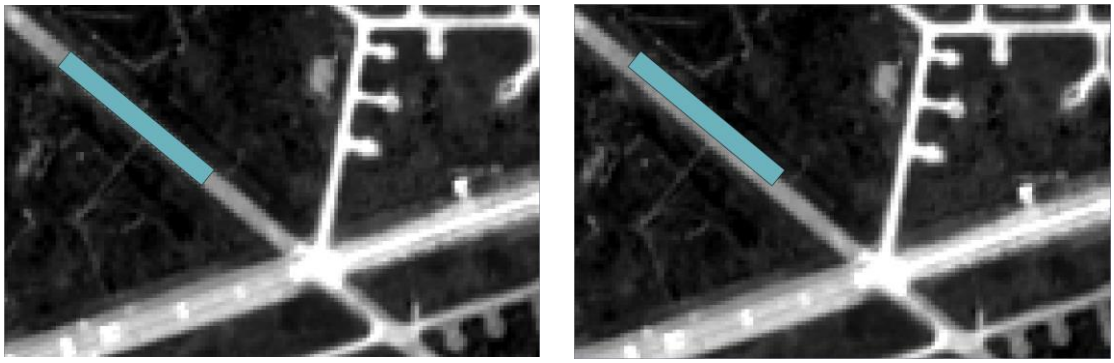
The calibrations regularly effected by ESA can still contain a non-negligible sum of geo-referencing errors, depending on the amount of precision required for the desired application. These discrepancies rarely exceed the 1.5-pixel mark, however, at a subpixel level they become quite prominent.

Figure 4.1 and Figure 4.2 depict a usual (0 to 0.5-pixel) geometric translation error and a simulated maximum deviation of 1.5-pixels respectively. The damaging effects of poor geometric precision can be seen here. Pixel and subpixel discrepancies can have devastating effects in land cover classification applications, particularly when the target area of interest is represented by a scarce number of pixels, as is the case with firebreak monitoring.



(a) **(b)**

Figure 4.1: Usual geometric translation error in band 04 of Sentinel-2 Level-1C products; 2017/11/11 (a) 2017/11/06 (b).



(a) **(b)**

Figure 4.2: Maximum expected geometric translation error in Sentinel-2 Level-1C products; 2017/11/11 (a) and an artificial translation of 1.5-pixel in the X and Y coordinates (b).

4.1.2 GIS software

The geographical nature of this dissertation motivated the use of GIS software. To manage, present and analyze the Earth observation data products required, QGIS, previously known as Quantum GIS, was chosen in virtue of its opensource nature and modular design. The specific version used is labeled as 2.18.28 (*Las Palmas*).

The software also incorporates opensource packages such as GRASS GIS, GDAL Tools, Orfeo ToolBox, among others. A modular design allows for download and incorporation of plugins written in Python or C++, extending QGIS's capabilities even further. Plugin creation alongside some of the packages included within QGIS will be examined in subsequent subchapters.

Early in development, Python was established as the core programming language for this dissertation. This decision stemmed from the inherent simplicity of its syntax and its overall popularity among QGIS users.

Support for Python (PyQGIS) was first introduced to QGIS back in version 0.9 and has since then been used to perform a variety of tasks within the software, such as scripting and plugin construction. A major advantage held by Python over C++ lies in its enhanced plugin distribution capabilities, a trait that synergises well with its straightforward development. Python version 2.7.14 was used in the creation of all algorithms presented throughout this dissertation.

4.2 Image preparation

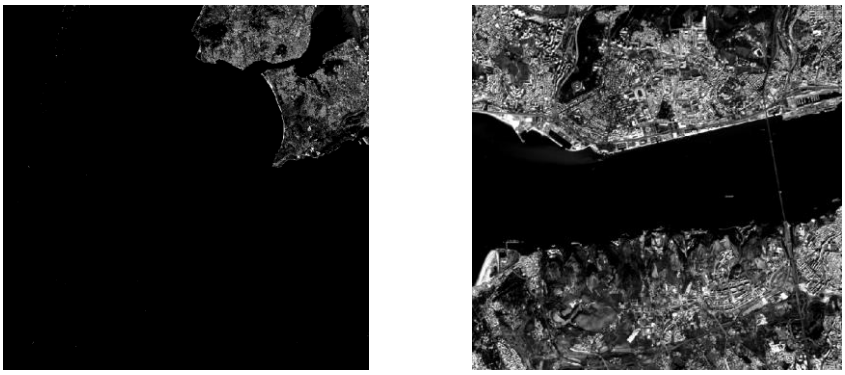
To increase the overall efficiency of the desired system, image data samples were put through a variety of pre-processing stages that would either help reduce computational costs or enhance performance in terms of practical results. Image transformations resulting from these pre-processing stages quickly became backbone components for the algorithms developed throughout the course of this dissertation.

This section will present, justify and explain these simple yet crucial operations responsible for preventing wasteful resource expenditure and increasing the overall accuracy of the system.

4.2.1 Cropping

Poor data management can lead to unnecessary and sometimes unreasonably high computational expenses associated with digital image processing. With an increase in pixel count, runtime for these applications will also increase accordingly.

Most Earth observation data products are provided in the form tiles, these being 100x100 km² in the case of the previously mentioned Landsat and Sentinel-2 missions, Figure 4.3. Despite the accentuated flexibility provided by this geographically encompassing size, the sheer volume of data may hinder or become a prohibitive factor in many land cover classification applications due to insufficient processing power. By removing irrelevant data from the process altogether, computational resources are given the chance to focus exclusively on the target observation area, circumventing a substantial portion of the issue. The processes responsible for removing any unwanted outer areas of a picture can be broadly referred to as cropping.



(a) **(b)**

Figure 4.3: Full tile from Sentinel-2 Level-1C product (a) and 1600% zoom over Lisbon (b); Band 04 – 2017/11/11.

A core part of the design philosophy behind this dissertation was centered around the use of readily available opensource tools, which would then be incorporated into a new modular product. This led to prioritized use of toolsets incorporated into QGIS over the creation of new tools from the ground up, which culminated in extensive use of GDAL Tools.

The created system was designed to allow the user to choose between two distinct cropping mechanisms, based on situational or personal preference. The first approach requires the user to input the X and Y upper left and lower right coordinates manually or by use of a script, forming a rectangle over the target observation area. The second approach allows the user to select a proxy shapefile (.shp) of the target observation area.

4.2.1.1 Coordinate method

Taking advantage of the GDAL Tools toolset, this method aimed to provide the user with the means to crop any given image data sample through simple use of the coordinate reference system present in QGIS.

Despite having been primarily designed as a function to convert raster data between different formats, *gdal_translate* is also capable of performing a variety of different operations. Among these utilities *-projwin, -outsize,* and *-r* were used to great effect in the creation of this cropping mechanism:

- **-projwin:** Arguably the most significant operation in this process, allowing the selection of a subwindow through the use of a coordinate system;
- **-outsize:** Fundamental in assuring equal resolutions between data samples, allowing them to be up or downscaled at will;
- **-r:** Synergising with the previous command, allowing the user to choose between a variety of resampling algorithms such as nearest, bilinear and cubic, among others.

By use of simple string manipulation techniques, it is possible to create a dynamic algorithm capable of taking full advantage of these commands, outputting a variety of results as requested by the user.

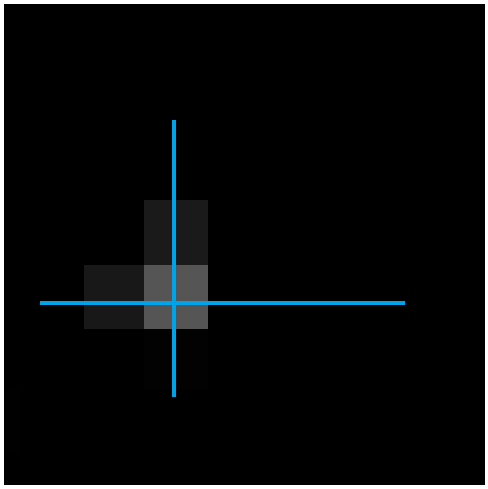


Figure 4.4: Possible horizontal/vertical subpixel mesuments in Sentinel-2 Level-1C product; Band 04 – 2017/11/11.

It is important to keep in mind that the coordinates selected might fall under the subpixel range of the raster image being processed, meaning that one or more pixels would be required to split into two separate entities, Figure 4.4. This, however, is not possible and under these circumstances *gdal_translate* assures that any pixel, whether fully or partially covered by the rectangle denoting the target area of observation, will be entirely represented in the output raster image.

4.2.1.2 Polygon method

The second cropping method was designed as an alternative, allowing users to take advantage of shapefile data as a means of representing the desired area of interest.

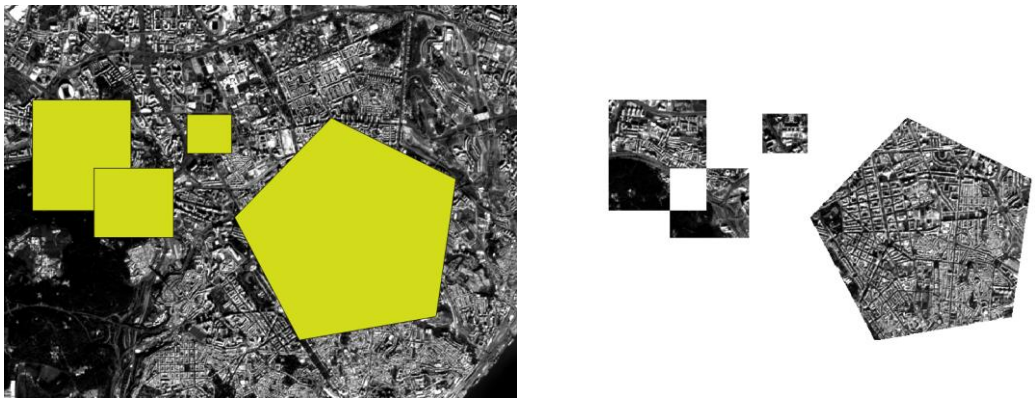
A function by the name of *gdalwarp,* included yet again in the GDAL Tools tool-set and originally designed as an image mosaicing, reprojection and warping utility, was selected as a foundation by virtue of the *-cutline* and *-crop_to_cutline* utilities:

- **-cutline:** This operation enables the use of a blend cutline, provided in the form of a shapefile (.shp);
- **-crop_to_cutline:** Complements the previous operation by cropping the extent of the data set to the extent of the previously defined cutline.

String manipulation techniques are employed once more, creating a dynamic algorithm capable of gathering the required elements and selecting or creating an output file location as requested by the user.

In similar fashion to the first cropping method presented, the polygon shape may end up having one or more edges located in subpixel regions, resulting in the aforementioned pixel division issue. To overcome this problem *gdalwarp* will only represent pixels in full, however, contrary to the first method which would output any pixels fully or partially covered by the rectangle representing the area of interest, this second method only outputs pixels fully encompassed by the polygon shape, Figure 4.5.

A benefit held by this method over the previously presented one stems from the fact that a single shapefile can represent a variety of polygons. This may result, if need be, in a 'composite' image that separates pixel clusters with a transparent layer of "empty pixels" devoid of any information. It is also important to note that the complexity and number of polygons used can affect performance.



(a) **(b)**

Figure 4.5: Shapefile with multiple polygons represented over Lisbon (a) and cropping output image (b); Sentinel-2 Level-1C product, band 04 – 2017/11/11.

4.2.2 Filters

In digital image processing, any technique used to transform or enhance an image by employing the use of kernels and convolution in the spatial domain or masking of specific frequencies in the frequency domain can be broadly described as filtering. When selected and calibrated properly, filtering techniques can vastly improve the performance and accuracy of many detection algorithms, which naturally led to the exploration of this topic early on in development.

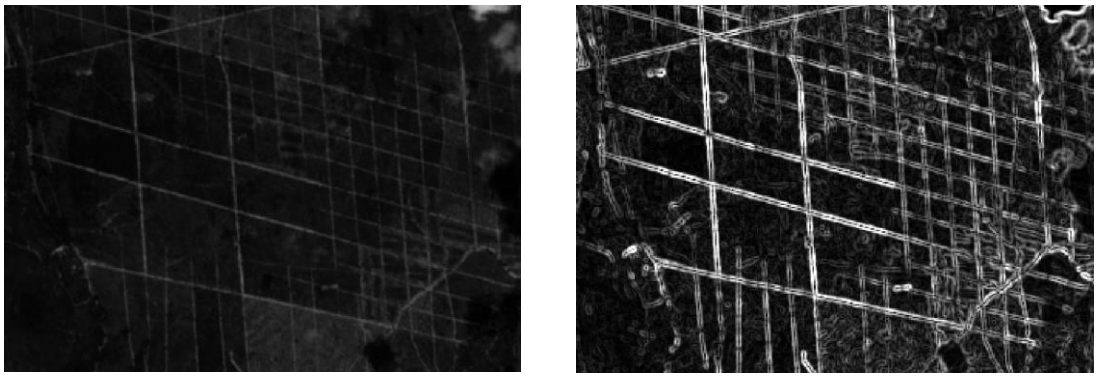
Neither Sentinel-2A or Sentinel-2B platforms can provide the parameters necessary to model the geometric discrepancies found between image data samples, resulting in the development of algorithms based on the use of GCPs. It was anticipated that filters

capable of enhancing the process of edge detection and extraction could prove beneficial to the overall accuracy of the system. Binarization methods were also explored as a way of possibly achieving further benefits in edge and shape detection as well as in correlation efforts.

4.2.2.1 Sobel

Use of this filter is mostly associated with edge detection algorithms in the field of computer vison. When properly applied, this operator enhances the target by convoluting the image in a 3x3 kernel, emphasizing its edges. Focus on the use of opensource and readily available tools was once more a priority, avoiding the necessity of designing a Sobel operator from the ground up, which would only be conducted as a last resort.

Included in the Orfeo ToolBox toolset, which can be downloaded and integrated into QGIS, are two edge extraction utilities that were experimented with during the prototype development phase of this dissertation, one of them being the Sobel operator, seen in Figure 4.6. This feature allows for little customization, as the only parameters that can be tampered with are the input and output locations, together with a channel index. Due to the relatively low amount of flexibility provided, coupled with some problems regarding Python integration, this method was soon abandoned and replaced.



(a) **(b)**

Figure 4.6: Original Sentinel-2 Level-1C product (a) and Sobel transformation from Orfeo ToolBox (b); Band 04 – 2019/04/18.

The second attempt at a Sobel filter derived from the use of the scikit-image Python library. Included in the *scipy.ndimage* class of this library is a Sobel operator that provides a greater variety in terms of customization options. Among these options, the standout features are the independently selectable X and Y input axis along which the

calculations for the transformation are performed and the mode parameter responsible for determining how the input array is extended when the filter overlaps a border, offering:

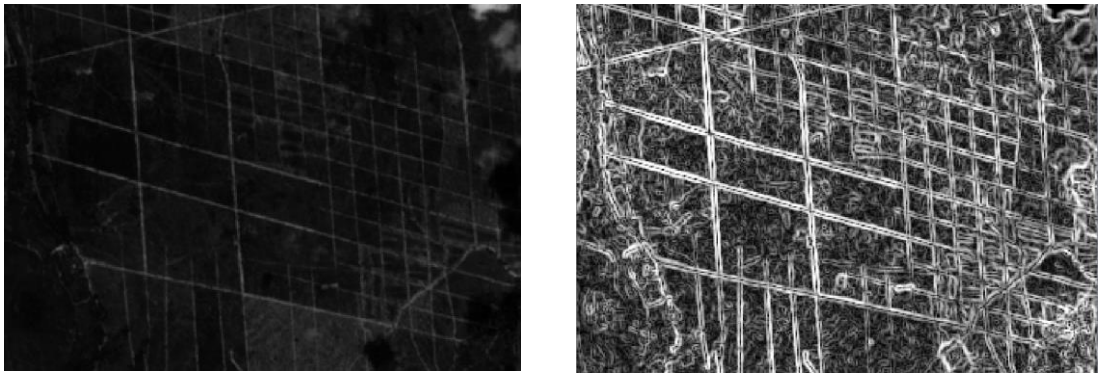
- **Reflect:** Extending the input by reflecting the edge of the last pixel;
- **Nearest:** Replicating the last pixel;
- **Mirror:** Reflecting about the centre of the last pixel;
- **Wrap:** Wrapping around the opposite edge;
- **Constant:** Filling all values beyond the edge with the same constant value, specified by the user.

This method was selected due to its easy implementation, customizability and performance, ultimately becoming a possible pre-processing transformation for any of the algorithms developed throughout this dissertation.

4.2.2.2 Touzi

This filter was the second and last edge detection utility experimented with in the Orfeo ToolBox toolset. In contrast to the Sobel operator also found in the same package, the Touzi filter allowed the user to customize two variables, these being the X and Y radius of the neighborhood. Increasing this length decreases filter sensitivity regarding speckle noise but may, in turn, cause it to miss the micro edge.

Despite demonstrating some potential, the edge enhancements attained were comparable, and deemed inferior, to the ones provided by the previously presented Sobel operand, Figure 4.7. Given the circumstances, this operand was abandoned due to its redundancy.



(a) **(b)**

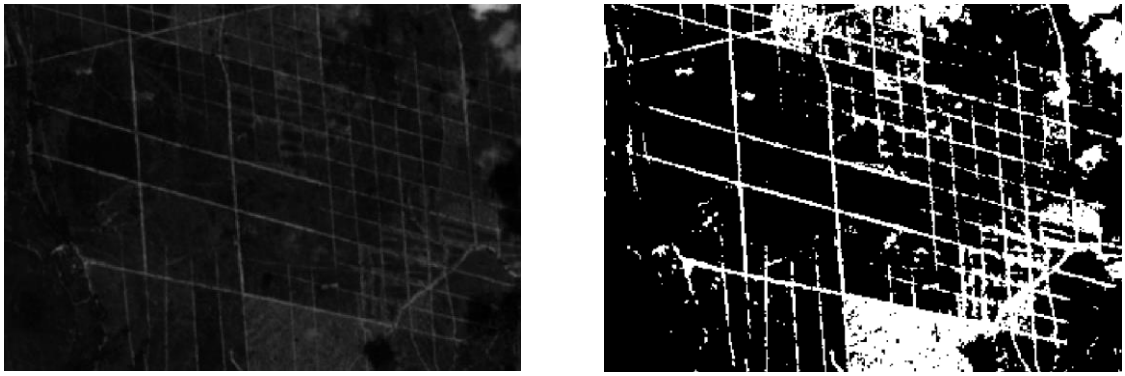
Figure 4.7: Original Sentinel-2 Level-1C product (a) and Touzi transformation from Orfeo ToolBox with an X and Y radius of 1 (b); Band 04 – 2019/04/18.

4.2.2.3 Otsu

In its simplest from, this method is intended to return a single intensity threshold which would then be used to separate pixels into two classes, background and foreground. Transformations of this nature are often referred to as binarization, as pixels intensities can only assume one of two values, usually represented as pure white and pure black.

Naturally, if an image possesses a histogram with bimodal tendencies this operation is more likely to isolate and distinguish relevant features. If the threshold is properly selected, discrepancies between data samples such as differences in lighting as well as thin cloud presence, may be mitigated or bypassed altogether. However, if applied incorrectly this filter can result in the destruction of relevant features, deteriorating the overall quality of data samples and diminishing the accuracy of correlation efforts.

Implementation of the Otsu threshold was done through the use of a function included the OpenCV Python library. The *cv2.THRESH_OTSU* function can be used to either manually select or automatically calculate the optimal threshold value of the data sample provided, depending on histogram distribution, Figure 4.8.



(a) **(b)**

Figure 4.8: Original Sentinel-2 Level-1C product (a) and Otsu binarization from OpenCV Python library (b); Band 04 – 2019/04/18.

4.2.2.4 Sobel-Otsu

A single data sample can be subjected to a multitude of pre-processing transformations which, depending on the operands and how they are stringed together, may result in greater feature sharpening. However, most image transformations will inevitably add a layer of noise that, when stacked, can undesirably deteriorate the sample and affect the overall accuracy of the system.

This dissertation explored the synergy between the Sobel operand and Otsu threshold, two of the previously presented transformations, as a possible method of increasing the contrast of edges and features. The order in which these pre-processing operations are applied to the image is homologous to the name given to the process, starting with the Sobel operand and concluding with the Otsu threshold.

4.3 Geo-rectification

Accurate geo-rectification of Earth observation data was a primary concern and, as a reflection thereof, assuring it became the first step in creating the automatic system proposed by this dissertation.

This subchapter is dedicated to all attempts at developing an algorithm capable of mitigating the issues caused by geo-location discrepancies between data sets. Throughout the subchapter, in-depth explanations concerning the components and operation of each individual approach will be provided, alongside a design perspective detailing the predictions and decisions that guided the development process of these algorithms.

4.3.1 Exhaustive method

The exhaustive method design was focused on the creation of a proof of concept and, as such, this method was never envisaged as a final solution to the geo-rectification aspect proposed by this dissertation. The didactic nature of the development process led to the design of a theoretically feasible solution loosely based around sequential similarity detection, intended to provide context to the results obtained by the final geo-rectification algorithm, while also helping to consolidate knowledge on Python programming, data manipulation by GIS software and plugin construction.

Planning and development phases were guided by the core hypothetical principle that with small incremental transformations, namely translations over the axes of the sample, it would be possible to adequately match the geometric parameters of the reference image. It was also hypothesized that if one of these incremental translations directed the sample closer towards an eventual full geometric match (in both axes), it would achieve a higher correlation result compared to all other translations in the cycle.

To implement such an approach, it was deemed necessary to compare both sample and reference images before and after each transformation, these transformations being a translation over the X, Y or a combination of both axes, followed by a comparison of all correlation values attained. After completion of the first cycle, the sample would be required to move towards the position associated with the highest correlation value,

followed by a new iteration of the entire process. The sample would only be considered sufficiently geo-rectified when all new translations resulted in lower correlation values when compared to the one obtained on the pre-transformation sate, meaning that any further adjustments would inevitably lead the sample further away from optimal geometric matching, Figure 4.9.

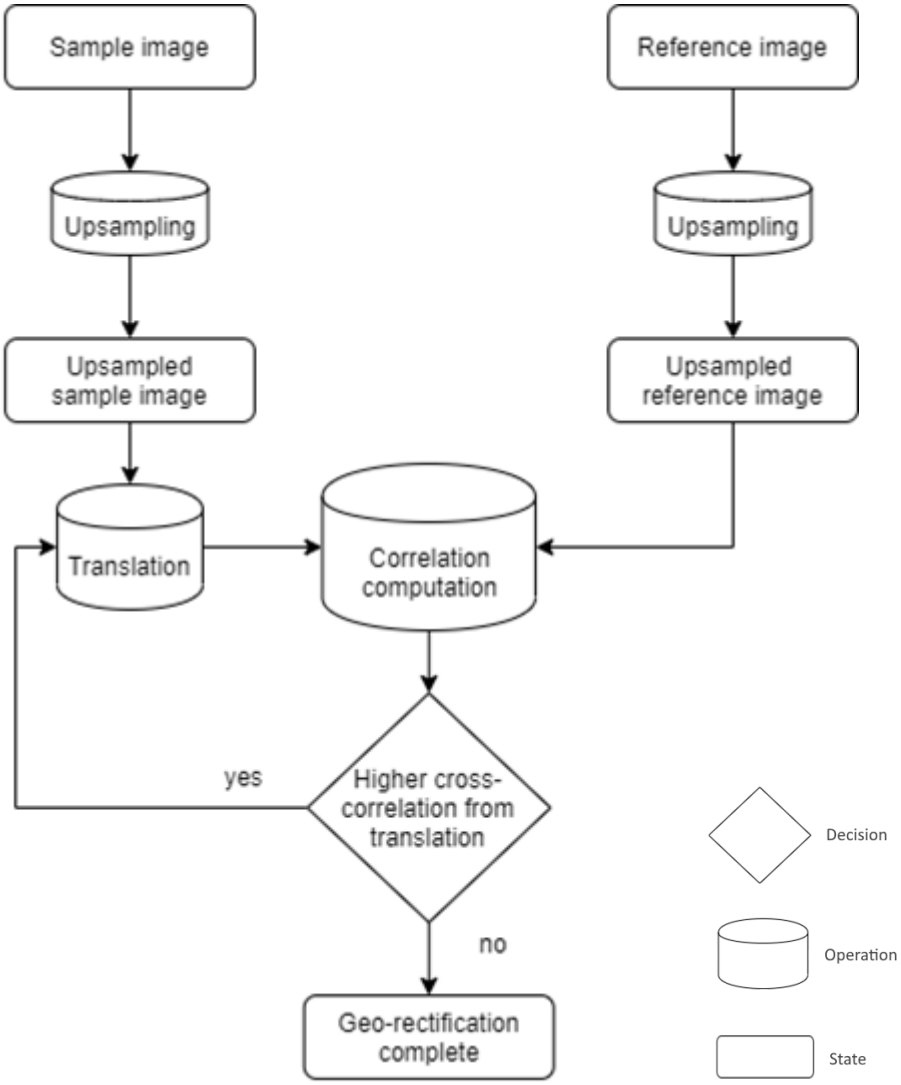


Figure 4.9: Exhaustive method diagram for multiple cycle implementation.

4.3.1.1 Procedure

In practice, this algorithm was designed to work with three sets of data, the sample and reference images, and an integer threshold value used for correlation measurements. Effective geo-rectification hinged on the ability to translate and correlate images on a subpixel level. The algorithm was hardwired to perform incremental translations of a half

pixel length due to its relative mathematical simplicity. This required both data sets to be upscaled by a factor of two. By using nearest neighbour interpolation, every individual pixel is represented by a 2x2 pixel grid, retaining all of the original information. Shifting this new matrix by a single pixel in any given direction is comparable to the same half pixel translation in the original sample. It is important to note that despite never suffering a translation, the reference image must also undergo the upscaling process to ensure an accurate correlation measurement.

Through the use of *gdal_translate*, included in the GDAL Tools toolset, it is possible to manipulate the X and Y output size of any sample provided. This can be achieved either by designating a percentage tied to the original sample size or by specifying the number of columns and lines desired for the output. The use of a percentage-based transformation was naturally chosen for this algorithm on account of its simplicity and certitude in the preservation of aspect ratios.

Once the sample and reference images have been upscaled, they are then treated as a matrix of pixels, each of them representing an individual integer value. This matrix is then shifted in all cardinal and intercardinal directions, resulting in a total of ten distinct matrices, these being the original sample and reference matrices along with all the subsequent translations. To achieve these translations in a Python setting, both *numpy.array* and *numpy.roll* functions from the NumPy library were used to convert sample and reference data into arrays and subsequently shift elements along the desired axes, respectively. The *numpy.roll* function avoids destroying information by guaranteeing that any element shifted beyond the last position will be orderly re-introduced back into the first. After finalizing the full translation process, pixels located around the edges of all matrices are set to zero, as a means of avoiding image distortion and possible correlation issues.

The last stage of the cycle involves the computation of the correlation values between all sample matrices and the reference matrix. This is achieved by comparing pixel intensities on a pixel-by-pixel basis. If the distance between pixel intensities is lower than the threshold integer value defined by the user, the variable responsible for storing the correlation measurement is incremented by one unit. Once all correlation values have been attained, the transformation associated with the highest value would become the starting point for the next iteration, unless the highest overall value belonged to the pre-translation sample, whereas the method would be considered complete. Due to the experimental nature of this algorithm, development was focus around the use of a single cycle, as it would be sufficient to confirm the hypothesis.

Preserving the original state of both sample and reference images is vital throughout the runtime of the algorithm. Direct manipulation of user provided data could result

in irreversible transformations and eventual loss of information, which was deemed as an unacceptable consequence. To circumvent this issue, all transformations are applied to copies of the original data, which are then treated as temporary files within the plugin folder.

Following the completion of the run cycle, the geo-rectification parameters obtained are then employed in the creation of geo-rectified copies of all requested bands, at original resolutions and devoid of any image preparation operations, not including cropping if applied. This final geo-rectification procedure involved the use of the GDAL Tools toolset *GetGeoTransform* and *SetGeoTransform* functions, that retrieve and overwrite the geographical parameters, respectively.

Summary. As previously stated, optimization efforts took a background role as the core design philosophy focused on the creation of a proof of concept. However, this method was expected to have inherent limitation due to its exhaustive nature and inflexible design.

Despite limitless upscaling being a distinct theoretical possibility capable of providing unlimited degrees of subpixel translation, in practice, relying on such methods would exponentially increase the processing power required with every subsequent upsampling factor used in the procedure. Increasing the subpixel precision of this method would not only increase its computational load but could also invalidate the core hypothesis, due to the ever increasingly small magnitudes of the translations.

The magnitude of the geometrical translation errors may also vary between axes, which would require the creation of more matrices increasing the number of correlation operations required to achieve only a slight increase in accuracy.

Advantageously, due to the nature of the correlation process, cloud presence may hardly affect the accuracy of this method. Even with part of the area of interest obstructed by noise, therefore resulting in a lower availability of useful pixels, each matrix will be affected equally and the independent pixel-by-pixel correlation technique should inevitably associate higher results to translations with a higher proximity to the optimal geometric matching.

4.3.2 Local DFT method

By analyzing both Level-1C and Level-2A data products provided by the Sentinel-2 mission, it was concluded in previous subchapters that the residual geometric errors present in these images were mostly, if not completely, composed of two-dimensional rigid translations not surpassing the 1.5-pixel mark on each axis. Further increasing the geometric accuracy of these samples would merely require the use of a method capable of detecting and correcting these simple displacements.

Techniques employing a fast Fourier transform are commonly used as a way of addressing the aforementioned reference issue. These procedures revolve around upscaling and computing a correlation measure by matching the peaks found in the frequency domain in both the sample and reference images. Expectedly, the computational burden is known to increase with higher accuracy requirements, notably in terms of memory (Guizar-sicairos, Thurman, and Fienup 2008).

Despite effective, use of a fast Fourier transform is restricted to the entirety of the upsampled array, likely resulting in massive wastes in terms of computational resources. This waste is often a non-negligible factor, since proper referencing only requires the computation of a small upsampled neighbourhood around the frequency peak location. A method by the name of single-step DTF, proposed by Guizar-Sicairos et al. (Guizar-sicairos, Thurman, and Fienup 2008), aims to reduce computational costs by initially estimating the frequency peak through a fast Fourier transform, using an upscale factor of two, and subsequently computing an upsampled correlation around the initial estimate, using a small 1.5x1.5 pixel neighbourhood. This operation is conducted through the product of three matrices, (1.5α, N), (N, M) and (M, 1.5α) with α being the upscale factor and N and M being the image dimensions, hence this process being known as matrix-multiply discrete Fourier transform (DFT), .

Akin to the single-step DFT approach previously described, a module from the scikit-image Python library was used as a foundation for the correction of the aforementioned geometric translation errors. Capable of subpixel registration, *register_translation* employs a fast Fourier transform to obtain an initial estimate of the peaks in the frequency domain, followed by the first adjustment. Subsequently, the algorithm refines the shift estimation by upsampling the discrete Fourier transform in the neighbourhood of the estimated peaks by means of matrix-multiply DFT.

This method requires the use of two data sets, these being both the sample and reference images. Assuming all required data is present within the system, the algorithm starts by utilizing the *numpy.array* function from the NumPy Python library to convert the data sets into arrays. Followed by this operation is the use of the aforementioned

register_translation module, which receives both arrays and computes the correlation between them on a hardwired upscale factor of 100, resulting in a $1/100^{th}$ of a pixel registration precision. The registration shift vector, provided in pixel scale, is then returned and subsequently used in the final geo-rectification process, Figure 4.10.

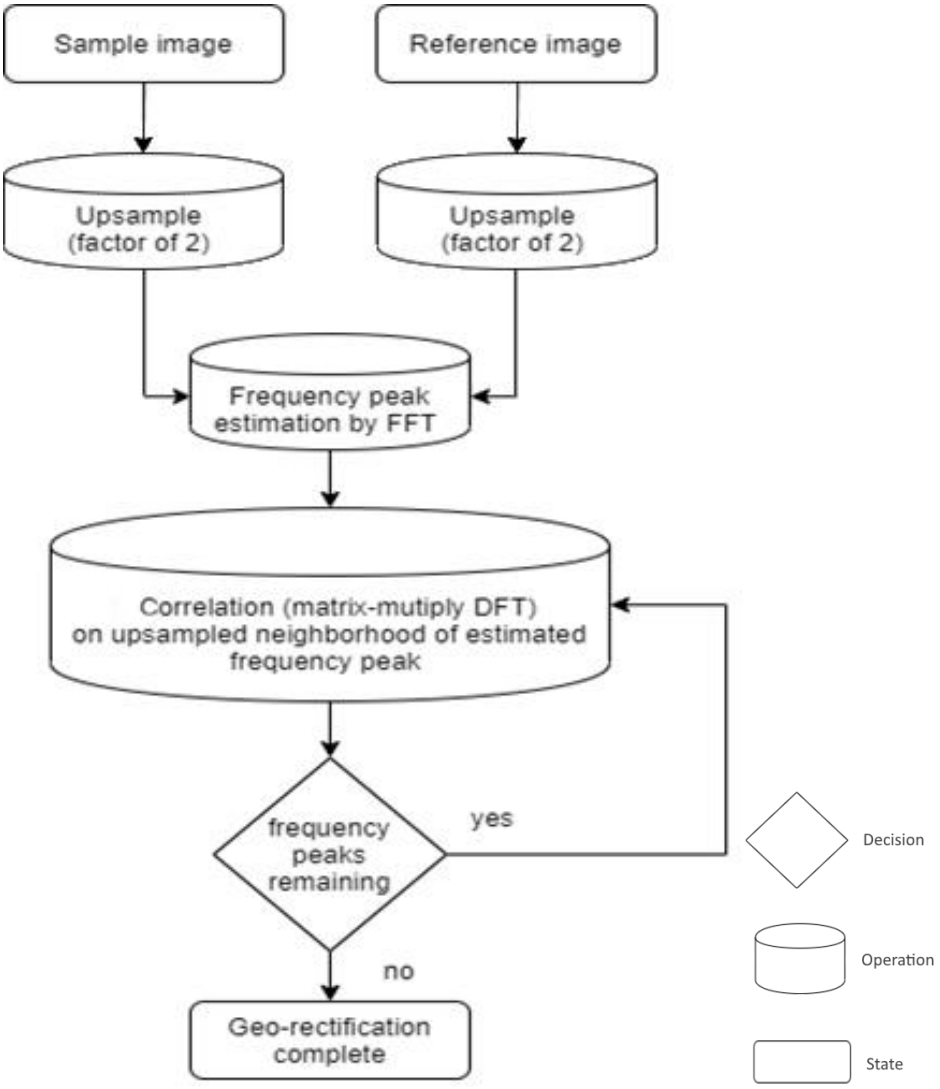


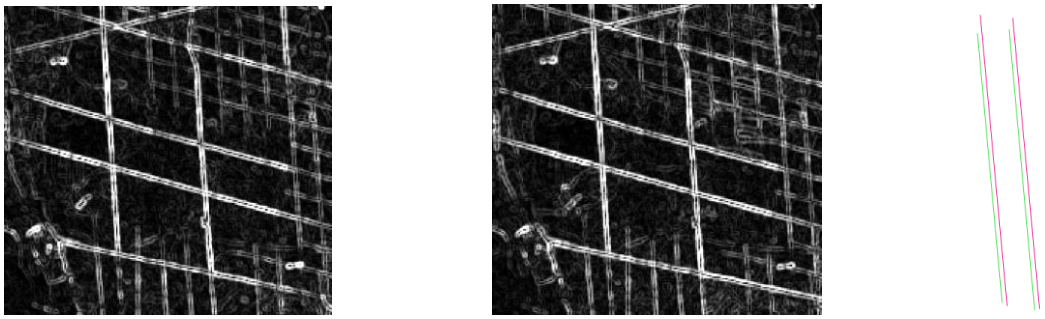
Figure 4.10: Diagram of the local DFT method proposed.

Summary. As an area-based matching method focused on the use of Fourier transforms, it is expected that geometric errors consisting of small to medium sized translations will be sufficiently corrected on a subpixel level. The literature examined for this dissertation indicates that phase correlation methods are capable achieving high accuracy even in the presence non-uniform, time varying lighting discrepancies, which can often be found within the data sets used. The effects of small to medium cloud presence cannot

be predicted at this point. It is speculated that atmospheric noise may negatively affect accuracy, possibly undermining the overall georeferencing process.

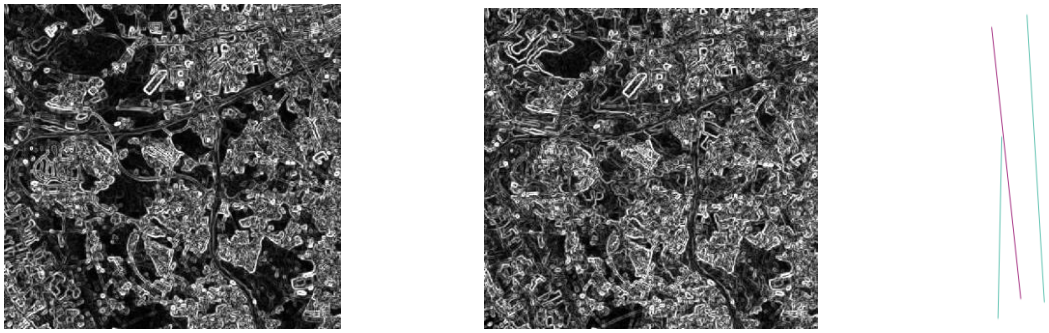
4.3.3 Line segment detection

Not including the main algorithms previously presented, a number of small attempts at designing a geo-rectification system led to the study of a variety of different possible approaches and solutions. One of the earlier geo-rectification approaches examined was centered around the use of feature-based matching. Taking into consideration the intended applications for this system, line features were expected as one of the defining characteristics in both reference and sample imagery.



(a) **(b)** **(c)**

Figure 4.11: Sobel transformed Level-1C imagery band 04 – 2017/04/08 (a), 2019/04/20 (b) and zoom in on LineSegmentDetection output (c).



(a) **(b)** **(c)**

Figure 4.12: Sobel transformed Level-1C imagery band 04 – 2017/04/08 (a), 2019/04/20 (b) and zoom in on LineSegmentDetection output (c).

The main focus was centered around a line detection algorithm capable of detecting roads, rivers, buildings, firebreaks and other similar features for subsequent use in the process of geometric correction. To that end, the LineSegmentDetection feature from the Orfeo ToolBox toolset was examined as a potential solution for the geo-rectification aspect of this dissertation. Designed to detect locally straight lines in a raster file, it is based on the Burns, Hanson and Riseman method (Burns, Hanson, and Riseman 1986) and uses a validation approach proposed by Desolneux et.al (Desolneux, Moisan, and Morel 2000). The customization options are limited to the selection of the input raster file and output file directory, toggle of rescaling and RAM allocation.

Summary. As a way of increasing effectiveness, this method was used together with the Orfeo ToolBox Sobel operand, Figure 4.11. Initial tests indicated that under certain conditions this method resulted in accurate detection of line features, reflecting the geometric discrepancies between the data sets analyzed. However, under sub-optimal conditions the detections became unable to accurately reflect the geometric translations due to the presence of false-positive or misaligned detections, Figure 4.12. The pursuit of geo-rectification based upon line feature detection was eventually abandoned due to its unreliable nature. Contrary to what was initially believed, data sets were not conductive to the sole use of this particular feature-based method.

4.4 Cloud-cover correction

Proper correction of atmospheric noise is not only dependent on the nature and effectiveness of the rectification algorithms but also on the preliminary aspect of cloud detection. Throughout the literature presented in previous chapters, many methods for cloud-cover detection were presented and briefly analyzed. Nonetheless, the focus of this dissertation was centered around the construction of a rectification algorithm which resulted in the acquisition and use of third-party cloud masks.

The design of this algorithm was based exclusively on a multi-temporal approach, excluding any of the non-complementation-based methods. After careful evaluation of the selected image data source, particularly in terms of availability, it was concluded that a temporal-replacement based method could become an adequate solution for the overarching theme of this dissertation, Figure 4.13.

The proposed technique requires manual selection of both data sets, these being the sample and reference images. In this context, the reference image is supposed to contain the cloud free patches used in the reconstruction of the sample, resulting in a cloudless

output. The overall design favours the use of Level-2A imagery by automatically acquiring the cloud mask included in these products.

Assuming all required data is present in the system, the algorithm starts by preparing the selected bands for reconstruction and performing a transformation over the could mask provided, in what amounts to a simple binarization process. This last operation is achieved by use of the raster calculator included in the GDAL Tools toolset, by the name of *gdal_calc*. This function allows the user to perform arithmetic operations such as additions and divisions, as well as simple logic over the pixels of one or more raster images. For the desired effect, every pixel not pertaining to the transparent layer must be normalized, which is accomplished by assuring that every non-zero pixel is set to zero. Following this procedure, one other function from the GDAL Tools toolset is used to convert the cloud mask into a collection of polygons, stored as a temporary shapefile. The *gdal_polygonize* function analyses the raster provided and creates vector polygons from all concentrated regions of pixels sharing a single value, outputting the results to a specified directory.

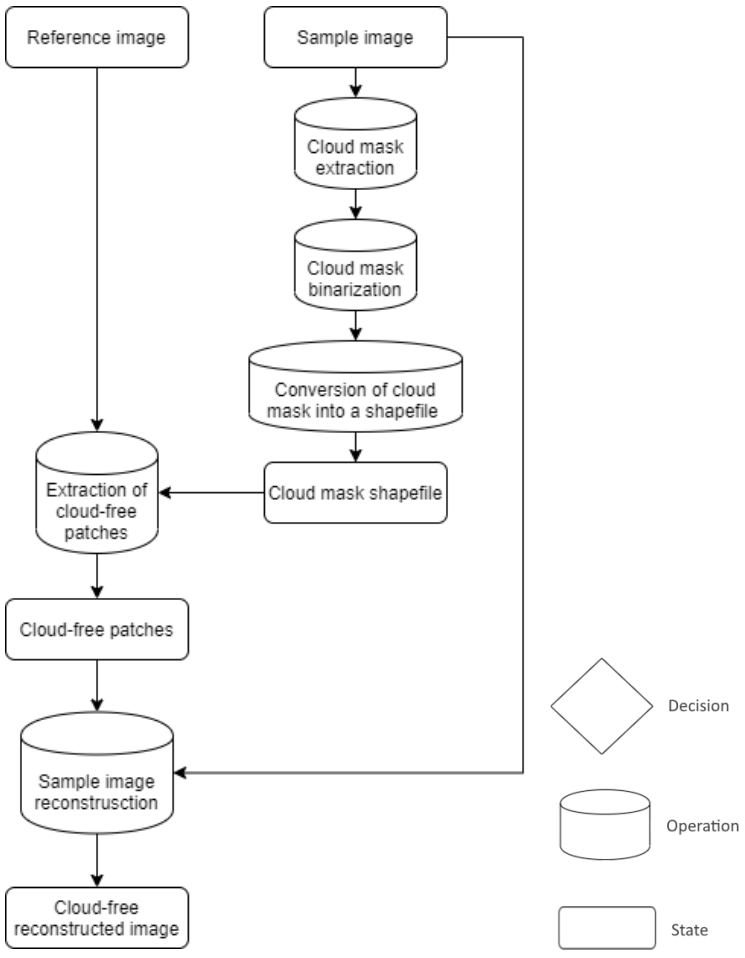


Figure 4.13: Diagram of the proposed multi-temporal cloud-cover correction technique.

After successful conversion of the cloud mask into a shapefile, the *gdalwarp* module is used as a way of extracting all the cloud free patches from the relevant reference image bands. To finalize the process, a mosaicking tool also present in the GDAL Tools toolset is used to merge the appropriate cloudless patches into the various bands previously prepared for reconstruction. The *gdal_merge* function, originally designed for automatic mosaicking, can be employed in the desired reconstruction due to the inclusion of an overlap feature capable of merging two separate rasters.

By the end of the run cycle, all temporary data is removed from the system and the rectification output is stored on the folder specified by the user.

Summary. As previously stated, the quality of cloud detection greatly affects the overall effectiveness of the correction algorithms, particularly so in temporal-replacement based approaches. Poor classification in both sample or reference images will result in reduced accuracy as either cloud patches in the sample avoid detection, and therefore subsequent correction, or as cloud patches become erroneous classified as cloudless in the reference image, leading to the insertion of atmospheric noise. Cloud shadows should also be considered by the detection algorithm, as they too can affect the accuracy of land cover classification applications.

Regarding the correction algorithm presented, temporal-replacement will always be assured, with accuracy being directly tied to both the cloud mask and reference sample selected for the operation.

4.5 Graphical user interface

Throughout development, one of the ever-present focal points was centred around the creation of a graphical user interface (GUI) capable of integrating all aspects and utilities designed for the proposed rectification system.

This subchapter will present the functionalities implemented throughout the course of this dissertation, all of which are ultimately reflected as GUI elements integrated into the final design of the QGIS plugin created. In addition, all extraneous components introduced in previous subchapters, such as data acquisition techniques and overall plugin operations, will be analysed in-dept alongside the aforementioned GUI elements.

4.5.1 Primary window

Included in what is described as *primary window* are all the common elements shared between the designed algorithms. These elements pertain to sample data selection, output pathing and identification, cropping and interpolation method selection and information display.

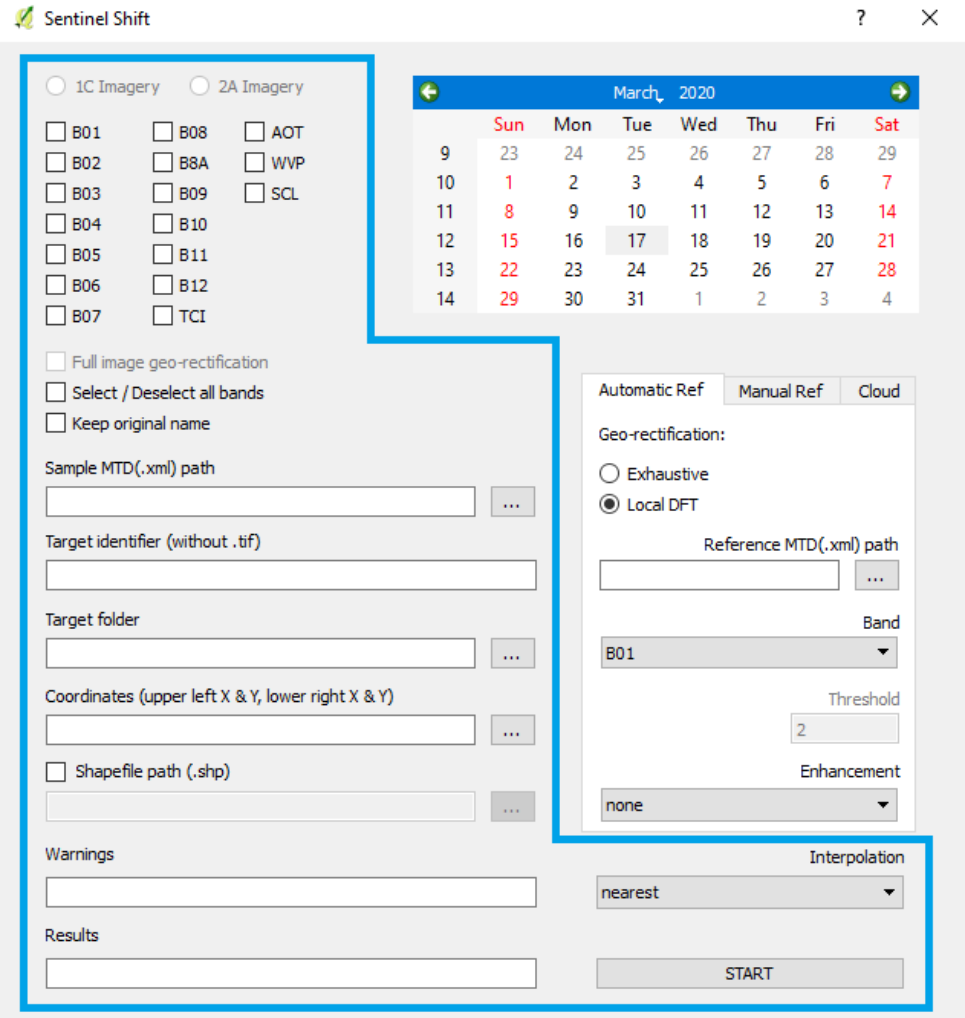


Figure 4.14: Final design of the QGIS plugin with all common elements shared between algorithms encompassed in blue.

Both of the data product levels intended for this system include a variety of bands that can be selected or deselected for processing by checking the checkboxes on the upper left corner of the main plugin window, as can be seen in Figure 4.14. This provides the user with the option to customize the output according to the land cover classification needs, avoiding wasteful use of computational resources. Due to the cumbersome number

of bands available, a “Select / Deselect all bands” checkbox was included to help expedite the selection process.

Data set selection was streamlined by allowing users to point the location of the MTD_MSIL1C.xml or MTD_MSIL2A.xml files (which act as a data product inventory), related to Level-1C and Level-2A imagery respectively. This process can either be done manually by writing the file path directly in the textbox or by clicking the button situated to its left, allowing the user to access a file explorer function. Selection of any unintended file will result in a warning, informing the user of the detected discrepancy. Upon selection of a valid file, the plugin will highlight the product level in the left uppermost corner of the main plugin window.

Sample data is selected through the “Sample MTD(.xml) path” component. After proper selection, the user must define a name for the output imagery by either referring to the “Target identifier” component or by checking the “Keep original name” checkbox. Checking this last box will grey out the previous component in order to avoid conflicting information. Regardless of the naming method selected, output images will always have the respective band name attached to avoid overwriting issues. Finally, the user must select the output file directory by either writing the path directly into the “Target folder” component textbox or by clicking the button to its left, in order to access a file explorer.

By default, the prioritized cropping method is based around the use of a coordinate system. The “Coordinates” component allows the user to select the area of interest by either typing the upper left and lower right X and Y coordinates directly into the textbox or by referring to a text file containing the required information. If manually introduced, coordinates must be separated by a comma and be internally consistent with the geographical system active in QGIS. Any unexpected inputs will result in a warning, informing the user of the detected discrepancy. The “Interpolation” component is directly tied to this cropping procedure, consisting of a dropdown menu featuring the nearest neighbour, bilinear and cubic interpolation algorithms.

Checking the “Shapefile path” checkbox will switch priority towards a polygon-based cropping method. Selecting this option will grey out both “Coordinates” and “Interpolation” components in order to avoid redundant and conflicting information. The user can then either type the shapefile path directly into the textbox or use a file explorer function as a way of referring to the file.

The final elements are non-interactable and serves only as a means of relaying information to the user. Both “Warnings” and “Results” components provide information regarding data input and output respectively.

4.5.2 Geo-rectification tabs

Located on the middle right section of the main plugin window is a group of three tabs, two of which are dedicated solely to geo-rectification procedures, Figure 4.15.

The first of the geo-rectification tabs is focused on automatic geometric correction, including both the exhaustive and local DFT methods presented in previous chapters, hence being referred to as the "Automatic Ref" tab. Elements exclusively pertaining to both these methods are also included in this tab, allowing the user to visualize and access all required input fields for the selected operation. The process for selecting reference data is homologous to the previously presented sample data selection method, requiring the user to reference either the MTD_MSIL1C.xml or MTD_MSIL2A.xml files by manually typing the path location into the "Reference MTD(.xml) path" component textbox or by using the file explorer function. If the reference data selected is not consistent with the previously selected sample, a warning will be issued informing the user of the discrepancy, since both sample and reference images must belong to the same data product level. Selection of a single reference band is also required and can be done so through the "Band" dropdown component.

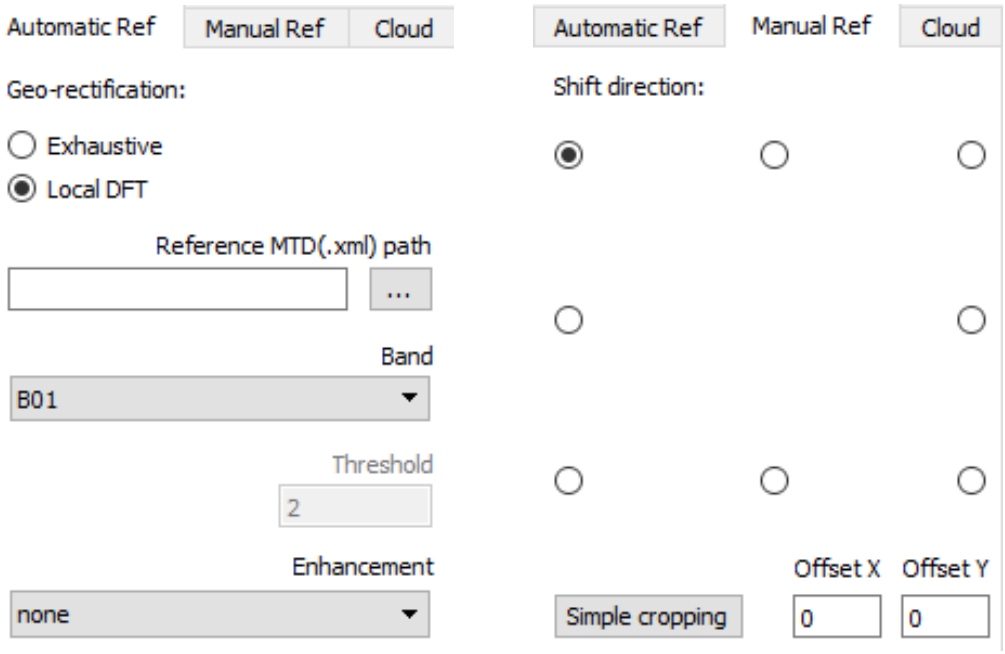


Figure 4.15: Automatic geo-rectification tab (a) and manual geo-rectification tab (b) from the final design of the QGIS plugin

The “Enhancement” component can be used to select any of the image preparation techniques presented in previous chapters. These transformations were designed exclusively as synergising element for the exhaustive algorithm however, if desired they can still be selected alongside the local DTF rectification method.

Associated solely with the exhaustive algorithm is the “Threshold” component, which can be used to alter the correlation threshold value, set as 2 by default. This component will be greyed out unless the “Exhaustive” radio button is selected.

The second geo-rectification tab is focused on manual/artificial geo-rectification, meaning the user is responsible for choosing the direction and magnitude of the desired translation, hence being referred to as “Manual Ref” tab. Contrary to the other main functionalities of the system, manual geo-rectification does not require a reference data set. Instead, the user must select one of the radio buttons representing each of the cardinal and intercardinal directions. The desired magnitude for each individual axis of the translation vector can be typed directly into the “Offset X” and “Offset Y” textboxes, with each unit representing the distance portrayed by a single pixel. Invalid inputs will result in a warning informing the user of the discrepancy detected. Clicking the “Simple cropping” button nullifies the translation vector, which can be used to crop the sample image without any further transformations.

4.5.3 Cloud-cover correction tab

Elements related to cloud-cover correction are exclusively located in the final “Cloud” tab, Figure 4.16. A geo-rectification feature, homologous to the local DFT method found in the “Automatic Ref” tab, is also present and can be selected by checking the “Geo-rectification” checkbox. Choosing this option will geometrically correct the samples prior to the cloud removal process.

The procedure used for reference data selection is identical to the one examined in the previous subsection. Band selection is not a necessary requirement unless the geo-rectification component is selected. In addition, both sample and reference images must be part of the Level-2A data product category.

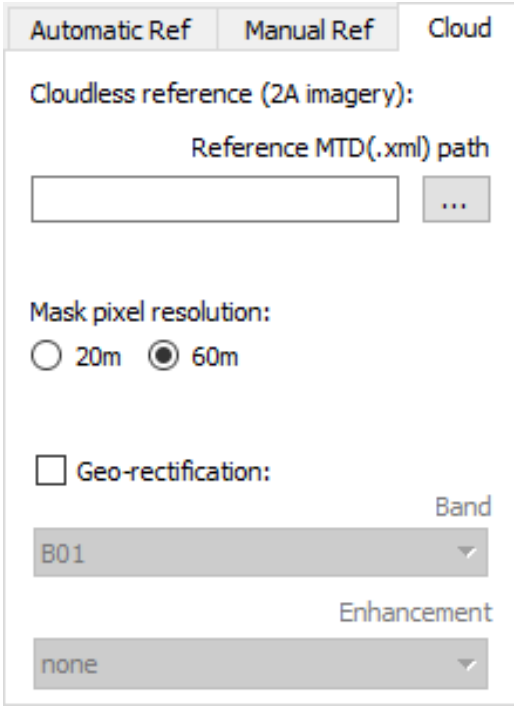


Figure 4.16: Cloud rectification tab from the final design of the QGIS plugin.

As a way of further synergizing with the Earth observation imagery provided by Sentinel-2, the system automatically acquires one of the two cloud masks included in Level-2A data products. A spatial resolution of 60 or 20 meters can be selected through the use of the “Mask pixel resolution” radial buttons, effectively providing the user with a choice between two distinct cloud masks.

5

Results and analysis

As a way of validating and measuring the performance of the methods proposed throughout this dissertation, a series of tests conducted under various conditions will be presented and analyzed. Initially, this chapter will be focused on the validation of the two final geo-rectification algorithms by presenting a series of tests under simulated and real-case scenarios. Testing will be followed by an error and performance estimation, which will be subsequently used in the analysis and comparison of both methods. The chapter will conclude with a demonstration of the cloud-cover correction technique developed, highlighting two separate results obtained through the use of different cloud masks.

5.1 Geo-rectification results and validation

Validation of the geo-rectification algorithms was achieved through the use of two distinct testing methodologies. The first method was designed with the intent of creating a controlled simulation of geometric distortion, namely pure translations, in order to obtain the error estimation for the local DFT method and to assess the accuracy of the exhaustive method. The second validation approach was based on a real-case error estimation, reflecting the subpixel adjustments ultimately required of these algorithms.

5.1.1 Artificial shift adjustment

As was previously stated, this method focuses on the use of artificial image translations as a way of minimizing the number of unknown variables present in the error estimation process. Both reference and sample data sets share the same original source

image, with the latter corresponding to the geometrically altered version. As a result, the effects of varying light intensities will not factor into the accuracy estimation.

Obtaining the various samples was done through a two-step process consisting of a manual translation and subsequent cropping, Figure 5.1 and 5.2. Due to the translations only affecting the geometric data of the images, as opposed to dealing with pixel intensities directly, the cropping tool developed for QGIS was used as a way of obtaining the desired results.

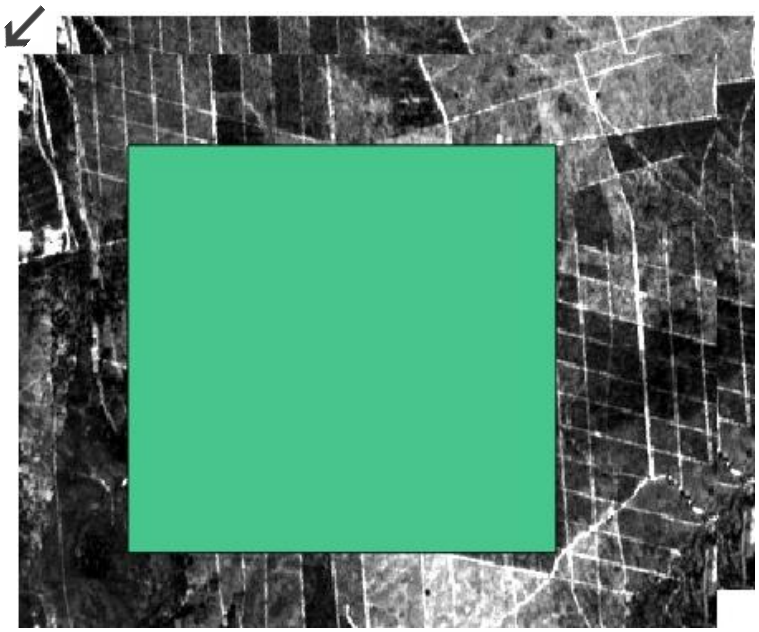
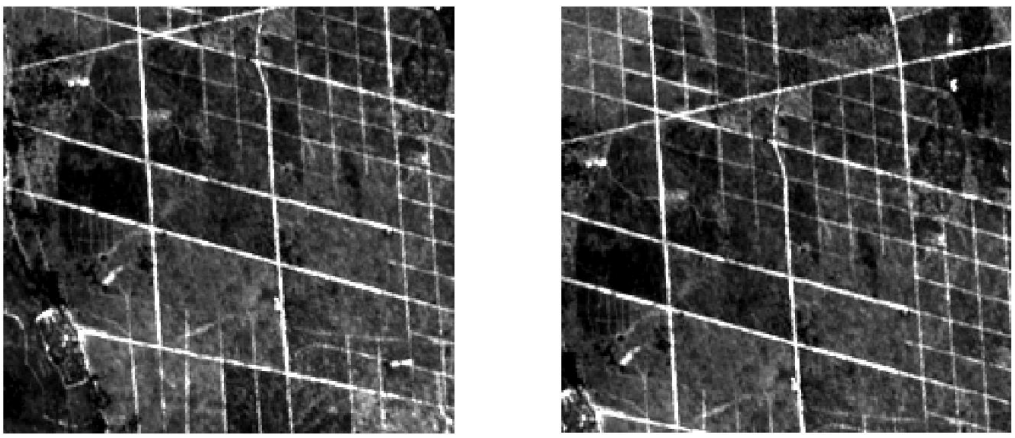


Figure 5.1: Demonstration of the sample acquisition process by translating the reference image in the SW intercardinal direction and subsequently cropping over the polygon shape.



(a) **(b)**

Figure 5.2: Artificial shift adjustment method reference (a) and shifted sample (b).

A total of 30 different tests over three distinct areas were conducted in order to obtain enough data for effective validation of the local DFT method. The error assessment was accomplished through measurement of the distance between the artificial and adjustment shifts by use of the root-mean-square deviation (RMSD) formula.

Table 5.1: Validation of the local DFT method by artificial translation.

ZONE 1	x_shift	y_shift	x_ref	y_ref	exec_time
T1a	0	0	0	0	0.32
T2a	1	1	-0.99	-0.98	0.33
T3a	1	-1	-0.99	0.99	0.32
T4a	3	1	-2.99	-0.98	0.30
T5a	-5	2	4.99	-1.99	0.33
T6a	10	0	-9.99	0	0.31
T7a	-10	-20	9.99	19.99	0.31
T8a	50	0	-49.99	0	0.32
T9a	-50	17	49.99	-16.99	0.33
T10a	70	60	-70	-59.98	0.30
ZONE 2					
T1a	0	0	0	0	0.07
T2a	1	1	-0.96	-0.97	0.12
T3a	1	-1	-0.95	0.96	0.12
T4a	3	1	-2.96	-0.97	0.14
T5a	-5	2	4.97	-1.98	0.14
T6a	10	0	-9.96	0	0.13
T7a	-10	-20	9.97	-19.97	0.12
T8a	50	0	-49.98	-0.01	0.13
T9a	-50	17	49.97	-16.99	0.12
T10a	70	60	-70	-59.95	0.12
ZONE 3					
T1a	0	0	0	0	0.18
T2a	1	1	-0.97	-0.98	0.06
T3a	1	-1	-0.97	0.98	0.25
T4a	3	1	-2.98	-0.98	0.25
T5a	-5	2	4.97	-1.99	0.24
T6a	10	0	-9.99	0	0.25
T7a	-10	-20	9.97	19.98	0.24
T8a	50	0	-50	0	0.25
T9a	-50	17	49.97	-16.99	0.43
T10a	70	60	-70	-59.97	0.25

Table 5.1 includes the same 10 artificial translations over the X and Y axes, represented by the *x_shift* and *y_shift* columns respectively, and the geometric adjustment performed over each axis by the local DFT method, represented by the *x_ref* and *y_ref* columns, in three distinct areas denoted by *zone 1* to *3*, Figures 5.3 and 5.4. Also included in the final column are the execution times (in seconds) for each of the tests performed with the local DFT algorithm, designated by *exec_time.*

By applying the RMSD formula, Equation 5.1, where $\widehat{x_t}$ and x_t, together with $\widehat{y_t}$ and y_t pertain to the X and Y axes value in both the artificial translation and adjustment columns respectively and *T* representing the total number of tests performed, it was possible to obtain an error estimation of 0.0256, representative of the pixel deviation. It is important to note that given the opposite nature of both the artificial and adjustment translations, absolute values for the $\widehat{x_t}$, x_t, $\widehat{y_t}$ and y_t variables were used to preserve the geometric integrity of this estimation.

$$\sqrt{\frac{\sum_{t=1}^{T}((\widehat{x_t}-x_t)^2+(\widehat{y_t}-y_t)^2)}{T}} \quad (5.1)$$

Validation of the exhaustive method was performed under similar circumstances, however, due to the experimental nature of this algorithm some procedures were changed to better reflect and confirm the expected capabilities of this technique. The same reference and sample images used in *zone 1* of the previous local DFT method validation were reused to ensure the highest possible level of consistency.

Table 5.2: Validation of the exhaustive method by artificial translation without the use of image enhancing pre-processing operations.

ZONE 1	x_shift	y_shift	x_ref	y_ref	exec_time
T1a	0	0	0	0	17.43
T2a	1	1	-0.5	-0.5	13.98
T3a	1	-1	-0.5	0.5	13.92
T4a	3	1	-0.5	0	13.69
T5a	-5	2	0.5	-0.5	13.68
T6a	10	0	0.5	-0.5	14.13
T7a	-10	-20	-0.5	0.5	13.58
T8a	50	0	0	-0.5	14.23
T9a	-50	17	0.5	0.5	13.52
T10a	70	60	0	0	14.32

Table 5.3: Validation of the exhaustive method by artificial translation with the use of a pre-processing Sobel operand image transformation.

ZONE 1	x_shift	y_shift	x_ref	y_ref	exec_time
T1a	0	0	0	0	18.80
T2a	1	1	-0.5	-0.5	17.46
T3a	1	-1	-0.5	0.5	17.24
T4a	3	1	0.5	0.5	16.59
T5a	-5	2	0.5	-0.5	16.66
T6a	10	0	-0.5	0.5	16.62
T7a	-10	-20	0.5	0.5	16.58
T8a	50	0	0	0	16.94
T9a	-50	17	0.5	0.5	16.49
T10a	70	60	0	0	17.21

Table 5.4: Validation of the exhaustive method by artificial translation with the use of a pre-processing Otsu binarization.

ZONE 1	x_shift	y_shift	x_ref	y_ref	exec_time
T1a	0	0	0	0	20.46
T2a	1	1	-0.5	-0.5	20.51
T3a	1	-1	-0.5	0.5	20.46
T4a	3	1	-0.5	0	20.94
T5a	-5	2	0.5	-0.5	20.80
T6a	10	0	-0.5	0.5	21.30
T7a	-10	-20	0.5	0.5	21.08
T8a	50	0	0.5	-0.5	21.97
T9a	-50	17	-0.5	0.5	24.00
T10a	70	60	0.5	0	15.99

Table 5.5: Validation of the exhaustive method by artificial translation with the use of a pre-processing Sobel operand image transformation followed by an Otsu binarization.

ZONE 1	x_shift	y_shift	x_ref	y_ref	exec_time
T1a	0	0	0	0	20.44
T2a	1	1	-0.5	-0.5	21.05
T3a	1	-1	-0.5	0.5	21.42
T4a	3	1	0.5	0	21.47
T5a	-5	2	-0.5	0.5	21.59
T6a	10	0	-0.5	0.5	21.30
T7a	-10	-20	-0.5	0.5	21.25
T8a	50	0	0	0.5	21.66
T9a	-50	17	0	0.5	21.30
T10a	70	60	0	0.5	22.36

Tables 5.2 through 5.5 are homologous to Table 5.1 presented earlier, with the *x_shift* and *y_shift* columns representing the artificial translations over the X and Y axes and with *x_ref* and *y_ref* columns representing the respective adjustments calculated by the exhaustive method algorithm. The execution time of each test is represented in seconds in the *exec_time* column.

Due to this technique running only one cross-correlation cycle, the RMSD formula will not be applied as it would prove to be fruitless under the described circumstances. Instead, results were marked green when the adjustments matched the artificial translation in both axes, as opposed to red, when one or both axes did not match the aforementioned translation.

Analysing the overall results, it is possible to conclude that the local DFT method vastly outclasses the exhaustive method both in accuracy and efficiency, proving to be reliable when faced with translations an order of magnitude higher than what can be expected in real-case applications. The exhaustive method seems to consistently lose reliability when presented with higher magnitude translations, however, with translations restricted to a single pixel on each axis, this technique manages to obtain matching results.

5.1.2 Real-case estimation

The second validation method is centered around the geo-rectification of multiple time series of Earth observation imagery, reflecting the conditions of a real-case application. A total of 10 distinct Sentinel-2 Level-1C images, spanning from 2017 to 2019, were used to test both the local DFT and exhaustive methods over the previous three locations, denoted by *zone 1* to *3*.

One of the images, which will be referred to as *image_01*, served as both a reference for all of the following tests and as a sample. The remaining images, denoted by *image_02* through *10*, acted exclusively as samples.

Table 5.6: Validation of the local DFT and exhaustive methods by real-case estimation without the use of image enhancing pre-processing operations.

ZONE 1	x_dft	y_dft	exec_t1	x_exh	y_exh	exec_t2
T1a	0	0	0.25	0	0	19.22
T2a	-0.45	-0.12	0.27	0	0.5	12.92
T3a	-0.67	0.06	0.27	0.5	-0.5	11.46
T4a	-0.57	-0.09	0.26	0.5	-0.5	11.52
T5a	-0.54	-0.06	0.26	0.5	-0.5	11.75
T6a	-0.55	-0.83	0.26	0	0	18.18
T7a	-0.13	-0.3	0.25	0.5	0.5	12.25
T8a	-0.02	-1.54	0.26	0	-0.5	18.83
T9a	-0.15	-1.27	0.25	-0.5	0.5	21.05
T10a	-0.28	-0.55	0.26	0	0	12.76
ZONE 2						
T1b	0	0	0.06	0	0	14.27
T2b	-0.45	-0.11	0.06	0.5	0.5	8.96
T3b	-0.41	0.12	0.06	0.5	-0.5	8.35
T4b	-0.31	-0.09	0.07	0.5	-0.5	8.53
T5b	-0.3	-0.1	0.06	0.5	0.5	8.56
T6b	-0.32	-0.84	0.07	0	-0.5	10.85
T7b	-0.13	-0.31	0.07	0.5	-0.5	8.70
T8b	0.16	-1.53	0.07	0.5	-0.5	11.23
T9b	-0.22	-1.37	0.08	0	-0.5	12.37
T10b	-0.12	-0.54	0.06	-0.5	0.5	8.64
ZONE 3						
T1c	0	0	0.18	0	0	58.52
T2c	-0.43	-0.08	0.18	0	0	39.60
T3c	-0.51	-0.13	0.20	0.5	-0.5	34.55
T4c	-0.41	-0.04	0.18	0.5	0.5	34.80
T5c	-0.39	-0.05	0.19	0.5	0.5	35.18
T6c	-0.37	-0.83	0.19	0.5	0.5	36.71
T7c	-0.14	-0.28	0.19	0	-0.5	46.90
T8c	-0.12	-1.54	0.18	0	-0.5	46.30
T9c	-0.17	-1.32	0.18	0	-0.5	52.97
T10c	-0.16	-0.5	0.18	0	-0.5	46.55

Table 5.7: Validation of the exhaustive method by real-case estimation with the use of a pre-processing Sobel operand image transformation and Otsu binarization.

ZONE 1	x_exhS	y_exhS	exec_t3	x_exhO	y_exhO	exec_t4
T1a	0	0	19.86	0	0	20.67
T2a	0	0	17.70	0	0	23.22
T3a	-0.5	0	17.57	-0.5	0	23.68
T4a	-0.5	0	17.51	-0.5	0	23.81
T5a	-0.5	0	17.63	-0.5	0	21.64
T6a	-0.5	-0.5	17.85	-0.5	-0.5	21.15
T7a	0	-0.5	18.08	0	0	23.17
T8a	0	-0.5	17.52	0	-0.5	21.50
T9a	0	-0.5	17.95	0	-0.5	23.03
T10a	0	-0.5	16.88	0	0	22.34
ZONE 2						
T1b	0	0	13.91	0	0	14.72
T2b	0	0	12.82	0	0	14.35
T3b	0	0	12.54	0	0	13.38
T4b	0	0	12.41	0	0	13.40
T5b	0	0	12.58	0	0	13.96
T6b	0	-0.5	12.62	0	-0.5	14.93
T7b	0	-0.5	12.70	0	0	14.40
T8b	0	-0.5	12.45	0	-0.5	15.21
T9b	0	-0.5	12.88	0	-0.5	15.03
T10b	0	-0.5	12.68	0	-0.5	13.99
ZONE 3						
T1c	0	0	57.65	0	0	61.46
T2c	-0.5	0	52.55	0	0	55.89
T3c	-0.5	0	50.74	-0.5	0	55.70
T4c	0	0	50.80	0	0	56.02
T5c	0	0	52.04	0	0	56.00
T6c	0	-0.5	52.52	0	-0.5	59.88
T7c	0	0	52.02	0	0	55.60
T8c	0	-0.5	51.61	0	-0.5	60.27
T9c	0	-0.5	53.43	0	-0.5	56.35
T10c	0	-0.5	51.43	0	-0.5	56.41

Table 5.8: Validation of the exhaustive method by real-case estimation with the use of a pre-processing Sobel operand image transformation followed by an Otsu binarization.

ZONE 1	x_exhSO	y_exhSO	exec_t5
T1a	0	0	20.64
T2a	0	0	20.85
T3a	-0.5	0	20.75
T4a	-0.5	0	21.05
T5a	-0.5	0	20.79
T6a	0	0	20.70
T7a	0	-0.5	20.89
T8a	0	-0.5	20.92
T9a	0	-0.5	20.82
T10a	0	-0.5	21.24
ZONE 2			
T1b	0	0	15.17
T2b	0	0	15.36
T3b	0	0	15.16
T4b	0	0	15.16
T5b	0	0	15.26
T6b	0	-0.5	15.09
T7b	0	-0.5	15.25
T8b	0	-0.5	15.26
T9b	0	-0.5	15.61
T10b	0	-0.5	15.33
ZONE 3			
T1c	0	0	62.14
T2c	0	0	62.49
T3c	-0.5	0	63.38
T4c	0	0	62.83
T5c	0	0	62.24
T6c	0	-0.5	64.16
T7c	0	-0.5	62.61
T8c	0	-0.5	63.40
T9c	0	-0.5	63.35
T10c	0	-0.5	63.56

Tables 5.6 through 5.8 represent a series of real-case estimations performed by both the local DFT and exhaustive methods. A series of 10 tests per area were conducted, using *image_01* as a reference. The test number is reflective of the sample image used,

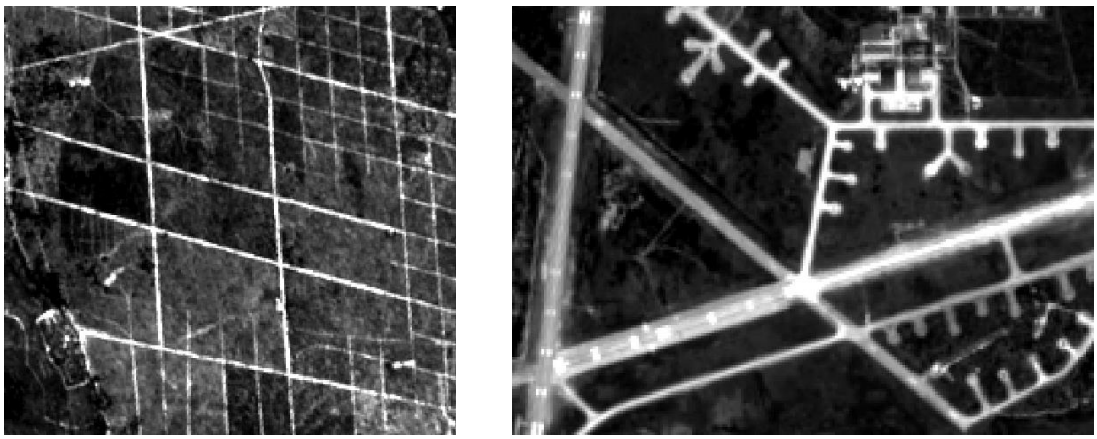
with T1 using *image_01* both as a reference and as a sample, so on and so forth. This sequence was maintained throughout *zone 1* to *3* as a way of maintaining consistency.

Unlike in the previous validation method, where the exact extent of the geometric translation was known a priori, in real-case estimations it is impossible to ascertain the exact adjustment needed for a perfect geometric match. For this reason, by taking into consideration the previous validation results it was assumed that in a real-case estimation the local DFT method will have a relative error around the 0.0256-pixel units. Figures 5.5 through 5.10 provide visual feedback on these transformations. Exhaustive method results were once again marked as green or red depending on the accuracy of the adjustments. As a way of validating the translations calculated, the results provided by the local DFT method, located in *x_dft* and *y_dft* columns of Table 5.6, were used as a baseline for the accuracy measurement.

Despite not using any exact measurement techniques, such as the previously employed RMSD formula, it is still possible to reach a variety of informative conclusions by analysing the data provided by this validation method. Focusing on the local DFT method, and assuming the accuracy rounds the 0.0256-pixel units, it is possible to observe that the adjustments, while sometimes similar between areas, may also vary as much as 0.25-pixel units. It is important to note that the areas used for testing were spread apart no more than 30 kilometres between them. This leads to the conclusion that performing a global geo-rectification on a 100x100 km tile of Sentinel-2 Earth observation imagery through the use of a single small region would prove to be highly unreliable.

Tested under the effects of smaller subpixel translations, the exhaustive method was able to provide consistent matching adjustments, given that both reference and sample images were pre-processed with image enhancing operations. Analysing the results obtained under no enhancements, located in *x_exh* and *y_exh* columns of Table 5.6, the adjustments performed seem to be virtually random, leading to the conclusion that correct matches were mostly incidental in nature, at the exception of tests T1a, T1b and T1c. Of the Sobel operand, Otsu binarization and Sobel-Otsu transformations, represented in the *x_exhS, y_exhS, x_exhO, y_exhO, x_exhSO* and *y_exhSO* columns of Tables 5.7 and 5.8 respectively, the Otsu binarization revealed to better fit the algorithm, having had only one false match in test T10a. Despite its lower accuracy, having had 3 false matches, the Sobel operand proved to have a faster performance, on an average of 3.7 seconds. Expectedly, combining the Sobel operand and Otsu binarization resulted in the slowest execution time, together with an accumulated total of 4 false matches.

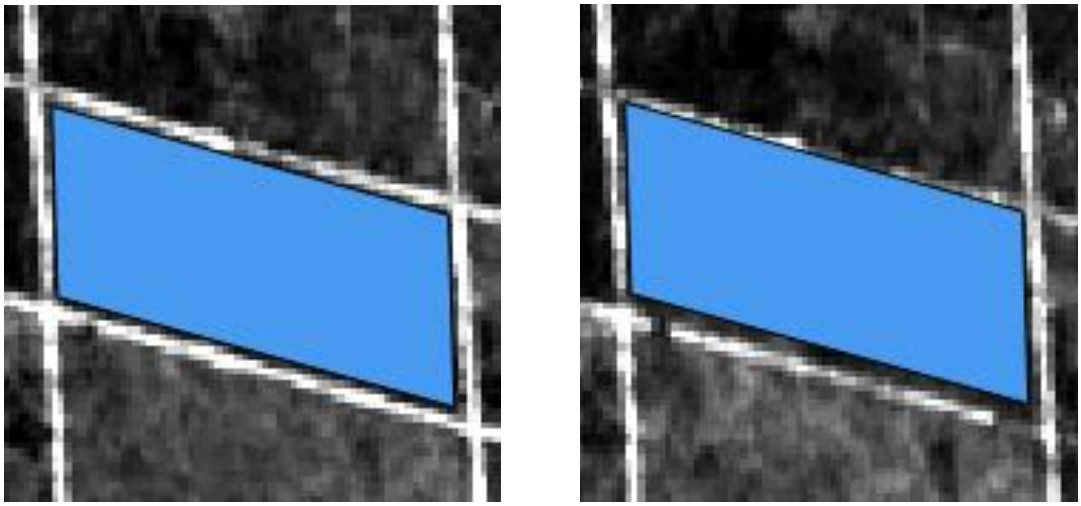
It is also possible to conclude from Figure 5.8, pertaining to test T5b, that the presence of high contrast lighting discrepancies can reduce accuracy, as seen by the ambiguous adjustment performed by the local DFT algorithm.



(a) **(b)**

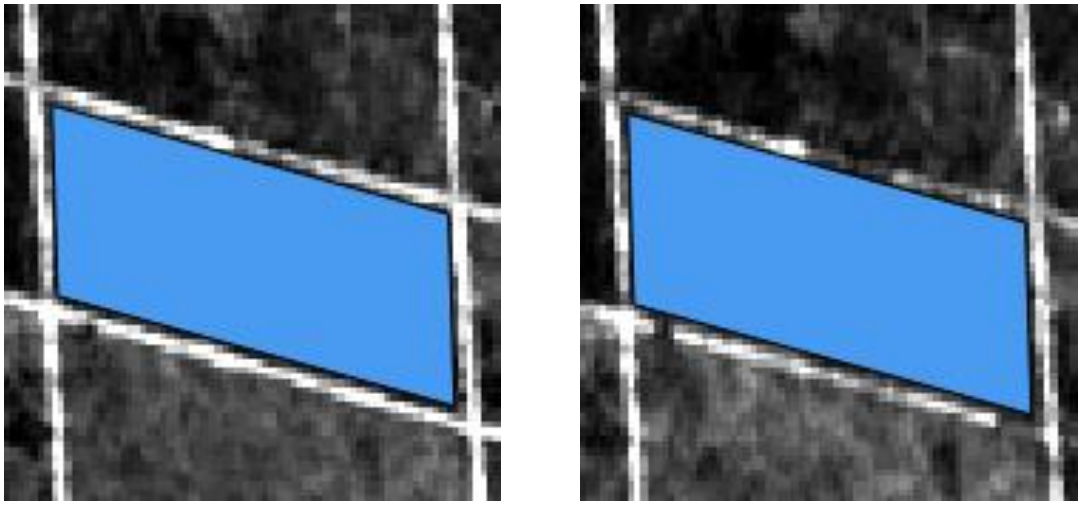
Figure 5.3: Depiction of zone 1 (a) and zone 2 (b) in image_01, band 04.

Figure 5.4: Depiction of zone 3 in image_01, band 04.



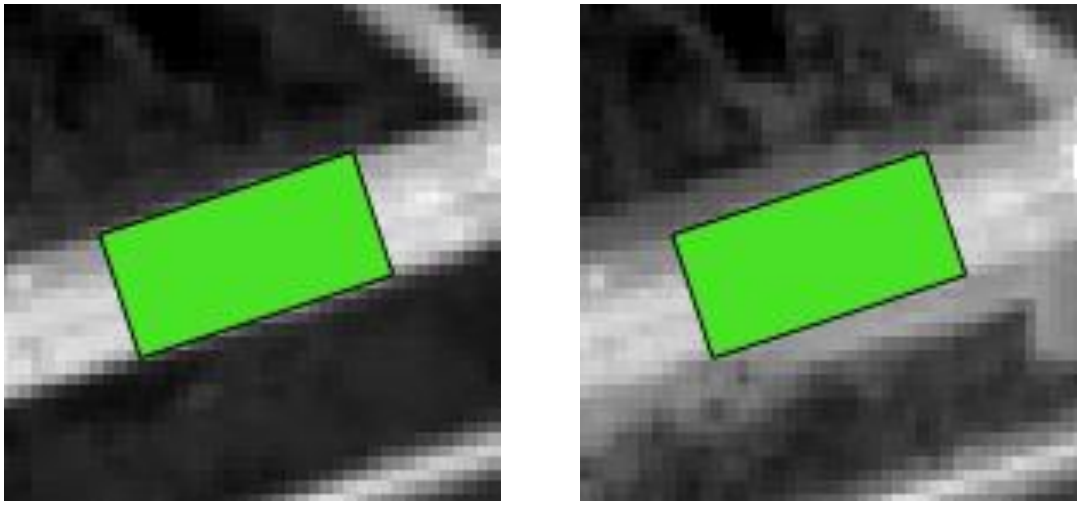
(a) **(b)**

Figure 5.5: Zoom on the reference image_01, band 04 – 2017/04/08 (a) and the sample image_08, band 04 – 2017/10/30 (b), in real-case estimation validation test T8a.



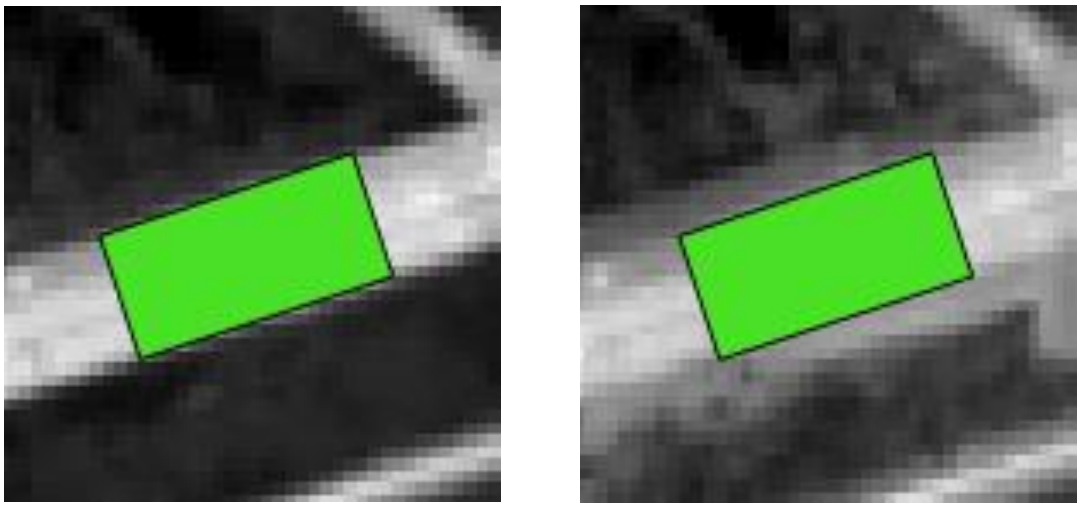
(a) **(b)**

Figure 5.6: Zoom on the reference image_01, band 04 – 2017/04/08 (a) and the adjusted sample image_08, band 04 – 2017/10/30 (b), in real-case estimation validation test T8a.



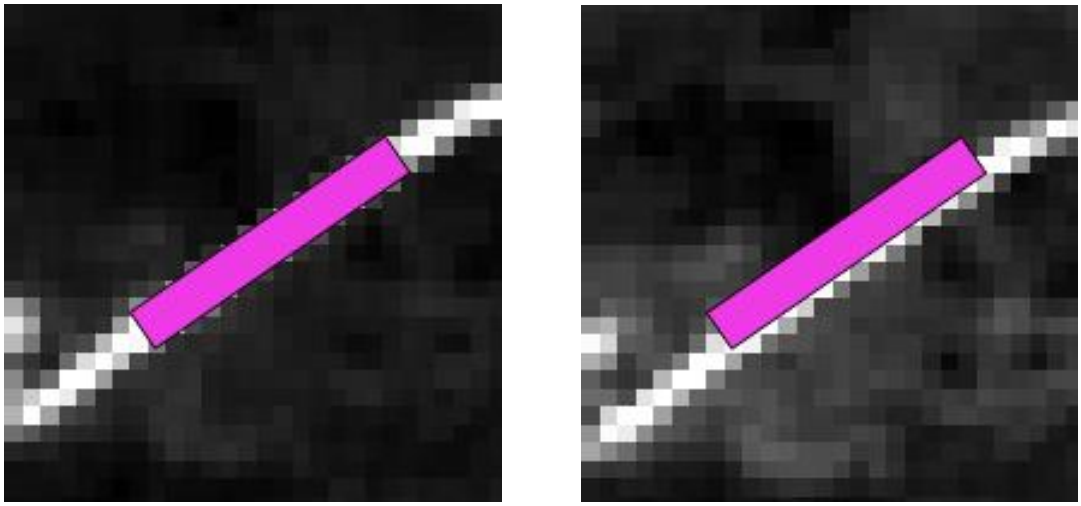
(a) **(b)**

Figure 5.7: Zoom on the reference image_01, band 04 – 2017/04/08 (a) and the sample image_05, band 04 – 2017/09/22 (b), in real-case estimation validation test T5b.



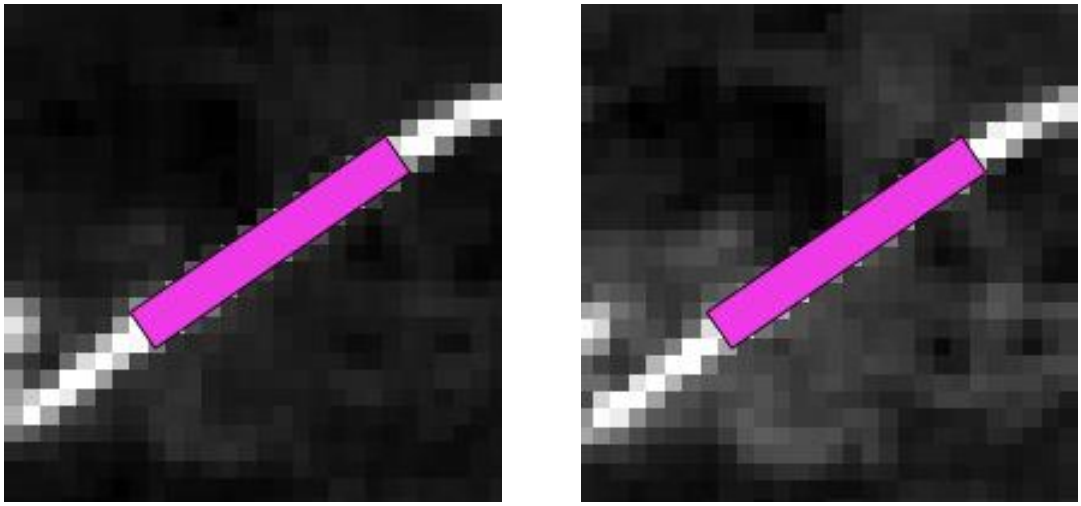
(a) **(b)**

Figure 5.8: Zoom on the reference image_01, band 04 – 2017/04/08 (a) and the adjusted sample image_05, band 04 - 2017/09/22 (b), in real-case estimation validation test T5b.



(a) **(b)**

Figure 5.9: Zoom on the reference image_01, band 04 – 2017/04/08 (a) and the sample image_06, band 04 - 2017/11/11 (b), in real-case estimation validation test T6c.



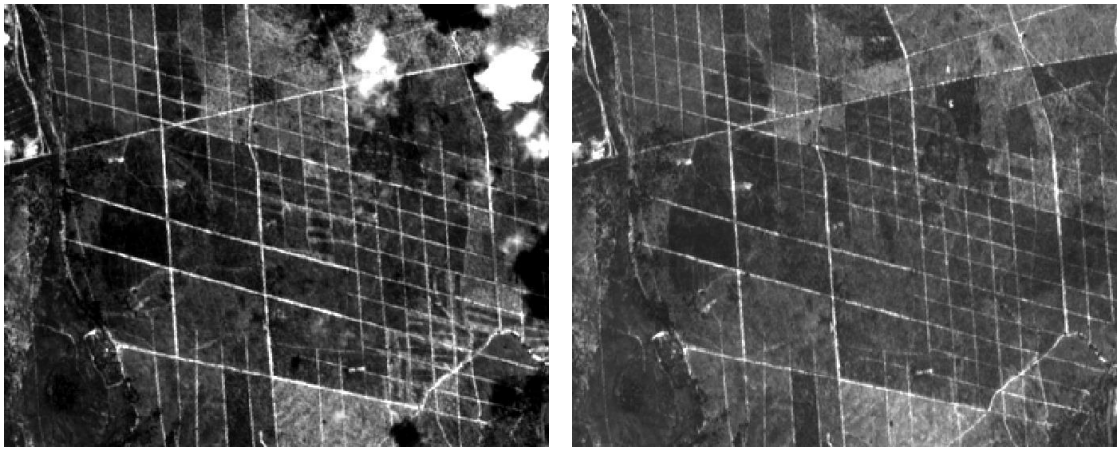
(a) **(b)**

Figure 5.10: Zoom on the reference image_01, band 04 – 2017/04/08 (a) and the adjusted sample image_06, band 04 - 2017/11/11 (b), in real-case estimation validation test T6c.

5.2 Real-case demonstration of cloud-cover correction

As previously stated in subchapter 4.4.1, the accuracy of the proposed cloud-cover correction technique is directly tied to both the cloud mask and reference data set selected for the operation. Due to the use of third-party cloud masks, validation of the algorithm will be performed via practical demonstration.

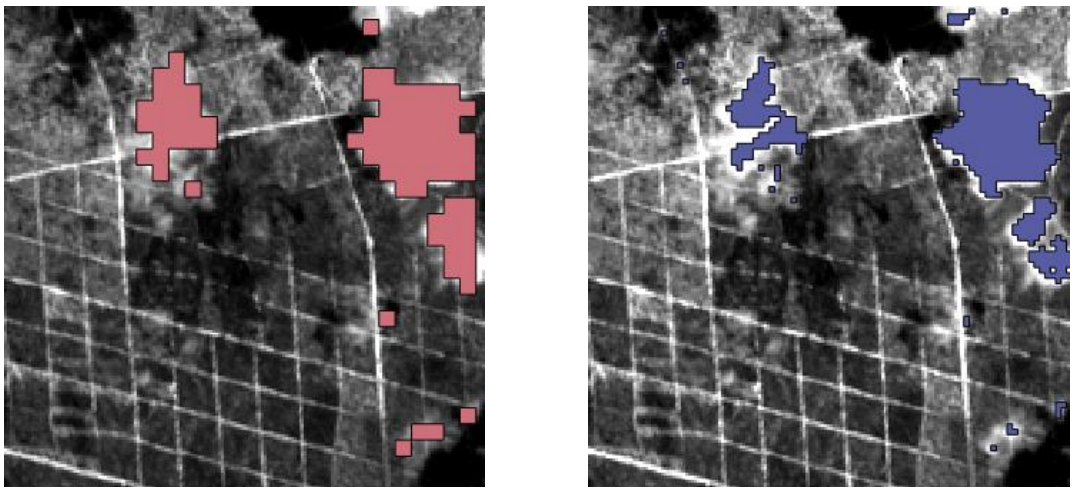
The demonstration will be composed of two individual tests, performed under equal conditions at the exception of the cloud mask used. The tests were conducted with the use of two Sentinel-2 Level-2A Earth observation images, depicted in Figure 5.11, together with two distinct cloud masks, compared in Figure 5.12.



(a) **(b)**

Figure 5.11: Level-2A sample image affected by cloud presence, band 04 – 2019/04/18 (a) and Level-2A cloud-free reference image, band 04 – 2019/04/20 (b).

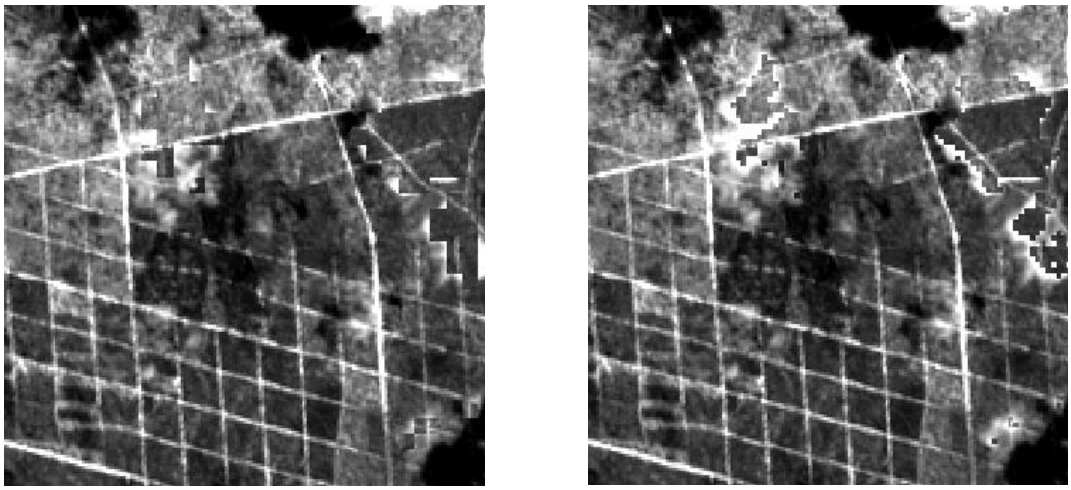
Of the two Level-2A data products used, the cloud-free image, which will be referred to as *image_12*, was used as a reference in all of the tests performed. The remaining sample, referred to as *image_11*, was selected for reconstruction.



(a) **(b)**

Figure 5.12: Sentinel-2 Level-2A cloud mask at a spatial resolution of 60m (a) and 20m (b).

The cloud masks used are part of Sentinel-2 Level-2A data products and were automatically acquired by the algorithm, as specified by the user. The differentiating element between them is the spatial resolution at which they were created, these resolutions being 20 and 60 meters. Despite having a higher resolution, the 20-meter cloud mask often leaves cloud pixels uncovered, something that is somewhat mitigated with the 60-meter variant. False positives may also occur in both masks, usually associated with bright man-made structures. Figure 5.13 depicts a reconstruction of *image_11*, using the 60-meter and 20-meter masks, respectively, to obtain the necessary cloud-free patches from *image_12*.



(a) **(b)**

Figure 5.13: Reconstruction of image_11 through the use of the Sentinel-2 Level-2A 60-meter cloud mask (a) and 20-meter cloud mask (b).

The cloud masks presented, despite being somewhat lacking in accuracy, are able to demonstrate the effectiveness of this technique, and how it can be utilized in the recovery of samples affected by various types of atmospheric noise.

It is also important to note that, by design, Sentinel-2 Level-2A cloud masks are not meant to detect cloud shadows. A brief analysis of Figure 5.13 is enough to ascertain the possible damaging effects of cloud shadows in land cover classification applications. A mask that not only detects clouds but also cloud shadows should be used in tandem with this cloud-cover correction system as to provide the best possible results.

6

Conclusion and future work

Geo-rectification and cloud-cover correction are two themes that expand well beyond what can be condensed into a single dissertation. The overwhelming number of different sensors, sensing platforms, geometric distortions and atmospheric variables have led to the creation of countless methods and approaches, each designed with a core set of strengths. Some of these techniques focus on niche applications, achieving highly accurate results but lacking in adaptability, while others focus on the opposite, obtaining moderate but sometimes sufficient results in a variety of scenarios.

Amidst all of the possible Earth observation sources, this dissertation specifically focused on the creation of a system capable of geo-rectifying and removing cloud-cover from Setinel-2 Level-1C and -2A data products, due to their innately high temporal resolution and geometric accuracy.

Throughout the analysis of these Earth observation products, it was possible to conclude that geometric anomalies were restricted to pure translations of no more than 1.5-pixel units per axis. Focus shifted towards the exploration of methods based in ABM, specifically ones centred around FFT techniques, which were described in most literature as cost efficient solutions for the correction of small to medium sized translations. These methods would also have the added benefit of being robust against multi-temporal non-uniform lighting discrepancies.

From this approach, a technique based on matrix-multiply DFT was originated. This method proved to be computationally light and accurate in the geo-rectification of Earth observation samples in a variety of different biomes and environments, showing only ambiguous judgements in the presence of major non-uniform lighting discrepancies.

Together with the matrix-multiply DFT method, an experimental technique loosely based around sequential similarity detection was proposed in order to provide a different perspective on the geo-rectification aspect of this dissertation. Despite being outclassed in every aspect by the FFT based approach, this technique managed to properly detect (on both axes) the direction of the displacement between sample and reference images. Positive results were relegated to pixel and subpixel translations of no more than 1.5-pixel units. The use of pre-processing image enhancements on both data sets, such as an Otsu binarization or Sobel operand, was also required to achieve the desired results.

Throughout the validation process it was also possible to conclude that Sentinel-2 Level-1C and -2A data products tiles cannot be effectively geo-rectified through the use of global mapping models, due to localized translations. Despite having no more than a 30-kilometre distance between them, test areas revealed to have a geo-location discrepancy of up to 0.25-pixel units. Such a deviation can hamper automatic land cover classification of small areas, as is the case with firebreak monitoring.

Cloud-cover correction through the use of multi-temporal fusion was successfully implemented, with the system being able to provide imagery akin to Level-3 data products, as defined by CEOS. Accuracy is heavily dependent on both the cloud mask and reference image selected for the process. The reference should ideally contain no cloud-cover overlap with the sample and the cloud mask must be able to correctly identify all types of atmospheric noise, these being *thick* and *thin* clouds as well as could shadows.

In summary, the culminating result of this dissertation can be condensed under the following conclusions:

- Data analysis was crucial to the development of a computationally light and accurate system for the geo-rectification and cloud-cover correction of Earth observation imagery;
- Integration of previously developed, open source software tools can be effectively used in the creation of a new modular product;
- Methods based on FFT and SSDA can be used effectively in the geo-rectification of small subpixel translation errors;
- Cloud-cover correction methods based on multi-temporal fusion can be an effective, computationally light and accurate solution in the presence of high temporal-resolution imagery.

Despite providing accurate and satisfactory results, the system developed can still be improved in both user friendliness and algorithm robustness and efficiency.

As stated throughout the literature, some of the proposed solutions for cloud-cover correction have started using a multi-disciplinary approach, combining multiple methods to help mitigate the weaknesses inevitably associated with any given sole technique.

Combining multiple techniques comes at the expense of a higher computational cost, which somewhat clashes with the core design philosophy behind this dissertation. However, despite the focus on the creation of a computationally light and accurate cloud-cover correction system, it still remains to be seen how far the algorithm can be pushed in terms of complexity until computational requirements become too high of a factor. Integration of a cloud classification algorithm can also be explored as a way of increasing the autonomy of the system, reducing the dependence on third-party cloud masks.

In terms of the plugin developed, further incorporation of QGIS utilities can still be achieved, increasing the number interactions and functionalities provided to the user, such as selection of the cropping area directly through the use of *Map Canvas*, direct inclusion of output files on the *Layers Panel*, and the option to automatically acquire cloud masks provided by the user.

In summary, future work should focus on:

- Exploring the possibility of integrating multi-spectral and/or integration-prediction methods to hopefully improve robustness of cloud-cover correction when multi-temporal reference availability is low;
- Increasing system autonomy through possible integration of a cloud classification algorithm;
- Enhancing the plugin created with further integration of QGIS functionalities.

7

References

Abraham, Werner, and Elisabeth Leiss. 2006. “Personal and Impersonal Passives: Definite vs. Indefinite Diatheses.” *Transactions of the Philological Society* 104 (2): 259–96.

A. Mora, T. Santos, S. Łukasik, J. Silva, A. Falcão, J. Fonseca, R. Ribeiro, Land Cover Classification from Multispectral Data Using Computational Intelligence Tools: A Comparative Study, Information. 8 (2017) 147.

Anuta, P. E. 1970. “Spatial Registration of Multispectral and Multitemporal Digital Imagery Using Fast Fourier Transform.” *IEEE Transactions on Geoscience Electronics* 8: 353–68.

Appledorn, C. R. 1996. “A New Approach to the Interpolation of Sampled Data.” *IEEE Transactions on Medical Imaging* 15: 369–376.

Barrow, H.G., Tenenbaum, J.M., Bolles, R.C., Wolf, H.C.: Parametric correspondence and chamfer matching: Two new techniques for image matching. In: Proc. International Joint Conference on Artificial Intelligence, vol. 2, pp. 659–668 (1977)

Bay, Herbert, Tinne Tuytelaars, and Luc Van Gool. 2008. “SURF: Speeded Up Robust Features” *Computer Vision and Image Understanding* 110 (3): 346-59.

Bjorkman, M., N. Bergstrom, and D. Kragic. 2014. “Detecting, Segmenting and Tracking Unknown Objects Using Multi-Label MRF Inference.” *Comput. Vision Image Understand.* 118: 111–27.

Borgefors, Gunilla. 1998. “Hierarchical Chamfer Matching a Parametric Edge Matching Algorithm.” *IEEE Transaction on Pattern Analysis and Machine Intelligence.* 10: 849-865.

Bracewell, R. N. 1965. “The Fourier Transform and Its Applications.” McGraw-Hill.

Burns, J. B., A. R. Hanson, and E. M. Riseman. 1986. “Extracting Straight Lines.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8 (4): 425–55.

Canny, John. 1986. “A Computational Approach to Edge Detection,” *IEEE Transaction on Pattern Analysis and Machine Intelligence* PAMI-8 (6): 679-98.

Castro, E. D., and C. Morandi. 1987. “Registration of Translated and Rotated Images Using Finite Fourier Transform.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9: 700–703.

Chandran, Geethu, and Christy Jojy. 2015. “A Survey of Cloud Detection Techniques For Satellite Images,” *International Research Journal of Engineering and Technology* 2: 2485–90.

Chen, Q., M. Defrise, and F. Deconinck. 1994. “Symmetric Phase-Only Matched Filtering of Fourier–Mellin Transform for Image Registration and Recognition.” *IEEE Transactions on Pattern Analysis and Machine Intellingence* 16: 1156–1168.

Cheng, Qing, Huanfeng Shen, Liangpei Zhang, Qiangqiang Yuan, and Chao Zeng. 2014. “Cloud Removal for Remotely Sensed Images by Similar Pixel Replacement Guided with a Spatio-Temporal MRF Model.” *ISPRS Journal of Photogrammetry and Remote Sensing* 92: 54–68.

Christodoulou, Christodoulos I., Silas C. Michaelides, and Constantinos S. Pattichis. 2003. “Multifeature Texture Analysis for the Classification of Clouds in Satellite Imagery.” *IEEE Transactions on Geoscience and Remote Sensing* 41 (11 PART I): 2662–68.

Desolneux, Agnès, Lionel Moisan, and Jean-Michel Morel. 2000. “Meaningful Alignments.” *International Journal of Computer Vision* 40: 7–23.

Dodgson, Neil A. 1997. “Quadratic Interpolation for Image Resamplinga” *IEEE TRANSACTIONS ON IMAGE PROCESSING* 6 (9): 0–6.

Dreschler, L., and H. Nagel. 1981. “Volumetric Model and 3-D Trajectory of a Moving Car Derived from Monocular TV-Frame Sequence of a Street Scene.” *Proceedings of the Interantional Joint Conference on Artificial Intelligence*, 692–697.

Du, Yong, Bert Guindon, and Josef Cihlar. 2002. “Haze Detection and Removal in High Resolution Satellite Image with Wavelet Analysis.” *IEEE Transactions on Geoscience and Remote Sensing* 40 (1): 210–16.

Ehlers, M., and D. N. Fogel. 1994. “High-Precision Geometric Correction of Airborne Remote Sensing Revisited: The Multiquadric Interpolation.” *Proceedings of SPIE: Image and Signal Processing for Remote Sensing* 2315: 814–824.

Enomoto, Kenji, Ken Sakurada, Weimin Wang, Hiroshi Fukui, Masashi Matsuoka, Ryosuke Nakamura, and Nobuo Kawaguchi. 2017. “Filmy Cloud Removal on Satellite Imagery with Multispectral Conditional Generative Adversarial Nets.” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* 2017–July: 1533–41.

Feng, C., J. -W. Ma, Q. Dai, and X. Chen. 2004. “An Improved Method for Cloud Removal in ASTER Data Change Detection.” *Proc. IEEE Int. Geoscience and Remote Sensing Sym. (IGARSS 2004)* 5: 3387–3389.

Fidrich, Â, Mike A Smith, Elizabeth Guest, Elizabeth Berry, and Richard A Baldock. 2001. “Robust Point Correspondence Applied to Two- and Three-Dimensional Image Registration” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23 (2): 165-79.

Flusser, Jan. 1992. “An Adaptive Method for Image Registration,” *Pattern Recognition* no. 1: 0–9.

Flusser, Jan, and T. Suk. 1993. “Pattern Recognition by Affine Moment Invariants.” *Pattern Recognition* 26: 167–174.

———. 1994. “A Moment-Based Approach to Registration of Images with Affine Geometric Distortion.” *IEEE Transactions on Geoscience and Remote Sensing* 32: 382–387.

Flusser, Jan, and Barbara Zitova. 2003. “Image Registration Methods : A Survey” *Image and Vision Computing* 21: 977–1000.

Forstner, W., and E. Gulch. 1986. “Precise, A Fast Operator for Detection and Location of Distinct Points, Corners and Centers of Circular Features.” *ISPRS Workshop on Fast Processing of Photogrammetric Data*, 281–305.

Gonzalez, Claudia I, and Patricia Melin. 2017. “Edge Detection Method Based on General Type-2 Fuzzy Logic Applied to Color Images,” *Fuzzy Logic for Image Processing*.

Goshtasby, A. 1986. “Piecewise Linear Mapping Functions for Image Registration.” *Pattern Recognition* 19: 459–466.

———. 1987. “Piecewise Cubic Mapping Functions for Image Registration.” *Pattern Recognition* 20: 525–533.

———. 1988a. “Image Registration by Local Approximation Methods.” *Image and Vision Computing* 6: 255–261.

———. 1988b. “Registration of Images with Geometric Distortions.” *IEEE Transactions on Geoscience and Remote Sensing1* 26: 60–64.

Goshtasby, A., and G. C. Stockman. 1985. “Point Pattern Matching Using Convex Hull Edges.” *Man and Cybernetics* 15: 631–637.

———. 1986. “A Region-Based Approach to Digital Image Registration with Subpixel Accuracy.” *IEEE Transactions on Geoscience and Remote Sensing* 24: 390–399.

Guizar-sicairos, Manuel, Samuel T Thurman, and James R Fienup. 2008. “Efficient Subpixel Image Registration Algorithms” 33 (2): 156–58.

Hagolle, O., M. Huc, D. Villa Pascual, and G. Dedieu. 2010. “A Multi-Temporal Method for Cloud Detection, Applied to FORMOSAT-2, VENμS, LANDSAT and SENTINEL-2 Images.” *Remote Sensing of Environment* 114 (8): 1747–55.

Hashimoto, Shutaro, Takeo Tadono, Masahiko Onosato, and Masahiro Hori. 2013. “GEO-REFERENCING OF AERIAL IMAGERY USING EMBEDDED IMAGE DENTIFIERS AND CROSS-REFERENCED DATASETS.” *International Geoscience and Remote Sensing Symposium (IGARSS)*, 3333–36.

He, Wei, and Naoto Yokoya. 2018. “Multi-Temporal Sentinel-1 and -2 Data Fusion for Optical Image Simulation,” 1–11.

Hedley, Mark. 1998. “Fast Corner Detection” *Image and Vision Computing*.

Helmer, E.H., and B. Ruefenacht. 2005. “Cloud-Free Satellite Image Mosaics with Regression Trees and Histogram Matching.” *Photogrammetric Engineering & Remote Sensing* 71 (9): 1079–89.

Hou, H. S., and H. C. Andrews. 1978. “Cubic Splines for Image Interpolation and Digital Filtering.” *IEEE Transactions on Acoustic, Speech and Signal Processing* 26: 508–517.

Hu, Ming-Kuei. 1962. “Visual Pattern Recognition by Moment Invariants,” *IRE Transactions on Information Theory* 66–70.

Huang, Lei, and Zhen Li. 2010. “Feature-Based Image Registration Using the Shape Context.” I*nternational Journal of Remote Sensing* 31: 2169-2177.

Huttenlocher, Daniel P, Gregory A Klanderman, and William J Rucklidge. 1993. “Comparing Images Using the Hausdorff Distance” *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE* 15 (9): 850–63.

J.E. Pereira-Pires, V. Aubard, R.A. Ribeiro, J.M. Fonseca, J.M.N. Silva, A. Mora, Semi-Automatic Methodology for Fire Break Maintenance Operations Detection with Sentinel-2 Imagery and Artificial Neural Network, Remote Sens. 12 (2020) 909.

Jenkinson, Mark, and Stephen Smith. 2001. “A Global Optimisation Method for Robust Affine Registration of Brain Images” *Medical Image Analysis* 5: 143–56.

Keys, Robert G. 1981. “Cubic Convolution Interpolation for Digital Image Processing,” *IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING* no. I: 1153–60.

Lehmann, T. M., C. Gonner, and K. Spitzer. 2001. “Addendum: B-Spline Interpolation in Medical Image Processing.” *IEEE Transaction on Medical Imaging* 20: 660–665.

Leutenegger, Stefan, Margarita Chli, and Roland Y Siegwart. n.d. “BRISK: Binary Robust Invariant Scalable Keypoints Stefan.”.

Li, H., B. S. Manjunath, and S. K. Mitra. 1995. “A Contour-Based Approach to Multisensor Image Registration.” *IEEE Transactions on Image Processing* 4: 320–334.

Li, Pengfei, Kaimin Sun, Deren Li, Haigang Sui, and Yong Zhang. 2017. “An Emergency Georeferencing Framework for GF-4 Imagery Based on GCP Prediction and Dynamic RPC Refinement,” no. 500

Liew, S. C., M. Li, and L. K. Kwoh. 2004. “Automated Production of Cloud-Free Cloud-Shadow Image Mosaics from Cloudy Satellite Imagery.” *XXth ISPRS Congr.*, 12–13.

Lin, Chao Hung, Po Hung Tsai, Kang Hua Lai, and Jyun Yuan Chen. 2013. “Cloud Removal From Multitemporal Satellite Images Using Information Cloning.” *IEEE Transactions on Geoscience and Remote Sensing* 51 (1): 232–41.

Liu, S J, and X H Tong. 2008. “TRANSFORMATION BETWEEN RATIONAL FUNCTION MODEL AND RIGOROUS SENSOR MODEL FOR HIGH RESOLUTION SATELLITE IMAGERY.” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXVII (Part B1): 873-77.

Long, Tengfei, Weili Jiao, Guojin He, and Zhaoming Zhang. 2016. “A Fast and Reliable Matching Method for Automated Georeferencing of Remotely-Sensed Imagery” *Remote Sensing* 8 (1): 56.

Lowe, David G. 2004. “Distinctive Image Features from Scale-Invariant Keypoints” *International Journal of Computer Vision* 60: 91-110.

Maalouf, Aldo, Philippe Carré, Bertrand Augereau, and Christine Fernandez-Maloigne. 2009. “A Bandelet-Based Inpainting Technique for Clouds Removal From Remotely Sensed Images.” *IEEE Transactions on Geoscience and Remote Sensing* 47 (7): 2363–71.

Maramatsu, K., and N. Fujiwara. 1999. “Automated Detection and Removal of Clouds and Their Shadows from Landsat TM Images.” *IEICE Trans. Inform. Syst. E82-D* 2: 453–460.

Marr, D., and Hildreth. 1980. “Theory of Edge Detection” *The Royal Society* 217: 187–217.

Mathur, Shashank, and Anil Ahlawat. 2008. “APPLICATION OF FUZZY LOGIC ON IMAGE EDGE DETECTION,” *Intelligent Technologies and Applications* 24: 24–28.

Meijering, Erik H.W., Karel J. Zuiderveld, and Max A. Viergever. 1999. “Image Reconstruction by Convolution with Symmetrical Piecewise Th-Order Polynomial Kernels.” *IEEE Transactions on Image Processing* 8 (2): 192–201.

Melgani, Farid. 2006. “Contextual Reconstruction of Cloud-Contaminated Multitemporal Multispectral Images” 44 (2): 442–55.

Modestino, J. 1979. “Image Enhancement by Stochastic Homomorphic Filtering.” *IEEE Trans. Acoust. Speech Signal Process.* 27: 625–637.

Montesinos, P, V Gouet, R Deriche, and D Pele. 2000. “Matching Color Uncalibrated Images Using Differential Invariants” *Image and Vision Computing* 18: 659–71.

Nguyen, Thanh. 2017. “Optimal Ground Control Points for Geometric Correction Using Genetic Algorithm with Global Accuracy” *European Journal of Remote Sensing* 48*: 101-20*.

Noble, A. 1988. “Finding Corners.” *Image and Vision Computing* 6: 121–28.

Oniga, Valeria-ersilia, Ana-ioana Breaban, and Florian Statescu. 2018. “Determining the Optimum Number of Ground Control Points for Obtaining High Precision Results Based on UAS Images,” *Proceedings* 2 (7): 352.

Pan, Hongbo, Chao Tao, and Zhengrong Zou. 2016. “Precise Georeferencing Using the Rigorous Sensor Model and Rational Function Model for ZiYuan-3 Strip Scenes with Minimum Control.” *ISPRS Journal of Photogrammetry and Remote Sensing* 119: 259–66.

Parker, J. A., R. V. Kenyon, and D. E. Troxel. 1983. “Comparison of Interpolating Methods for Image Resampling.” *IEEE Transactions on Medical Imaging* 2: 31–39.

Peli, T. 1981. “An Algorithm for Recognition and Localization of Rotated and Scaled Objects.” *Proceedings of the IEEE* 69 (4): 483–485.

Peng, L., F. Chen, Z. Zhao, and D. Yan. 2005. “Clouds and Cloud Shadows Removal from High-Resolution Remote Sensing Images.” *Proc. IEEE Geosci. Remote Sens. Symp.* 6: 4256–4259.

Pérez, Patrick, Michael Gangnet, and Andrew Blake. 2016. “Poisson Image Editing.” *Image Processing On Line* 5: 300–325.

Pratt, W.K. 1974. “Correlation Techniques of Image Registration.” *IEEE Transactions on Aerospace and Electronic Systems* 10: 353–358.

Ritter, Nicola, Robyn Owens, James Cooper, Robert H Eikelboom, and Paul P Van Saarloos. 1999. “Registration of Stereo and Temporal Images of the Retina” *IEEE TRANSACTIONS ON MEDICAL IMAGING* 18 (5): 404–18.

Rossow, W. B., R. J. Gurney, J. L. Foster, and C. L. Parkinson. 1993. “‘Clouds,’ in Clouds Atlas of Satellite Observations Related to Global Change.” *U.K. Cambridge Univ. Press*.

Rosten, Edward, Reid Porter, and Tom Drummond. 2010. “Faster and Better: A Machine Learning Approach to Corner Detection.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32 (1): 105–19.

Rublee, E, V Rabaud, and K Konolige. 2011. “ORB : An Efficient Alternative to SIFT or SURF About Local Feature and Matching Motivation OFAST – Oriented FAST BRIEF (Calonder et Al . 2010).” *Intl. Conf. Computer Vision*, 1–5.

Saitwal, Kishor, Mahmood R. Azimi-Sadjadi, and Donald Reinke. 2003. “A Multichannel Temporally Adaptive System for Continuous Cloud Classification From Satellite Imagery.” *IEEE Transactions on Geoscience and Remote Sensing* 41 (5 PART II): 1098–1104.

Sawhney, H. S., and R. Kumar. 1999. “True Multi-Image Alignment and Its Applications to Mosaicing and Lens Distortion Correction.” *IEEE Transactions on Pattern Analysis and Machine Intellingece* 21: 235–243.

Sedaghat, A., M. Mokhtarzade, and H. Ebadi. 2011. “Uniform Robust Scale-Invariant Feature Matching for Optical Remote Sensing Images.” *EEE Trans. Geosci. Remote Sens.* 49: 4516–4527.

Sester, Monika, Heiner Hild, and Dieter Fritsch. 1992. “DEFINITION OF GROUND-CONTROL FEATURES FOR IMAGE REGISTRATION USING GIS-DATA.” *International Archives of Photogrammetry and Remote Sensing* 32: 538-543.

Shaohong, Shen, Mo Xiaocong, and Zhang Qian. 2014. “Land Use/Cover Classification of Cloud-Contaminated Area by Multitemporal Remote Sensing Images.” *Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics* 1: 156-59.

Sharma, Ravi K, and Misha Pavel. 1997. “MULTISENSOR IMAGE REGISTRATION” *SOCIETY FOR INFORMATION DISPLAY* XXVIII (May): 951–54.

Shekarforoush, Hassan Foroosh, Josiane B Zerubia, Senior Member, and Marc Berthod. 2002. “Extension of Phase Correlation to Subpixel Registration.” *IEEE Transactions on Image Processing* 11: 188-200.

Shen, Huanfeng, Huifang Li, Yan Qian, Liangpei Zhang, and Qiangqiang Yuan. 2014. “An Effective Thin Cloud Removal Procedure for Visible Remote Sensing Images.” *ISPRS Journal of Photogrammetry and Remote Sensing* 96: 224–35.

Silverio, Walter, and Jean-michel Jaquet. 2005. “Glacial Cover Mapping (1987 – 1996) of the Cordillera Blanca (Peru) Using Satellite Imagery” *Remote Sensing of Environment* 95: 342–50.

Smith, S M, and J M Brady. 1995. “SUSAN - A New Approach to Low Level Image Processing,” *International Journal of Computer Vision* 23: 45–78.

Srinivasa, B., and B. N. Chatterji. 1996. “An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration” *IEEE TRANSACTIONS ON IMAGE PROCESSING* 5 (8): 1266–71.

Stockham, Thomas G. 1972. “Image Processing in the Context of a Visual Model.” *Proceedings of the IEEE* 60 (7): 828–42.

Stockman, G., S. Kopstein, and S. Benett. 1982. “Matching Images to Models for Registration and Object Detection via Clustering.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 4: 229–241.

Syst, J Electr Electron, and Vikram Mutneja. 2015. “Methods of Image Edge Detection: A Review” *Journal of Electrical & Electronic Systems* 4 (2).

Systeme, Arbeitsbereich Kognitive, Fachbereich Informatik, and Hamburg University. 1994. “Localization Properties of Direct Corner Detectors” *Journal of Mathematical Imaging and Vision* 4: 139-50.

Takahashi, Masuo, Kenlo Nishida Nasahara, Takeo Tadono, Tomohiro Watanabe, Masanori Dotsu, Toshiro Sugimura, and Nobuhiro Tomiyama. 2013. “JAXA HIGH RESOLUTION LAND-USE AND LAND-COVER MAP OF JAPAN.” *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2384–87.

Tao, CV, and Y Hu. 2000. “A Comprehensive Study of the Rational Function Model for Photogrammetric Processing.” *PE & RS- Photogrammetric Engineering & Remote Sensing* 67 (403): 1–26.

Thkvenaz, Philippe, and Michael Unser. 1997. “Spline Pyramids for Inter-Modal Image Registration Using Mutual Information” *Proceedings of the SPIE* 3169 : 236–47.

Thtvenaz, Philippe, and Michael Unser. 1998. “An Efficient Mutual Information Optimizer for Mutual Image Registration.” *Proceedings 1998 International Conference on Image Processing* 1: 833-37.

Tian, B., M. Shaikh, and M. Azimi-Sadjadi. 1999. “A Study of Cloud Classification with Neural Networks Using Spectral and Textural Features.” *IEEE Trans. Neural Netw.* 10: 138–151.

T.M.A. Santos, A. Mora, R.A. Ribeiro, J.M.N. Silva, Fuzzy-fusion approach for land cover classification, in: 2016 IEEE 20th Jubil. Int. Conf. Intell. Eng. Syst., IEEE, 2016: pp. 177–182.

Topan, Huseyin, and Hakan S. Kutoglu. 2009. "Georeferencing Accuracy Assessment of High-Resolution Satellite Images Using Figure Condition Method." IEEE Transactions on Geoscience and Remote Sensing 47 (4): 1256–61.

Tschumperlé, D., and R. Deriche. 2005. "Vector-Valued Image Regularization with PDE's A Common Framework for Different Applications." *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (4): 506–17.

Tseng, Din Chang, Hsiao Ting Tseng, and Chun Liang Chien. 2008. "Automatic Cloud Removal from Multi-Temporal SPOT Images." *Applied Mathematics and Computation* 205 (2): 584–600.

Unser, Michael. 1996. "A PYRAMID APPROACH TO SUB-PIXEL IMAGE FUSION BASED ON MUTUAL INFORMATION" *IEEE Int. Conf. on Image Processing* I: 265–68.

Ventura, Anna Della, Anna Rampini, and Raimondo Schettini. 1990. "Image Registration by Recognition of Corresponding Structures" *IEEE Transactions on Geoscience and Remote Sensing* 28 (3): 305-314.

Viola, Paul, and William M. Wells. 1997. "Alignment by Maximization of Mutual Information" *International Journal of Computer Vision* 24 (2): 1–29.

Visa, A., and J. Iivarinen. 1997. "Evolution and Evaluation of a Trainable Cloud Classifier." *IEEE Trans. Geosci. Remote Sens.* 35: 1307–1315.

Vujovic, Nenad, Student Member, and Dragana Brzakovic. 1997. "Establishing the Correspondence Between Control Points in Pairs of Mammographic Images" *IEEE TRANSACTIONS ON IMAGE PROCESSING* 6 (10): 1388–99.

Wang, Han, and Fei Shen. 2000. "Real Time Gray Level Corner Detector." *Conf Rob Vision* 6: 1-4.

F. Shen, H. Wang, Real time gray level corner detector, in: Proceedings of 6thInternational Conference on Control, Automation, Robotics, and Vision(ICARCV), 2000.

Wang, J.-F., H.-J. Hsu, and S.-C. Liao. 2007. "A Hybrid Algorithm with Artifact Detection Mechanism for Region Filling after Object Removal from a Digital Photograph." *IEEE Trans. Image Process.* 16 (6): 1611–1622.

Wang, Qunming, and Peter M. Atkinson. 2018. "Spatio-Temporal Fusion for Daily Sentinel-2 Images." *Remote Sensing of Environment* 204 (October 2017): 31–42.

Wang, Wen-hao, and Yung-chang Chen. 1997. "Image Registration by Control Points Pairing Using the Invariant Properties of Line Segments" *Pattern Recognition Letters* 18: 269–81.

Wiemker, Rafael, Karl Rohr, Lutz Binder, Rainer Sprengel, and H Siegfried Stiehl. 1971. "APPLICATION OF ELASTIC REGISTRATION TO IMAGERY FROM AIRBORNE SCANNERS" *Proceedings of the XVIII. Congress of the International Society for Photogrammetry and Remote Sensing* XXXI Part B4: 949-54.

Wolf, Paul, and Bon DeWitt. 2000. *Elements of Photogrammetry with Applications in GIS 3rd Edition*. Mc Graw.

Wu, C. 2007 "SiftGPU A GPU Implementation of Scale Invariant Feature Transform (SIFT)."

Ye, Yuanxin, and Jie Shan. 2014. "A Local Descriptor Based Registration Method for Multispectral Remote Sensing Images with Non-Linear Intensity Differences." *ISPRS Journal of Photogrammetry and Remote Sensing* 90: 83–95.

Yi, Z., C. Zhiguo, and X. Yang. 2008. "Multi-Spectral Remote Image Registration Based on SIFT." *Electron. Lett.* 44: 107–108.

Zeng, Chao, Huanfeng Shen, and Liangpei Zhang. 2013. "Recovering Missing Pixels for Landsat ETM+ SLC-off Imagery Using Multi-Temporal Regression Analysis and a Regularization Method." *Remote Sensing of Environment* 131: 182–94.

Zhang, Chengyue, Zhiwei Li, Qing Cheng, Xinghua Li, and Huanfeng Shen. 2017. "CLOUD REMOVAL BY FUSING MULTI-SOURCE AND MULTI-TEMPORAL IMAGES" *IEEE International Geoscience and Remote Sensing Symposium* 5: 6–9.

Zheng, Qinfen, and Rama Chellappa. 1993. "A Computation Vision Approach to Image Registration," *IEEE Transactions on Image Processing* 2 (3): 311-26.

Zhou, F., Z. Wang, and F. Qi. 2006. "Inpainting Thick Image Regions Using Isophote Propagation." *Proc. IEEE ICIP*, 689–692.

Zitov, Barbara, Jaroslav Kautsky, Gabriele Peters, and Jan Flusser. 1999. "Robust Detection of Significant Points in Multiframe Images" *Pattern Recognition Letters* 20: 199–206.

Zitova, Barbara, Jan Flusser, and F. Sroubek. 2002. "Application of Image Processing for the Conservation of the Medieval Mosaic." *Proceedings of the IEEE International Conference on Image Processing ICIP'02*, 993–996.