UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingenieros Informáticos

Spatio-temporal Analysis of Agricultural Landscape
Images: A Superpixel-based Approach

PhD Thesis

Angel M. García Pedrero
MSc Computer Science

2016

Departamento de Arquitectura y
Tecnología de Sistemas Informáticos
Escuela Técnica Superior de Ingenieros Informáticos
Universidad Politécnica de Madrid

INTERNATIONAL
CAMPUS OF
EXCELLENCE

PhD Thesis

Spatio-temporal Analysis of Agricultural Landscape Images: A Superpixel-based Approach

Angel M. García Pedrero
MSc Computer Science

Supervisors: Consuelo Gonzalo Martín, PhD
Mario Lillo Saavedra, PhD

2016

Angel M. García Pedrero, MSc Computer Science

Spatio-temporal Analysis of Agricultural Landscape Images: A Superpixel-based Approach

PhD Thesis, 2016

Supervisors: Consuelo Gonzalo Martín, PhD

Mario Lillo Saavedra, PhD

Universidad Politécnica de Madrid

Escuela Técnica Superior de Ingenieros Informáticos

Departamento de Arquitectura y

Tecnología de Sistemas Informáticos

Campus Montegancedo, Pozuelo de Alarcón

28223 Madrid, Spain

Thesis Committee

Chairman Ferrán Marques Acosta, PhD
Universitat Politècnica de Catalunya

Member Jan Dirk Wegner, PhD
Eidgenössische Technische Hochschule Zürich

Member María González de Audicana Amenábar, PhD
Nafarroako Unibertsitate Publikoa

Member Francisco Javier Marcello Ruíz, PhD
Universidad de Las Palmas de Gran Canaria

Secretary Ernestina Menasalvas Ruiz, PhD
Universidad Politécnica de Madrid

Substitute Dionisio Rodríguez Esparragón, PhD
Universidad de Las Palmas de Gran Canaria

Substitute Angel Rodríguez Martínez de Bartolomé, PhD
Universidad Politécnica de Madrid

Abstract

In the last hundred years, the world population has tripled and is still growing dramatically but resources have remained the same, causing changes in food supply outlook. According to the Food and Agriculture Organization (FAO), global food production will need to grow by 70% in order to satisfy the food and feed demand of a population of 9 billion people by 2050. In the current scenario, with limited arable lands and water scarcity, this means a greater pressure than ever before on productive land; thus today's main challenge is to put agriculture on a more sustainable and productive long-term path. In this context, the study of the variability within plots represents an opportunity for improving agricultural management. For most farmers, agricultural plots are variable which implies that not all areas within them require the same management; for example: some areas have specific needs of water, fertilizers, or pesticides in order to be more profitable. Therefore, adjusting the management of the plots based on the variability within them opens a route to a variable farm management able to improve cost/yield relationship.

New generation of optical remote sensors placed on aircraft, satellite platforms and drones, offers accessible and useful data of very-high resolution for monitoring and determining spatial variability of agricultural fields at plot level. However, the landscape complexity makes the manual analysis of the variability within an agricultural plot a highly time-consuming and expensive task. At the same time, it hinders considerably the monitoring of a high number of agricultural plots simultaneously. Moreover this very-high spatial resolution represents a challenge for traditional approaches of analysis based on pixels, unable to handle the within-class spectral variability, intrinsic to this type of images. Therefore, it exists a critical need to develop methodologies for efficiently and automatically extracting and analyzing information from very-high resolution images, that would allow stakeholders to enable variable farm management. In this context, it is of particular interest to automatically detect region dynamics within the agricultural plots.

The aim of this work is the development of a methodology for automatic generation of spatial and temporal information on the dynamics of agricultural land, particularly at plot level. In this sense, the study of agricultural scenes has been addressed through an object-based approach, exploring superpixel methods, which are seen as a link between the pixels of the image and objects of interest. To this end a superpixel method for multi-spectral images has been proposed. This method has been exploited in different applications at three different scales of interest:

(i) to analyze fragmented agricultural scenes, (ii) to delineate agricultural plots, and (iii) to capture the internal variability of agricultural plots.

The analysis of fragmented agricultural scenes has been approached by two methodologies. The first one focuses on providing a framework for multi-scale segmentation and, at the same time, a way to identify the best scale according to criteria of spectral variability of the regions at each scale. While this methodology provides spectrally similar regions at all scales, the variability of agricultural covers hinders to correctly establish the plot boundaries. This issue is further discussed in later approaches. The second methodology is aimed at generating thematic maps by combining only two scales which correspond to segments generated by a segmentation based on edges and the other based on superpixels. For mapping lands covers, it has combined the results of classifying the superpixels through a supervised classifier and the segments based on edges by a set of rules. The combination of both scales has yielded results with an accuracy better than those obtained on both scales separately.

Delineation of agricultural plots has been approached by two different methodologies. The first approach uses a supervised classification method to segment the image by agglomerating superpixels. This methodology represents an alternative to traditional methods of segmentation, which is based on learning how agricultural plots in a similar way as a human operator, as it does. In this regard, unlike conventional methods that require to find a suitable combination of segmentation parameters, the proposed method has the advantage that the classifier is able to find relationships in a multidimensional space to facilitate an adequate segmentation. The main drawback of this method is that it requires large amounts of annotated information for training the classifier. In this regard, a second method which uses only the image information for the delimitation has been proposed. From the basis that segmentations (superpixels) obtained by different parameters can capture various phenomena at different scales of an image, the proposed methodology allows to exploit its edges to obtain by consensus a segmentation of agricultural plots. It was found that superpixels can reduce the image noise, while most of the edges corresponding to the plots are kept in different segmentations. Both approaches are complementary, and depending on the availability of annotated data one or the other can be used, the first method can be used when there is ground truth, and the second one when there is no such information.

Finally, to capture the internal variability of agricultural parcels, it has proposed a method based on superpixels that considers spatial and temporal components of multiple images to find temporarily homogeneous regions within plots. Further analysis of the behavior of generated superpixels through the different dates provided the information about the spatio-temporal variability inside the plots. In this regard, three behaviors have been considered to determine the variability: (i) no variability, (ii) low-persistence variability, and (iii) high-persistence variability. This considerations allowed to create a map of spatio-temporal dynamics that provides an overview of the variability inside plots.

Resumen

En el último siglo, la población mundial se ha triplicado y continúa creciendo de forma alarmante; sin embargo, los recursos disponibles se mantienen igual, lo que provoca cambios en la perspectiva del suministro de alimentos. De acuerdo a la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), la producción mundial de alimentos tendrá que aumentar en un 70% con el fin de satisfacer la demanda de alimentos de una población de 9 mil millones de personas en 2050. En el escenario actual con limitadas tierras cultivables y la escasez de agua, lo anterior significa una presión cada vez mayor sobre las tierras productivas; por tanto, actualmente el principal desafío es conducir a largo plazo a la agricultura hacia un camino más sostenible y productivo. En este contexto, el estudio de la variabilidad dentro de las parcelas representa una oportunidad para mejorar la gestión agrícola. Para la mayoría de los agricultores, las parcelas agrícolas presentan variabilidad, lo cual implica que no todas las áreas dentro de las mismas requieren el mismo manejo; por ejemplo, algunas áreas tienen necesidades específicas de agua, fertilizantes o pesticidas que les permitirían ser más rentables. Por lo tanto, el ajuste de la gestión de las parcelas en función de su variabilidad interna abre una ruta hacia una gestión agrícola variable, es decir que varíe de acuerdo a esas necesidades permitiendo mejorar la relación coste/rendimiento.

La nueva generación de sensores ópticos de teledetección, instalados en aviones, plataformas de satélite y aviones no tripulados, ofrece datos de muy alta resolución accesibles y útiles para el seguimiento, así como la determinación de la variabilidad espacial de los campos agrícolas a nivel de parcela. Sin embargo, la complejidad del paisaje hace que el análisis manual de la variabilidad interna de las parcelas agrícolas sea una tarea costosa que requiere un tiempo considerable. Al mismo tiempo, el seguimiento simultáneo de un alto número de parcelas se dificulta considerablemente. Por otra parte, la muy alta resolución espacial de los datos representa un desafío para los enfoques tradicionales de análisis basados en píxeles, los cuales son incapaces de manejar la variabilidad espectral dentro de las clases. Por lo tanto, existe una necesidad crítica de desarrollar metodologías para extraer y analizar la información de las imágenes de muy alta resolución de manera eficiente y automática que permita a las partes interesadas la gestión agrícola variable de las parcelas. En este contexto, es de particular interés la detección automáticamente de las dinámicas dentro de las parcelas agrícolas.

El objetivo de este trabajo es el desarrollo de una metodología automática que permita la generación de información espacial y temporal sobre la dinámica de las tierras agrícolas, particularmente a nivel parcela. En este sentido, el estudio de las escenas agrícolas se ha abordado

mediante un enfoque basado en objetos, explorando técnicas de generación de superpíxeles, los cuales sirven como un enlace entre los píxeles de la imagen y los objetos de interés. Con este fin se ha propuesto una extensión de un método de superpíxeles para el procesamiento de imágenes multispectrales. Este método ha sido explotado en diferentes aplicaciones a tres diferentes escalas de interés: (i) para analizar escenas agrícolas fragmentadas, (ii) para delinear parcelas agrícolas, y (iii) para capturar la variabilidad interna de las parcelas agrícolas.

El análisis de escenas agrícolas fragmentadas se ha abordado mediante dos metodologías. La primera está centrada en proveer un marco para la segmentación multiescala y al mismo tiempo la identificación de la mejor escala de trabajo de acuerdo a criterios de variabilidad espectral de las regiones. Si bien esta metodología provee regiones espectralmente similares en todas las escalas, la variabilidad de las cubiertas agrícolas dificulta establecer correctamente los límites de las parcelas. Este difícil problema se ha abordado posteriormente en capítulos posteriores. La segunda metodología está orientada a la generación de mapas temáticos mediante la combinación de dos escalas, las cuales corresponden a segmentos generados mediante una segmentación basada en bordes y la otra basada en superpíxeles. Para obtener el mapa de las cubiertas se han combinado los resultados de clasificar los superpíxeles mediante un clasificador supervisado y los segmentos basados en bordes mediante un conjunto de reglas. La combinación de los resultados en ambas escalas ha permitido obtener resultados con una precisión superior a los obtenidos en ambas escalas por separado.

La delimitación de parcelas agrícolas ha sido abordada mediante dos diferentes metodologías. La primera metodología utiliza un método de clasificación supervisado para segmentar la imagen mediante la aglomeración de superpíxeles. Esta metodología plantea una alternativa a los métodos tradicionales de segmentación, la cual se basa en aprender como segmentar parcelas de una forma similar a cómo una persona lo hace. En este sentido, a diferencia de los métodos de segmentación convencionales que requieren encontrar una combinación de parámetros adecuada, el método propuesto tiene la ventaja de que el clasificador es capaz de encontrar relaciones en un espacio multidimensional permitiendo así una segmentación adecuada. El principal inconveniente de este método radica en que requiere grandes cantidades de información anotada con el fin de realizar el entrenamiento del clasificador. En este sentido, se ha propuesto una segunda metodología que utiliza solamente la información de la imagen para realizar la delimitación. Partiendo de la premisa de que segmentaciones (superpíxeles) obtenidas mediante diferentes parámetros pueden capturar diversos fenómenos de una imagen a diferentes escalas, la metodología propuesta permite explotar los bordes de estas segmentaciones para obtener mediante un consenso las parcelas agrícolas. Los superpíxeles permiten reducir el ruido de la imagen, mientras mantienen la mayoría de los bordes de las parcelas a diferentes escalas. Ambos enfoques son complementarios, así en función de la disponibilidad de datos anotados, el primer método puede ser utilizado cuando existe verdad de terreno, y el segundo cuando no existe esta información.

Finalmente, para capturar la variabilidad interna de las parcelas agrícolas, se ha propuesto un método basado en superpíxeles que considera las componentes espaciales y temporales de múltiples imágenes con la finalidad de encontrar regiones temporalmente homogéneas en el interior de las parcelas agrícolas. Así, un análisis detallado del comportamiento de los superpíxeles generados a través de las diferentes fechas mediante este método permite generar la información sobre la variabilidad espacio-temporal dentro de las parcelas. En este sentido, se consideraron tres comportamientos para determinar la variabilidad: (i) sin variabilidad, (ii) variabilidad de baja persistencia, y (iii) variabilidad de alta persistencia. Las consideraciones anteriores han permitido crear un mapa de la dinámica espacio-temporal que proporciona una visión general de la variabilidad dentro de las parcelas.

Acknowledgement

” *Caminante, son tus huellas el camino, y nada más;
caminante, no hay camino, se hace camino al andar. Al
andar se hace camino, y al volver la vista atrás se ve la
senda que nunca se ha de volver a pisar. Caminante, no
hay camino, sino estelas en la mar.*

— **Antonio Machado (1875–1939)**

Spanish poet.

Resulta siempre difícil poder agradecer a todas las personas que de alguna manera me han acompañado a lo largo de este camino. En primer lugar, quiero agradecer a mis directores de tesis Consuelo Gonzalo Martín y Mario Lillo Saavedra, por ser los primeros responsables de que haya conseguido concluir este trabajo. Gracias por haber confiado en mi y tenerme paciencia, por el trato cercano y buen carácter en todo momento, pero sobre todo en los momentos difíciles, así como los consejos que me han dado durante todo este tiempo. Por todo esto, y por los proyectos que puedan venir, muchas gracias.

Agradezco también a todos los miembros (antiguos y actuales) del grupo de MIDAS de la UPM, por haberme recibido, las siempre interesante pláticas durante la comida y por ayudarme en lo que he necesitado, por el trato familiar que siempre han tenido conmigo y sobre todo por su amistad. No puedo dejar de agradecer al grupo ITECMA² de la UdeC, por haberme hecho sentir como en casa durante mi breve estancia en Chile, la cual espero se repita en el futuro. Quisiera agradecer también a los miembros del proyecto ARTEMISAT (ULPGC y UPC), por las enriquecedoras experiencias generadas durante las reuniones del proyecto.

Quiero expresar mi profundo agradecimiento a mi madre, a mi hermano, y a toda mi familia, por haberme siempre apoyado y ayudado en lo posible. También quisiera agradecer a todos mis amigos de México, España y Chile por haberme aguantado todos estos años, y haberme dado la posibilidad de desconectar y reponer las fuerzas para continuar.

Igualmente quiero agradecer al Consejo Nacional de Ciencia y Tecnología (CONACyT) de México por el apoyo brindado a través de la beca de estudios doctorales (beca 216146). Así como a la Comisión Nacional de Investigación Científica y Tecnológica (CONICYT) de Chile por

el apoyo brindado para la divulgación en conferencias y revistas *open-access* de los resultados derivados de esta tesis.

Contents

I	Introduction	1
1	Introduction	3
1.1	Motivation and Problem Statement	7
1.2	Hypothesis and Objectives	9
1.3	Thesis Structure	10
II	State of the Art	11
2	Object-based Image Analysis	13
2.1	Introduction	13
2.2	Definition of Object-based Image Analysis Paradigm	15
2.3	Phases Involve in Object-based Image Analysis	17
2.4	Image-objects Generation	19
2.4.1	Image Preprocessing	20
2.4.2	Image Segmentation	21
2.5	Analysis Phase	26
2.5.1	Object-based Attributes	28
2.5.2	Supervised Methods for Image Analysis	29
2.6	Object-based Image Analysis in Agriculture	33
2.7	Segmentation not as End-itself	36
2.7.1	Supapixel Segmentation	38
2.8	Discussion	43
III	Supapixel-based Approaches for the Spatio-Temporal Analysis of Agricultural Landscapes	45
3	A Supapixel Method for Remote Sensing Applications	47
3.1	Introduction	47
3.2	SLIC for Remote Sensing Applications	48
3.3	Conclusions	50
4	Optimal Scale of Analysis in a Hierarchical Supapixel-based Segmentation Method	51

4.1	Introduction	51
4.2	Materials and Methods	53
4.2.1	Data Set	53
4.2.2	Methodology	53
4.3	Results and discussion	58
4.4	Conclusions	59
5	A Dual-Scale Segmentation Methodology for Analyzing Fragmented Agricultural Landscapes	63
5.1	Introduction	63
5.2	Materials and Method	63
5.2.1	Data Set	63
5.2.2	Methodology	65
5.2.3	Image segmentation scheme	66
5.3	Results and Discussion	70
5.4	Conclusions	76
6	Delineating Agricultural Plots in VHR Remote Sensing Images	77
6.1	Introduction	77
6.2	Materials and Methods	79
6.2.1	Data Set	79
6.2.2	Approach based on Classification	79
6.2.3	Approach based on Consensus of Segmentations	81
6.3	Results and Discussion	83
6.3.1	Approach based on Classification	83
6.3.2	Approach based on Consensus	85
6.3.3	Conclusions	88
7	A superpixel-based approach for temporal analysis of intra-plot variability	91
7.1	Introduction	91
7.2	Dataset and Methods	93
7.2.1	Data Set	93
7.2.2	Methodology	93
7.3	Results and Discussion	96
7.4	Conclusions	98
IV	Conclusions and future work	101
8	Conclusions and future work	103
8.1	Future Work	105
8.2	Selected Publications	106
A	Appendixes	109

A.1 Superpixels in Remote Sensing	109
Bibliography	113

Part I

Introduction

Introduction

“ *Farming looks mighty easy when your plow is a pencil,
and you're a thousand miles from the corn field.*

— **Dwight D. Eisenhower (1890–1969)**

34th president of US 1953–1961

In the last hundred years, the world population has tripled and is still growing dramatically but natural resources have remained almost the same, causing changes in food supply outlook. According to the [FAO \(2009\)](#), global food production will need to grow by 70% in order to satisfy the food and feed demand of a population of 9 billion people by 2050. Thus, the pressure on agricultural land due to the limited resources of arable land and water scarcity, is greater than ever (Seelan et al., [2003](#)). Under this scenario, the use of modern technologies in developed and developing countries can allow to (Seelan et al., [2003](#)): provide information to enable better management decisions, reduce costs associated with water and fertilizer application, improve crop yield and increase profit margin. Or in other words, helping to put agricultural lands on a more sustainable and productive long-term path.

In order to develop a sustainable agriculture, technological efforts must provide inexpensive and efficient solutions adapted to the following intrinsic characteristics of agriculture (Liang, [2008](#)):

- agricultural activities are usually performed on large areas, therefore, field surveys are time consuming and often expensive;
- most crops have annual patterns with diverse stages of growth and development in different seasons of the year, where intra-annual changes can be very drastic, hence, the recognition of the phenological status of crops is essential; and
- human activities and management strongly affect agriculture, therefore, it is required timely and accurate monitoring information.

In this regard, [Remote Sensing \(RS\)](#) satellite imaging technology offers a spatio-temporal monitoring capability of large areas, noted for its speed, accuracy, and economy. Moreover, optical sensors are able to capture diverse spectral properties of vegetation that make phenological monitoring possible (Jones and Vaughan, [2010](#)). These features make RS a useful tool for monitoring and management of agricultural lands.

RS technology has long been used in agriculture since its early stages. The launch in 1972, by the National Aeronautics and Space Administration (NASA), of the first civil Earth observation satellite (Landsat-1), devoted to study Earth's surface, represented a breakthrough for studying and monitoring worldwide croplands (Henebry and Beurs, 2013). Landsat-1 was firstly used by Bauer and Cipra (1973) to classify Midwestern US agricultural landscapes into corn and soybean fields. It provided information in the green, red, and two infrared bands, with a resolution of 80 m and a return frequency of 18 days (Mulla, 2013). Subsequent satellites, belonging to the Landsat program, have been improving these features. Today, Landsat-8, launched in 2013, collects imagery in the coastal, blue, green, red, near infrared, and two infrared bands at a resolution of 30 m, in addition, two thermal bands at a resolution of 100 m (Roy et al., 2014). Its character of free data for users has facilitated its adoption for realizing diverse agricultural studies (Miller et al., 2013).

Other US government initiatives to provide free RS data are the Advanced Very High Resolution Radiometer (AVHRR) and the Moderate-Resolution Imaging Spectroradiometer (MODIS) missions, started by the National Oceanic and Atmospheric Administration (NOAA) in 1978 and NASA in 1999, respectively. AVHRR provides information within a spectral range between 0.58 μm and 12.50 μm through five bands with a resolution of 1 km (Odeh and McBratney, 2000), whereas MODIS offers 36 bands, ranging in wavelength from 0.4 μm to 14.4 μm , at three resolutions 250, 500 and 1000 m (Maccherone and Frazier, 2010). In the European side, the Copernicus Program (Harris and Baumann, 2015), started in 2008, offers free RS data through the Sentinel satellite series. For example, Sentinel-2A Satellite (launched in 2015) currently provides data at three resolutions 10, 20 and 60 m, with 13 channels covering a spectral range between 0.44 μm and 2.19 μm (Gascon et al., 2014).

Several agricultural applications have been developed with the data coming from all these missions. These applications bounded primarily to the management and production of wide agricultural areas are categorized by Atzberger (2013) as follow: (i) biomass and yield estimation (Wiegand et al., 1979; Liu et al., 2010); (ii) vegetation vigor and drought stress monitoring (Xu et al., 2012; Jackson et al., 2004; Gao et al., 2011); (iii) assessment of crop phenological development (González-Sanpedro et al., 2008); (iv) crop acreage estimation and cropland mapping (Wardlow and Egbert, 2008; Galford et al., 2008), (v) mapping of disturbance and land use/land cover (LULC) changes (Verbesselt et al., 2010; Alcantara et al., 2012); and (vi) precision agriculture and irrigation management (Edlinger et al., 2012).

These applications provide a useful overview of what is going on agricultural lands at the regional and global scales; however it exists an inherent disadvantage associated with the capability of analysis which is determined by the smallest area represented by a pixel (spatial resolution). Thus, the low¹ spatial resolution (i.e., pixel sizes above typical field sizes) of the data captured by aforementioned sensors produces that spectral radiances registered in a pixel

¹In the case of Landsat, 30 m is considered as medium or even high resolution depending on the year of the publication.

contain mixed information from several land covers (Henebry and Beurs, 2013). This seriously complicates the identification, interpretation and validation of the data, and in consequence the reliability of the derived information products (Rembold et al., 2013). Moreover, this resolution does not make it possible to properly analyze the phenomena occurring at agricultural plot level, and therefore it hinders the application of precision agriculture techniques (Mulla, 2013; Yang et al., 2013). Site-specific agricultural applications require information with greater accuracy, and higher resolution to adequately survey crop conditions and thus contribute to improve agricultural productivity (Shi et al., 2016). This is particularly relevant for fragmented agricultural systems (i.e., farming systems based on smallholder producers with less than 10 ha land, Jayne et al., 2010), which represent the 80% of the world's estimated 500 million small farms and provide over 80% of the food consumed in the developing world (IFAD and UNEP, 2013). Figure 1.1 shows the farming system distribution over the world and described in Table 1.1.

Fig. 1.1.: Farming system maps of the developing regions of the world (modified from Bégue et al., 2015).

Tab. 1.1.: Farming system description (Bégue et al., 2015).

Farming system	Characteristics
Irrigated farming systems	Dominated by smallholder producers
Wetland rice based	Dominated by smallholder producers, dependent upon seasonal rains supplemented by irrigation
Rainfed farming systems in humid (and subhumid) areas	Dominated by smallholder producers, characterized by specific dominant crops or mixed crop-livestock systems
Rainfed farming systems in steep and highland areas	Dominated by smallholder producers, often mixed crop-livestock systems
Rainfed farming systems in dry or cold areas	Dominated by smallholder producers, with mixed crop-livestock and pastoral systems merging into systems with very low current productivity
Mixed large commercial and small holder	Dualistic, across a variety of ecologies and with diverse production patterns
Coastal artisanal fishing mixed	Dominated by smallholder producers, incorporates mixed farming elements
Urban based	Dominated by smallholder producers, typically focused on horticultural and livestock production

Fragmented agricultural landscapes, dominated by smallholder producers, are characterized by the diversity in sizes, shapes, and crops of the different agricultural plots. This high spatial heterogeneity makes agricultural studies using remote sensing are conditional on the spatial resolution of the images, thus higher resolutions allow more accurate analysis of the agricultural fields. Nowadays, technological progress makes it possible to gather information not only from aircraft and satellite platforms, but also from Unmanned Aerial Vehicles (UAVs), providing spatially and temporally distributed information on agricultural land at sub-metric resolutions. While the spatial resolution improves, the area represented by a pixel decreases and consequently the land-cover mixing within it (Mulla, 2013). This improved spatial resolution allows the use of images for assessing within-field yield variability and site-specific agricultural applications (Yang et al., 2013). Generally, the resolutions of RS images can be classified into (DigitalGlobe, 2015):

- low resolution: larger than 30 meters
- medium resolution: between 5 and 30 meters
- high resolution: between 1 and 5 meters
- very-high resolution: under 1 meter

As can be observed in Figure 1.2, finer resolution images provide a detailed overview of what happens in an agricultural plot, unlike coarser resolution images where hardly can be appreciated those details.

During the last decade, it has been observed an upward trend for the number of satellites that provide high resolution images. Table 1.2 shows the evolution of satellite image resolutions and their suitability for applications that study the variability within agricultural plots. Nowadays eight operational satellites provide Very-High Resolution (VHR) and High Resolution (HR) images suitable for analyzing the internal variability of agricultural plots.

Fig. 1.2.: An overview of agricultural plots using different image resolutions. Figures 1.2b and 1.2a have spatial resolutions of 30 m and 2 m, respectively.

Tab. 1.2.: Suitability of satellite images for studying the internal variability of agricultural plots. PAN refers to Panchromatic, C to coastal, B to blue, G to green, R to red, IR to infrared, NIR to near infrared, MIR to mid infrared, TIR to thermal infrared. Suitability for fragmented agricultural landscape analysis is represented by letters Y (Yes) and N (No). Resolution class L refers to low, M to medium and H to high and VH to very-high resolution.

Satellite (year)	Spectral bands (spatial resolution)	Resolution	Suitable
Landsat 1 (1972)	G, R, two IR (56 x 79 m)	L	No
AVHRR (1978)	R, NIR, two TIR (1090 m)	L	No
Landsat 5 TM (1984)	PAN, B, G, R, two NIR, MIR, TIR (30 m)	M	No
SPOT-1 (1986)	G, R, NIR (20 m)	M	No
IKONOS (1999)	PAN, B, G, R, NIR (1–4 m)	H	Yes
Terra EOS ASTER (2000)	B, G, R, two NIR, MIR, TIR (30 m)	M	No
EO-1 Hyperion (2000)	400–2500 nm, 10 nm bandwidth (30 m)	M	No
QuickBird (2001)	PAN, B, G, R, NIR (0.61–2.4 m)	VH	Yes
EOS MODIS (2002)	36 bands in VIS-IR (250–1000 m)	L	No
RapidEye (2008)	B, G, R, red edge, NIR (6.5 m)	H	Yes
GeoEye-1 (2008)	PAN, B, G, R, NIR1, NIR2 (1.6 m)	VH	Yes
WorldView-2 (2009)	PAN, C, B, G, Y, R, red edge, two NIR (0.5 m)	VH	Yes
Pléiades 1A (2011)	PAN, B, G, R, NIR (0.5m)	VH	Yes
Pléiades 1B (2012)	PAN, B, G, R, NIR (0.5m)	VH	Yes
Landsat 8 OLI (2013)	PAN, C, B, G, R, NIR, two MIR, two TIR (30–100 m)	M	No
WorldView-3 (2014)	PAN, C, B, G, Y, R, red edge, two NIR (0.25 m)	VH	Yes
Sentinel-2A (2015)	C, B, G, R, four red edge, NIR, IR, three MIR (10–60m)	M	No

1.1 Motivation and Problem Statement

This work is focused on the use of VHR satellite images for the analysis of fragmented agricultural landscapes at agricultural plot level. In this context, Govaerts et al. (2007) point out that “a non-uniform spatial distribution of crop performance within a field indicates the inefficient use of available resources, with consequent yield losses”. For most farmers, agricultural plots are variable which implies that not all areas within them require the same management. For example: some areas have specific needs of water, fertilizers, or pesticides in order to be more profitable; or because they are affected by a disease. Therefore, adjusting the management of the plots based on their internal variability opens a route to a variable rate management of tasks as irrigation, fertilization and pesticide application, that is able to improve cost/yield relationship. The identification of within-field spatial variability is not only useful for variable rate applications, but also for evaluating the economic returns of different farming management strategies, and realizing field-level improvements, such as drainage, land leveling, and irrigation (Yang et al., 2013).

VHR satellite images offer accessible and suitable data for monitoring and determining spatial variability of agricultural fields at plot level. However, the landscape complexity makes the manual analysis of the variability within an agricultural plot a highly time-consuming and expensive task. At the same time, it hinders considerably the monitoring of a high number of agricultural plots simultaneously. Moreover this very-high resolution represents a challenge for traditional approaches of analysis based on pixels, unable to handle the within-class spectral variability, intrinsic to this type of images (Blaschke et al., 2014; Hay et al., 1996; Kim et al.,

2008). Therefore, it exists a critical need of methodologies for efficiently and automatically extracting and analyzing information from VHR images of agricultural landscapes. These methodologies would allow stakeholders to enable variable farm management. In this context, it is of particular interest to automatically detect dynamics inside of regions of interest within the agricultural plot that have an abnormal behavior and hence require specific management.

Traditional image processing uses individual pixels as the basic unit in analysis of remotely sensed imagery. Pixels carry an integrated spectral signature and are characterized by their ground sampling size. Essentially, the set of pixels making up an image defines its entire amount of information. Nevertheless, working from individual pixels is a limiting strategy, since pixels are not treated as constituent elements of landscape units with a particular spatial context, but rather as independent samples (Schöpfer et al., 2010). Blaschke and Strobl (2001) highlight that pixels do not adequately represent the basic structural and functional units of the landscape, which are ontologically and epistemologically essential to understand natural phenomena. They argue that human perception focuses to discover and interpret patterns and mosaics in images rather than analyze individual pixels. In the same vein, Ren and Malik (2003) establish that pixels are not natural entities; they are simply a consequence of the discrete representation of images, and therefore they lack of semantic meaning in the real world.

Over the last two decades, Object-based Image Analysis (OBIA) has emerged as a new paradigm when processing and analyzing remotely sensed images of very-high resolution (Blaschke et al., 2014). Objects are the basis for OBIA as they represent meaningful entities in both the image and in reality and are closely aligned to how we perceive the world. Broadly speaking, the OBIA approach starts with a segmentation process, followed by successive analyses, usually at different hierarchical levels (scales) in order to create relationships between segments (Kim et al., 2008). Unlike pixel-based approach, OBIA provides additional information, by accessing segments features, such as mean spectral response per bands, statistical moments, contextual, spatial, textural, and morphological. These additional features can contribute to enhance the result of the analysis step (e.g., classification), particularly when two different image-objects are not separable using only the available information at pixel level.

Segmenting an image into objects opens up an opportunity to apply various geographical concepts to the processing of images, such as scale and context, that are not easily transferable to individual pixels (Blaschke and Strobl, 2001). The main idea behind generating segments is to reflect semantic objects of interest from the real world in a more appropriate way than purely spectrally defined agglomerations of individual pixels (Schöpfer et al., 2010). Starting from image objects rather than individual pixels, the analysis process can utilize spatial and geometrical properties as well as relationships among objects. And somehow try to imitate the processes carried out by a person during the interpretation of an image. In this regard, it results natural to think about agricultural plots as the main units of interest, and therefore it is crucial to have an adequate delineation of its boundaries in order to perform a correct

analysis of the variability within them. On this basis, OBIA offers the possibility of realizing better image analysis centered only on objects of interest, as is the case of the agricultural plots. However, delineating agricultural plots in VHR images by automatic or semiautomatic means is a challenge, particularly for fragmented agricultural landscapes, where fields are small and irregularly shaped (Forster et al., 2010). In this context, most of plots exhibit vague boundaries between neighboring fields, as well as a spectral ambiguity between plots and natural vegetation (Debats et al., 2016), hindering their segmentation. Moreover, they show a high spectral and phenological variability due to several factors such as sub-optimal management, low cropping intensity, rainfed farming, and the inclusion of large trees within fields (Debats et al., 2016).

1.2 Hypothesis and Objectives

The central hypothesis is that an approach based on image-objects would allow the automatic analysis of the spatial and temporal dynamics of the variability occurring at plot level in fragmented agricultural landscapes. The rationale for this study is that the generation of a reference framework where agricultural plots are the minimal units of interest, would facilitate processing very-high resolution data, and enable the analysis focused on the internal variability thereof.

The objective in this proposal is to develop an automatic methodology for generating information about spatial and temporal region dynamics of the agricultural fields at plot level (unit of interest). This research is predicated upon the use of object-based processing techniques which consider image-objects instead of pixels as basic processing units, providing a more efficient and “intelligent” approach, when compared with traditional pixel-based paradigm, for handling the intrinsic complexity of very-high resolution images.

Two specific objectives have been proposed for this project:

Objective #1: Develop an automatic object-based method for delineating agricultural plots from very-high resolution satellite images. The first step for the any automatic analysis is the identification of the object of interest, in this case the agricultural plot. In this regard, the correct plot delineation is essential to the analysis of the variability within parcels.

Objective #2: Develop an automatic object-based method for the joint spatial and temporal analysis of the variability at plot level. Once agricultural plots are identified, variability within them will be analyzed, in space and time in order to identify areas that could require specific management.

1.3 Thesis Structure

The structure of the document is as follows: Chapter 2 reviews the state of the art in Object-based Image Analysis, and its applications in remote sensing. In Chapter 3, a method for generating superpixels from multi-spectral remote sensing images is presented. Chapter 4 describes and analyzes a proposed approach for image analysis in the spatial domain, particularly a hierarchical multi-scale method based on superpixels for multi-spectral images, while in Chapter 5 a dual segmentation for analyzing fragmented agricultural landscapes is proposed. In Chapter 6, two approaches for delineating agricultural plots in VHR satellite images are presented and analyzed. Chapter 7 describes and analyzes the proposed approach for identifying the spatio-temporal variability within plots. In Chapter 8, the main conclusions of this study are developed, including a definition of the possible lines for the continuation of this thesis.

Part II

State of the Art

Object-based Image Analysis

“Objects don’t exist for me except insofar as a rapport exists between them or between them and myself.

— Georges Braque (1882–1963)

French painter and sculptor.

2.1 Introduction

Image pixels are the base of most of image processing tasks (Neubert and Protzel, 2012). Thus the most of algorithms for image analysis are based on classifications of individual pixels in a multi-dimensional feature space (Schöpfer et al., 2010). Although there have developed a number of sophisticated and well-established pixel-based techniques for image analysis, the processing on VHR images represents a restraint for these techniques due to the intrinsic characteristics of these images (Schöpfer et al., 2010). In VHR the pixel size is significantly smaller than the average size of the object of interest (Figure 2.1) which increases the within-class spectral variability and, consequently, decreases the potential accuracy of a pixel-based approach for image analysis (Hay et al., 1996; Kim et al., 2008; Blaschke et al., 2014).

Fig. 2.1.: Objects and sensor resolutions: the higher the resolution, the better the display of objects in the scene (modified from Blaschke et al., 2015).

Most of traditional RS image analysis methods process pixels as independent samples without considering their spatial context neither if they belongs to a particular semantic unit of the landscape (Schöpfer et al., 2010). In this context, pixels lack of semantic meaning in the real world; they are a merely consequence of the discrete representation of an image (Ren and Malik, 2003). Therefore they fail to represent the basic structural and functional units of

the landscape, which are essential to understand it (Blaschke and Strobl, 2001). Moreover it produces a poor performance during the analysis of VHR images, generating adverse results such as low overall classification accuracy and the salt-and-pepper effect (Blaschke et al., 2014), particularly if the mapping area has diverse units with high spatial heterogeneity (Zhang, 2015).

Shortcomings of pixel-based analysis of VHR images have stimulated new analysis concepts to be investigated in RS area. In this regard, the use of images segmentation algorithms to divide the VHR images into objects allows to apply geographical concepts to the processing of images that are not easily transferable to individual pixels, such as scale and context (Blaschke and Strobl, 2001). Image segments are intended to reflect semantic objects of interest from the real world in a more appropriate way than purely spectrally defined agglomerations of individual pixels (Schöpfer et al., 2010). Starting from image-objects rather than single pixels, the analysis process can access to spatial and geometrical properties as well as relationships among objects, not available in pixel-based processing. In addition, the results of object-based analysis are often easier to interpret than the obtained by pixel-based analysis. These factors have led to a paradigm shift from a pixel-based analysis to an object-based approach, referred to as Object-based Image Analysis in the scientific literature (Blaschke et al., 2008; Blaschke, 2010; Blaschke et al., 2014).

Object-based Image Analysis (OBIA) approach is not exclusive of RS area and is commonly used in areas such as Computer Vision and Biomedical Imaging. In fact, OBIA processing techniques have significantly improved as a result of research efforts during the last decades in those fields (Castilla and Hay, 2008). However, the intrinsic characteristics of RS data make it difficult transferring the results of such areas. The main obstacles are the lack of typical shape and crisp boundaries of the objects of interest, and the multi-band and multi-scale nature of RS images (Castilla and Hay, 2008; Schiewe, 2002). This has led to the creation of specific methodologies and software for RS area.

The advent of eCognition¹, the first commercial software for object-based analysis of RS data, in 2000 has meant a breakthrough in the adoption of object-based approach in RS community. Several works have been developed using this software. As reported by Blaschke (2010), about 50%-55% of 800 papers published until 2010 had been used this software to perform an object-based analysis of RS images. Nowadays several OBIA software alternatives, both commercial and free, are available for use (Meinel and Neubert, 2004). These factors and others described by Blaschke et al. (2014) point out the acceptance of OBIA as a new paradigm of RS image processing.

¹www.ecognition.com. Last access on May 2016.

2.2 Definition of Object-based Image Analysis Paradigm

According to Castilla and Hay (2008), Object-based Image Analysis (OBIA) aims to imitate human interpretation of RS images by computational means. Although results similar or better than human perception can be considered as the ultimate benchmark of OBIA, obtaining these results is far from trivial (Blaschke et al., 2014). Human perception is a complex issue that involves biological and cognitive processes to perform tasks such as filtering the relevant signals from noise, selective processing of detailed data, and interpreting this information (Lang, 2008; Blaschke et al., 2014). Moving from image perception to image interpretation, other factors such as the experience come into play. Image interpretation of RS images depends largely on the knowledge and experience of the domain that often goes beyond what is contained in an image. In fact, a successful visual interpretation is only possible through the iterative process of recognition and interpretation of objects (Figure 2.2), which depends heavily on the knowledge and experience of the analyst about the scene under analysis (Özdoğan, 2015). During the visual interpretation, the image content is not always identified and thus, the analyst has to deal with a semantic gap between the output of the analysis and the identified content available.

Fig. 2.2.: Schema of image interpretation process (Özdoğan, 2015).

In RS images, geographical objects are defined in terms of how they reflect solar radiation that result in their spectral signatures recorded in the image. Therefore, the visual elements such as color, shape, size, texture, and contextual information are the tools available that help recognition and interpretation of objects in a scene (Özdoğan, 2015). Moreover, when dealing a landscape, RS image interpretation often requires the recognition of various geographical objects or phenomena at various scales in order to understand what is happening in the whole scene (Lang and Tiede, 2015). Some of these aspects are that somehow the analysis tries to automate.

Formally, OBIA in a geographic context is defined as (Hay et al., 2014): “a sub-discipline of GeoInformatics² focused on (i) developing automated methods to **extract meaningful image-object** attributes from Remote Sensing imagery, (ii) assessing their **multi-scale** characteristics through time and space, (iii) and generating new geo-intelligence (i.e., geospatial content in ‘context’) from related multi-source datasets...”.

²According to Ehlers (2008): Geoinformatics is the discipline designed for handling of geospatial information. It encompasses the acquisition and storing of geospatial data, the modeling and presentation of spatial information, geo-scientific analyses and spatial planning, and the development of algorithms and geospatial database systems.

Here, terms as object and scale are indispensable to understand an image. Therefore, these concepts have been interpreted and adapted to object-based processing of RS images by the use of segmentation algorithms.

Object term is related to the output of some digital procedure, and by convention referred to as image-object. It is formally defined as “a discrete region of a digital image that is internally coherent and different from its surroundings” (Castilla and Hay, 2008). There are three characteristics that a region must possess in order to qualify as an image-object (Castilla and Hay, 2008): (i) discreteness; (ii) coherency; and (iii) contrast. An image-object is said to be discrete if it has defined boundaries to be manipulated by Geographic Information System (GIS) programs. Thus, each pixel of the image must have a unique identifier of the object within which is included. This makes it possible to represent it as an polygon in vector format, and facilitates its integration to GIS tools. On the other hand, coherency establishes that a region qualifies as an image-object, if the pixels within it are perceptually similar. This does not necessarily imply that the region has a low variance. For example, regions that show uniform texture patterns throughout its extent are coherent despite having a high variance. Finally, an image-object must have some contrast with the surroundings so that the region can be perceived as a distinct entity. Contrast may be produced by differences in color or texture between one object and its neighbors, or by the existence of edges separating the regions. Thus, image-objects attempt to represent meaningful geographical entities (scene components) that are distinguishable in an image at determined scales (Blaschke et al., 2014). In this context, image-objects are generally obtained through a segmentation process. It is important to highlight that generated image-objects (segments) are only perceptual entities with no meaning attached until a real-world concept (object class) is assigned to them (Castilla and Hay, 2008). This assignment is carried out usually by a classification process, in which an object-image is compared to a set of object classes through the analysis of their attributes, and finally classified as belonging to the more similar class.

Scale is an ambiguous and field-dependent term. In cartography, scale is considered as the ratio between the size of a map representation and the actual size in real world. In geography, it is related with the amount of detail of a geographic coverage; and in physics, biology and ecology, it is taken to refer to the size or extend of a process, or phenomenon under study (Atkinson and Tate, 2000). Later interpretation is closer to the idea of spatial scale, which provides a link between a property distributed in space and its representation. In RS images, visually distinct spatial patterns tend to emerge at specific scales, representing natural entities that form part of hierarchically organized systems, in this context, real-world objects exhibit an intrinsic scale and can be composed of structurally connected parts belonging to finer scales (Hay et al., 2003). As can be seen in Figure 2.3, a scene is composed of objects (phenomena, processes) of different size, shape and spatial arrangement, giving rise to multiple scales of analysis (multi-scale analysis). In OBIA, the scale refers to the size of the segments generated, whereas

multi-scale involves the generation of different size segments that may or may not be hierarchically related.

Fig. 2.3.: Scale representation in ecology: a forest patch nested within a landscape mosaic (Wu, 2016).

Other OBIA terms, commonly used in a geographical context, are defined in Table 2.1. In addition, Hay and Castilla (2008) provide a number of fundamental guiding principles to better understand OBIA frameworks. They describe the following core capabilities (Blaschke et al., 2014; Hay and Castilla, 2008): (i) its data sources is geocentric, i.e., it is originated from the surface of this planet (ii) its analytic methods are multi-source capable, (iii) object-based delineation is a prerequisite, and (iv) its methods are contextual, allowing for surrounding information and attributes.

2.3 Phases Involve in Object-based Image Analysis

OBIA relies on two main interlinked phases, often in a cyclic process (Hernando et al., 2012): (i) segment generation and (ii) an analysis phase. Segmentation process, in an initial step, serves as a bridge to spatial concepts applied in landscape analysis, and GISs, as well as the synergy between image-objects and their radiometric properties (Blaschke et al., 2008; Blaschke et al., 2014). This step is seen as a pre-attentive vision first stage in which object primitives (segments) are generated at one or multiple scales and then through a posterior analysis phase, they are turned into objects with a geographical meaning. Analysis phase commonly involves the study of a determined phenomenon in space/time (e.g. classification and change detection). It includes the construction of a model image interpretation that allows to establish a relationship between the analyst's knowledge about the scene and the features of the segments obtained in the previous phase. In this regard, it is clear that OBIA process depends on a correct delineation of the image-objects that are critical to recognizing geographical entities during the analysis phase.

Tab. 2.1.: Definitions of certain terms related to Object-based Image Analysis of Remote Sensing data (modified from Veljanovski et al., 2011).

Term	Definition
Geographical entity	Geographical entities are the constituent elements of landscape units, with a geographical meaning and their own characteristics in real world. They are real objects and phenomena on Earth's surface that represent the minimum unit of analysis, i.e., they can not be further divided into phenomena of the same type under analysis. Within object-based processing of satellite data, it is a synonym for geographical object, topographical object and spatial entity.
Object-based analysis, object-oriented analysis	Series of processing steps applied to a RS image for recognition, identification, assessment and analysis of spatial entities instead of individual pixels. These steps involve the segmentation and the spatio-temporal analysis (e.g. changes, comparison, classification, mapping) of homogeneous areas, structures and phenomena that can be observed in an image, as well as the evaluation of results obtained in these steps.
Segment	It represents the structural delineation of a geographical entity in accordance to the characteristics presented on the image. It is homogeneous in content, closed, discreetly limited to the area on the image. It defines the area for calculating object attributes.
Attributes, object attributes	They describe the characteristics of a segment and serve as an element to define or classify it regarding an object class. Attributes can have different nature such as spectral, geometric, textural, contextual, and temporal.
Object	It is an individual sample of one or more segments classified with regards to their attributes. Every object belongs to a thematic or object class and represents a geographical entity.
Object class, thematic class	It is a category that incorporates objects of the same type. It is the results of an image classification, therefore it often integrates elements of land use/land cover.
Object-based classification	It is the process of assigning a label to each segment in an image according the thematic classes that are relevant for a study. It involves the recognition of objects as well as relations between them.

Depending on how image-object generation (image segmentation) is performed, OBIA frameworks can be classified as: per-field and object-based approaches. The per-field approach, also known as per-parcel approach (Walter, 2004), typically involves the integration of vector data and raster images within a GIS (Berberoglu and Curran, 2004). In this approach, “field” or “parcel” refers to homogeneous patches of land, e.g., agricultural fields, urban structures or roads (Berberoglu and Curran, 2004) which are manually delineated and stored as a vector dataset belonging generally to a cadastral GIS. Thus vector data serves to extract image regions, i.e., pixels inside a polygon representing a spatial entity, which are later analyzed as a whole (Zhan, 2003). The value of this approach in improving analysis has been shown in several studies (Aplin and Atkinson, 2001; Berberoglu and Curran, 2004; Ruiz et al., 2009; Forster et al., 2010; Conrad et al., 2010; Ruiz et al., 2011). However, in many cases accurate and up-to-date vector datasets are rarely available (Zhan, 2003). Field boundaries may have change between the time of producing vector data and the time of acquisition of a new image (Zhan, 2003). In addition, datasets publicly disseminated are often generated on scales too coarse making difficult their use for analyzing phenomena occurring at finer scales (Duro, 2012). These drawbacks force to manually generate/update vector datasets, which is a subjective and, in many cases, highly time-consuming task. On the other hand, the automated segmentation procedure replaces manual per-field delineation in object-based approach (Forster et al., 2010). In this regard,

automated segmentation methods facilitate the objective generation of image-objects as well as reduce considerably the time spent on their delineation. In both cases, accurate field boundary delineation is a prerequisite for carrying out a successful analysis. However the aforementioned drawbacks of per-field approaches encourage the use of automated methods for image-object generation. This work focuses on the study of automated methods for object delineation.

Image-object generation and analysis phases are composed by several steps, which may vary depending on the characteristics of the image and the purpose of the analysis. A conventional object-based processing scheme is represented in Figure 2.4. The basic input data is represented by a multispectral RS image. However in accordance with the aim of the analysis, other ancillary data layers can be included in order to refine the process of sorting segments into object classes (Veljanovski et al., 2011). Furthermore, ancillary data can be also used in other subsequent parts of the analysis process, e.g., for characterizing image-objects.

After image-object generation process (discussed in Section 2.4), which involves preprocessing and segmentation steps, the next step is to assign segments to a certain class by comparing them with patterns previously defined, e.g., spectral signatures, textures, shapes. This later step is known as the analysis phase. Generally, analysis phase involves the identification of patterns from a huge set of attributes, which is difficult to perform by manual means. Thus it is necessary the use of Data Mining (DM) techniques that enable the automatic generation of structures of knowledge in order to interpret the scene under analysis (Vieira et al., 2012). Analysis phase is discussed in detail in Section 2.5.

Fig. 2.4.: Flowchart illustrating the main stages that are part of the OBIA approach.

2.4 Image-objects Generation

According to scientific literature, the generation of image-objects is a critical and initial step in any OBIA process. In this regard, Schöpfer et al. (2008) point out that the analysis phase depends on the quality of the generated image-objects, which in turn depends on the data under analysis and its preprocessing, and the segmentation algorithm. It therefore makes sense to think that the better the quality segments (closer to real-world objects) the better the quality of the later analysis steps.

2.4.1 Image Preprocessing

In an object-based approach, generating image-objects often requires a set of preliminary steps to segmentation, in order to improve the quality of the image and, consequently, that of the segmentation itself. These steps constitute a sub-phase called image preprocessing. The main purpose of the preprocessing step is to enhance image properties in order to facilitate the generation of segments, e.g., image-objects can be easily extracted by increasing the homogeneity and reducing noise in the image. Several methods, involving the modification or transformation of the original data properties, can be used depending on the desired result. Traditional RS image preprocessing techniques involve tasks such as image filtering, contrast adjustment, and pansharpening.

Image filtering is a set of techniques for enhancing an image through operations that allows emphasizing certain features or removing others. These operations include smoothing, and edge enhancement. Smoothing aims at reducing noise in the image, while conserving most representative image features. However, it usually blurs edges, which can severely reduce the performance of segmentation algorithms. Examples of smoothing filters are the Gaussian filter, box filter, and statistical filters such as average filters. In this context, filters that preserve most edges are preferred for OBIA applications. This type of filters aims to preserve the high-contrast edges and remove low-contrast or gradual changes, such as the median filter, and bilateral filter. Recently scale-aware filters, i.e., filters that can remove different levels of detail in images with the complete control of detail smoothing have been proposed in literature, an example of these methods is the Rolling Guidance Filter proposed by Zhang et al. (2014). Edge enhancement aims at enhancing the local discontinuities (gradients) at the boundaries of different objects (edges) in the image. Most of the edge enhancement filters are based on first and second order derivatives and different gradient filters are also common to use. Examples of edge enhancement filters include Roberts, Prewitt, Sobel, and Laplacian filters. While edge enhancement contributes to detect natural object boundaries useful during image-object generation, noise and non uniform lighting can produce adverse effects such as detecting false-positive edges and discontinuities in edges. Image filtering can be defined and applied in both frequency and spatial domains. A detailed review of implementations of these filters can be found in Gonzalez and Woods (2007) and Schowengerdt (2006).

Contrast adjustment expands the range gray values of the image to fill the available display gray level range by means of a mapping function. This approach attempts to improve the visual quality of a displayed image by highlighting some image structures. Thus, edges are better visualized when the contrast is higher. However, contrast adjustment is generally not desirable for the quantitative analysis of multispectral data because it modifies the spectral signature of the data (Schowengerdt, 2006). Contrast adjustment techniques are usually classified according its scope of operation into two groups: global and local transforms. While in global transforms, the mapping function is determined by the global (full image) statistical properties and is the same for every pixel, in local

transforms the mapping function changes from pixel-to-pixel according to the local image contrast inside a window centered on the pixel. Examples of contrast enhancement algorithms are the min-max, histogram equalization, and normalization methods. These methods are described in detail in Schowengerdt (2006).

Pansharpening allows to combine high-resolution panchromatic and lower resolution multispectral bands to generate a high-resolution multispectral image (Stathaki, 2011), that preserves both spatial and spectral quality of the source images to be useful in RS tasks. It exists several pansharpening algorithms in the literature, in this Section only the following two approaches are studied (Vivone et al., 2015; Aiazzi et al., 2012): component substitution and multiresolution analysis. Some examples of component substitution approaches are Intensity-Hue-Saturation, Principal Component Analysis, Grand-Schmidt, and Brovey Transform. Most of these techniques provide superior visual high-resolution multispectral images but ignore the requirement of high-quality synthesis of spectral information (Aiazzi et al., 2012). The multiresolution approach is based on the multiresolution decomposition of the panchromatic and multi-spectral images, with the aim of extracting spatial details, that are injected into the multispectral image. Most of these approaches use the Discrete Wavelet Transform, both in its pyramidal and redundant versions (Lillo-Saavedra et al., 2011). Even though these methods increase the spatial resolution of multi-spectral images to the spatial resolution levels of panchromatic images, while maintaining their spectral characteristics, they tend to inject false details into the pansharpened images (artifacts) (Gonzalo and Lillo-Saavedra, 2008). In this regard, it is necessary a compromise between spatial and spectral resolution of the image depending on the intended application (Xu et al., 2014). A detailed description of pansharpening methods can be found in Stathaki (2011) and Vivone et al. (2015).

2.4.2 Image Segmentation

Image segmentation is the process of partitioning a scene into meaningful segments (or image-objects). Each segment is characterized by a huge set of features that define its spectral properties, texture, shape, context, among others. The generated image-objects may or may not correspond to real-world objects (or in a geographical context, geographical entities), which are subsequently analyzed during analysis phase. There are several image segmentation algorithms in literature but most of them are developed for specific tasks and are not suited for RS imagery which is different from other images in several ways (Schiewe, 2002), e.g., several layers corresponding to different spectral channels, and redundancy and complexity of information.

Definition of Image Segmentation

Image segmentation is defined as grouping pixels into a set of n segments (Bhanu and Lee, 2012; Zhang, 2009; Zhang, 2006):

$$S = S_1, S_2, \dots, S_n \quad (2.1)$$

where S refers to the set of image segments S_i such that,

$$\bigcup_i^n S_i = I \quad (2.2)$$

$$\forall i, S_i \text{ is connected} \quad (2.3)$$

$$\forall i \neq j, S_i \cap S_j = \emptyset \quad (2.4)$$

The first condition (Equation 2.2) implies that every pixel of image I belongs to a segment. This means that a segmentation algorithm does not end until all pixels in an image are processed. The Equation 2.3 establishes that segments are composed of contiguous pixels (spatial proximity criterion). In other words, each spatially continuous unit (referred as connected component in machine vision literature) needs to be assigned a unique label (Blaschke et al., 2004). The third condition points out that different segments could not overlap each other. Note that pixels in the same segment should have similar properties (e.g., intensity, color, texture) or corresponds to some semantic interpretation, while pixels belonging to neighboring segments should have some different properties (Bhanu and Lee, 2012; Zhang, 2006). These conditions suggest important aspects that should be fulfilled for an image segmentation algorithm, however they do not lead to a unique algorithm, resulting in a large number of methods available.

Types of Segmentation Algorithms

Image segmentation methods have been developed within Pattern Recognition and Computer Vision since 1980, with successful applications in disciplines like medicine or telecommunication engineering. Traditional image segmentation algorithms can be broadly grouped into the following categories (Marpu, 2009; Blaschke et al., 2004; Schiewe, 2002):

Thresholding algorithms are the simplest and most commonly methods of image segmentation. Thresholding methods creates binary images from grey-level ones by turning all pixels below some threshold to zero and all pixels about that threshold to one. It exits many methods to select a suitable threshold value for a segmentation task. Perhaps the most commonly automatic technique for determining an adequate threshold is the proposed by Otsu (1975). Several variants can be found in the literature (Sezgin et al., 2004), which propose different ways to exploit the information to select a threshold. The major problem with thresholding is that we consider only the intensity, not any relationships between the pixels. Therefore these methods do not necessarily produce a result which fulfills the requirement and definition of segmentation, and therefore

the resulting output needs to be post-processed in order to obtain connected segments (Blaschke et al., 2004).

Clustering algorithms first identify a predefined number of clusters in a multidimensional feature space where each dimension corresponds to a band of the image. Clusters are generated by methods such as K-means, Kohonen self organizing map, Expectation Maximization algorithm. Canty (2011) and Schowengerdt (2006) offer a description of the most commonly used clustering algorithms for segmentation in literature. This type of methods is computationally expensive in terms of time, since they need to deal with all pixels during processing due to their lack of spatial proximity criterion. As with thresholding methods, clustering algorithms require a post-processing step to generate connected segments.

Edge-based algorithms exploit image gradients to find edges between the regions and determine the segments as pixels within these edges. Edges are considered as boundaries between image objects and they are located where changes in the gradient occur (Blaschke et al., 2004). There are various ways to delineate boundaries, but in general the first step of any edge-based segmentation methods is edge detection, which consists of three steps (Blaschke et al., 2004; Gonzalez and Woods, 2007): filtering, enhancement and detection. Filtering stage is often necessary to reduce the noise in images. The enhancement techniques aim at revealing local changes in intensities. One option to implement the enhancement step is to use high-pass filtering. Finally, the actual edges are detected from the enhanced data using, for example, a threshold technique. The detected edge points need to be linked to form region boundaries and the regions have to be labeled. Nevertheless, problems arise in edge detection due to noise effects, and discontinuities in the edges because of nonuniform lighting and other effects such as occlusions and blurring. Moreover, false edges tend to be detected in images with many textures.

Region growing methods are generally based on the growth of random seeds (initial regions) distributed throughout the image using predefined criteria. These methods, also known as bottom-up segmentation algorithms, joint neighboring pixels to the initial regions until a certain threshold is reached. This threshold is normally set up according to a homogeneity criterion combined with a desired size of segments. A region grows until no more pixels can be assigned to any of the segments and new seeds are placed, then the process is repeated. This continues until the whole image is segmented. These methods are highly dependent on the input seeds, sensitive to noise, and sometimes suffer from lacking control over the cut-off criterion for the growth of a region (Blaschke et al., 2004). Region-growing algorithms result in hierarchical segmentations where the images can be represented at different levels. Some examples of region-growing methods are given by Shih and Cheng (2005), Adams and Bischof (1994), and Chang and Li (1994).

Region splitting algorithms are classified as the top-down segmentation algorithms. Unlike region-growing algorithms which start from a set of seed points, in region splitting

methods the input usually consists of large segments which are divided into smaller units if they are not sufficiently homogeneous (Blaschke et al., 2004). Commonly, they start with the entire image and iteratively split regions into regions of pre-defined shape (squares or rectangles) (Marpu, 2009). The final segmentation is achieved when all segments meet the criteria of homogeneity. Their main drawback is that they produce segmentation errors at the edges of the split regions. An example of these type of methods is the Quad-tree segmentation algorithm proposed by Spann and Wilson (1985).

Graph-based algorithms use graph theory for partitioning an image. The image is represented as a weighted graph where each pixel is a node and an edge is formed between pairs of pixels with a weight defined according to a similarity measure. Segments are obtained through the generation of sub-graphs by removing the edges that connect the pixels that are not similar. Several algorithms based on this approach have been proposed in literature, (e.g., Felzenszwalb and Huttenlocher, 2004; Boykov and Jolly, 2001; Shi and Malik, 2000). One major drawback of these methods is related to the memory usage which increases quickly as the image size increases. This limits their use in applications where processing large images is required.

The last three categories can produce a hierarchical representation of the scene with segments at different scales (Figure 2.5) if the order in which regions are merged is stored. Depending on the direction in which the hierarchical segmentation is performed, bottom-up and top-down approaches can be identified. Bottom-up approaches start by grouping pixels and can end with the complete image as a single segment, whereas top-down approaches follow the opposite way, starting with the entire image and dividing until the pixel level can be reached.

Fig. 2.5.: Scheme of a hierarchical segmentation: top-down and bottom-up approaches.

More extensive reviews on image segmentation techniques can be consulted in the following papers: Cheng et al. (2001), Pal and Pal (1993), Fu and Mui (1981), and Haralick and Shapiro (1985).

Although it exists a large number of image segmentation algorithms, most of them can not be transferred directly to the domain of RS mainly because of the following reasons (Schiewe, 2002):

- RS sensors provide data on multiple wavelength bands and also at different resolutions, which increases not only the complexity but also the redundancy of information.
- A large variety of additional data (e.g., information from GIS systems) is available and can be integrated into the process.
- The objects in RS data can have heterogeneous properties with respect to size, form, spectral behavior.
- It is very difficult to employ model-based interpretation due to the inherent heterogeneity of the object classes.

The above reasons have led to modify or create new algorithms that can deal with RS data.

Multi-resolution Segmentation

The multi-resolution segmentation (MRS) algorithm, available in the eCognition software, is one of the most widely used in the RS community for generating image-objects during glsobia studies. MRS method, proposed by Baatz and Schäpe (2000), is based on the assumption that individual image-objects are similar in color and maintain a spatial proximity (Nussbaum and Menz, 2008). This method follows a bottom-up approach to generate image-objects. It joints pixels/regions according to a criteria of local heterogeneity which is evaluated by an optimization procedure that minimizes the average heterogeneity³ of pixels inside image-objects while encourages a determined size (scale parameter) of segments (Baatz and Schäpe, 2000). The criteria for calculating the heterogeneity are defined as follows (Marpu, 2009; Nussbaum and Menz, 2008).

Spectral heterogeneity. It is defined as a weighted sum of the standard deviations of the spectral values of the object pixels over all the spectral channels:

$$h_c = \sum_{b=1}^{NB} w^b \sigma^b \quad (2.5)$$

where σ^b is the standard deviation of the object in the layer b , w^b are the weight assigned to the layer, and NB represents the total number of spectral channels.

Spatial heterogeneity. A shape criterion for each segment is defined as the relative deviation of the edge length from a circle in two ways:

- Area criterion (compactness) represents the deviation from the ideal shape, given by the ratio between the border length of the object (l) and the segment size (n) in pixels.

$$h_{area} = \frac{1}{\sqrt{n}} \quad (2.6)$$

³Heterogeneity and homogeneity are opposite terms, both are used to describe a segment depending on the situation.

- Edge criterion (smoothness) is the ratio between the border length of the segment (l) and the edge length of a bounding box ($bbox$).

$$h_{edge} = \frac{l}{bbox} \quad (2.7)$$

Starting from pixels, image-objects are iteratively merged until a termination condition is met. At each step, a fusion value is calculated in order to decide when to combine two segments. This fusion value expresses the increase of heterogeneity resulting from merging two segments. Before a merging operation, the fusion value is calculated for each neighbor of the selected segment. The neighbor with the smallest fusion value is chosen for merging, but only if it is lower than a certain threshold, defined as the square of the scale parameter. This scale parameter determines the desired size of the image-objects. During merging process, the information of merged segments is stored to create a hierarchical network of objects. The fusion value (f) for merging two segments ($obj1$ and $obj2$) considers the heterogeneity of both segments as well as the segment resulting of their union ($merge$), and is calculated as follows:

$$f = w_{spectral}h_{spectral} + (1 - w_{spectral})h_{shape} \quad (2.8)$$

where spectral and spatial heterogeneities are defined in Equations 2.9 and 2.10, respectively.

$$h_{spectral} = h_{merge_c} - (h_{obj1_c} - h_{obj2_c}) \quad (2.9)$$

$$h_{shape} = w_{spatial}n_{merge}h_{merge_area} - (n_{obj1}h_{obj1_area} + n_{obj2}h_{obj2_area}) + (1 - w_{spatial})n_{merge}h_{merge_edge} - (n_{obj1}h_{obj1_edge} + n_{obj2}h_{obj2_edge}) \quad (2.10)$$

where $w_{spectral}$, $w_{spatial}$ are the weights assigned to the spectral and spatial heterogeneities, respectively. Both weights fall in the interval $[0, 1]$. The MRS algorithm has three main user-set parameters: scale, shape, and compactness.

2.5 Analysis Phase

Depending how the interpretation of the scene is performed, two procedures can be distinguished during analysis phase: manual and automatic interpretation. The manual analysis requires the analyst skills to identify and delineate geographical objects in an image. Although this procedure tends to generate precise results, it is expensive and highly time-consuming, involving thousands of analyst hours. On the other hand, automatic approach uses Machine Learning/Pattern Recognition techniques to identify geographical entities. Unlike manual approach, in which the knowledge of the whole landscape is needed, automatic approach only requires information on a subsample (set of observations) of the landscape for analyzing an image (Özdogan, 2015). These approaches can be used in both the pixel-based and object-based analysis. Unlike pixel-based analysis, in which each pixel represents an observation, in object-based analysis, image-objects are considered thematically homogeneous, which implies that the set of pixels within an image-object is considered as an individual training pattern

by these methods. Unsupervised and supervised methods are commonly used in automatic approach.

Unsupervised methods establish, with the help of an algorithm, a relationship between the observed patterns, in the feature space, and a predetermined number of statistical clusters representing the information about object classes. These methods are often applied in situations where prior knowledge about a scene is not available to provide information about statistical homogeneous clusters in the image. After clusters are created, the analyst uses a sample of known class to generate associations between the information clusters and actual object class labels. Examples of clustering methods are k-means (Arthur and Vassilvitskii, 2007), DBSCAN (density-based spatial clustering of applications with noise; Ester et al., 1996), and mean-shift (Comaniciu and Meer, 2002) algorithms. On the other hand, in the case of supervised methods, a set of labeled samples is collected to train a classifier in order to identify the class label of every unlabeled image-object in the image. Only a fraction of the population is needed to estimate the characteristics of every individual in the population (image-objects), i.e., knowledge on less than 5% or 15% of the landscape is used to extract information on the entire landscape under analysis (Schowengerdt, 2006). Examples of supervised methods commonly used in the scientific area are SVM (support vector machines; Vapnik, 1998), k-NN (*k*-nearest neighbors; Altman, 1992), Decision Tree (Loh, 2011), and Random Forest (Breiman, 2001). Note that the accuracy of the unsupervised approaches is generally lower than supervised ones, especially in complex landscapes where the characteristics of the different thematic categories may be similar. In this regard, supervised methods are more commonly used in RS applications.

Figure 2.6 shows a general scheme of the analysis using supervised methods (also known as classifiers or predictors). Three steps can be identified during this process (Schowengerdt, 2006): (i) feature generation, (ii) training, and (iii) labeling.

Feature extraction represents a fundamental step of any application that uses traditional supervised methods. It basically consists of extracting features that provide a good representation of the information contained in the image. In this context, feature is synonymous of input variable or attribute (Guyon and Elisseeff, 2006). To maintain consistency with the OBIA literature, the term attribute instead of feature is used in the rest of this document. Each data observation consists of a vector $\mathbf{x}$ of dimension n , $\mathbf{x} = [x_1, x_2, \dots, x_n]$, where x_i represents an attribute which can be binary, categorical or continuous. In the case of RS images, primary attributes are the spectral indexes (e.g., vegetation and water indexes), which are obtained from a linear combination of the spectral values (stored as pixels) of the image. It does not exist a clear definition of what is a “good” attribute but there are some desirable properties that must have (Aichert, 2008): perceptually meaning, consistent over several images of the same scene, and insensitive to noise. Most common attributes in the context of OBIA are addressed in more detail in Section 2.5.1.

Training consists of the selection of samples to train a model able to recognize the desired thematic classes from labeled data. During training process, the model attempts to learn the relationship between a set of attribute vectors and a target variable of interest by means of a learning algorithm (Aggarwal, 2014). As result, it is produced an inferred function which can be used for mapping new observations, allowing to determine the target value for unseen instances. Formally, supervised methods seeks a function $h : \mathcal{X} \rightarrow \mathcal{Y}$ that allows to approximate a set of observations ($\mathcal{X}$) to a set of labels ($\mathcal{Y}$), from a given set of n training set $\mathcal{D} = \{(\mathbf{x}^{(1)}, y^{(1)}), (\mathbf{x}^{(2)}, y^{(2)}) \dots (\mathbf{x}^{(n)}, y^{(n)})\}$, where $\mathbf{x}^{(i)} \in \mathcal{X}$ represents an attribute vector and $y^{(i)} \in \mathcal{Y}$ its corresponding target value. When y is a discrete value, the supervised problem is known as classification, whereas if y is a real value, it is called regression. These methods ideally require a large amount of labeled samples (training set) with good statistical properties to generate models with a good discrimination among classes. There are several supervised methods in literature such as decision trees, probabilistic methods, SVM methods, instance-based methods, and neural networks. A good review of these can be found in Aggarwal (2014). In Section 2.5.2, the most common supervised methods for object-based analysis are addressed.

Labeling consists of the application of the trained model to the entire image to label all unlabeled image-objects, and consequently the pixels within them. It is also known as testing or prediction phase. The output map consists of one label for each pixel.

Fig. 2.6.: General scheme of the analysis using supervised methods (Tang et al., 2014).

2.5.1 Object-based Attributes

One of the strengths of OBIA, compared with traditional pixel-based approaches, is that provides an improved basis of attributes for image analysis (Marpu, 2009). In this regard, image-object exhibits useful attributes that pixel lacks (Hay and Castilla, 2008). Thus a lot of information in addition to the spectral characteristics of the region can be extracted from image-objects. Several types of attributes can be used to describe image-objects. Only a brief review of the object attributes is presented here (for more detailed description, the reader is directed to consult the following works: Definiens (2009), and Gonzalez and Woods (2007).

Broadly speaking, OBIA attributes can be grouped according to its nature into:

Layer-value attributes. Statistical moments such as the mean, minimum, maximum, standard deviation, variance, median, skewness, and kurtosis, among others can be calculated from the pixels that form each image-object in each data source layers. Thus, hundred of potentially useful variables can be created.

Shape attributes. Geometric properties like object area, perimeter, compactness, centroid, axis lengths, elongation, eccentricity, orientation among others offer additional metrics that can help differentiate certain object classes. Additional shape descriptors for regions are detailed in the book of Gonzalez and Woods (2007).

Texture attributes. Texture describes the visual “roughness” of a surface and it is manifested in an image as local, quasi-periodic variations in intensities caused by the spatial distribution of scene reflectance (Schowengerdt, 2006). Examples of texture descriptors are the Haralick’s texture features based on the Gray-Level Co-occurrence Matrix (GLCM; Haralick, 1979), Fractal Dimension (FD; Bisoi and Mishra, 2001), and Local Binary Pattern (LBP; Heikkilä et al., 2009). Besides it can be generated layers containing the texture information of a window around each image pixel by using spatial filters (Gonzalez et al., 2009) such as local entropy, local range, and local standard deviation.

Context attributes. They describe an image-object in terms of its relations with the other images-objects, i.e., an image-object corresponding to a shadow has a bigger difference to the brighter neighboring segments (Marpu, 2009). Various types of context attributes can be generated depending on the type of the established relationships. Thus, relationships can be established regarding layer-values, shapes, and texture attribute at one scale or multiples scales (Definiens, 2009). In addition, other way of defining context is the position of the image-object with respect to the entire scene or any particular class objects.

Although it is claimed in literature that object-based attributes can improve image analysis (Marpu, 2009), this huge number of attributes poses a problem: how to distinguish what are the best them to analyze a scene. The use of irrelevant attributes can often result in poor modeling, since they are not well related to the target label. Typically such features worsen the analysis accuracy. Therefore, it is critical to use the correct attributes during the training process. In this regard, the analyst’s experience plays an important role in selecting good attributes.

2.5.2 Supervised Methods for Image Analysis

Because attributes derived image-objects not necessarily obey parametric statistical distributions that allow an easy separation of the analyzed object classes, it is necessary to use methods based on Machine Learning theory to perform the analysis. As mentioned above, several supervised methods are available, which have different properties better suited to certain problems and

thus may show a different analysis potential. In fact, it is not possible to declare one classifier better than another for all applications; their performance is strongly dependent on the data especially if there is some overlap of class attributes (Schowengerdt, 2006).

Regarding object-based analysis, k -nearest neighbors (k -NN) method, implemented in eCognition, is the most commonly used due to its simplicity and flexibility (Li et al., 2016). However, it may be less effective than other more complex methods in the presence of high-dimensional data (Maxwell et al., 2015), as found Mallinis et al. (2008) when a Decision Tree (DT) outperformed k -NN classification. Other supervised methods already used in OBIA are Random Forest (RF; Stumpf and Kerle, 2011), and Support Vector Machines (SVM; Heumann, 2011).

Decision Tree

DT creates a hierarchical partitioning of the feature space, which relates the different partitions at the leaf level to the different classes. A partition is created at each level by using of a split criterion on a single attribute. The general approach is to try to recursively divide the training data to maximize the separation between the different classes over different nodes. The separation between the different classes is maximized when the level of skewness between different classes in a given node is maximized. To quantify this skew in a node N a measure such as Gini-index or entropy can be used. Gini-index is defined as follows (Aggarwal, 2014):

$$G(N) = 1 - \sum_{i=1}^k p_i^2 \quad (2.11)$$

where p_i represents a fraction of the samples belonging to the k different classes. The value of the Gini-index lies between 0 and $1 - 1/k$, where the smaller the value, the greater the skew. On the other hand, entropy is calculated as follows (Aggarwal, 2014):

$$E(N) = - \sum_{i=1}^k p_i \log(p_i) \quad (2.12)$$

The value of the entropy lies between 0 and $\log(k)$. The smaller the entropy, the greater the skew in the data.

During the construction of the training model, a split it made in order to minimize the weighted sum of the Gini-index or entropy of the two resulting nodes. This step is performed recursively until a stopping criterion is satisfied, which can be that all samples in the node belong to the same class. The termination criterion requires either a minimum level of skew or purity to avoid overfitting (Aggarwal, 2014). However, during decision tree construction is not possible to predict the best time to stop its growth in order to prevent overfitting. Therefore, in most cases, decision trees are pruned to remove nodes that may correspond to overfitting. A minimum description length principle is commonly used to decide when to prune a node from the tree (Aggarwal, 2014).

Random Forest

Random forest (RF) is an ensemble classification algorithm which uses decision trees as base classifiers (Breiman, 2001). Each tree is generated using a dataset $\mathcal{D}_t$ sampled independently with replacement from the training set $\mathcal{D}$. Each $\mathcal{D}_t$ usually contains $2/3$ of dataset $\mathcal{D}$. Trees are grown to the maximum depth on new training data using a combination of features. These fully grown trees are not pruned. Each of them casts a vote for the most popular class to classify a new sample. Formally, a RF is defined as an ensemble H method consisting of n trees classifiers h_t (Aggarwal, 2014):

$$H(\mathbf{x}) = \arg \max_{y \in \mathcal{Y}} \sum_{t=1}^n (h_t(\mathbf{x}) = y) \quad (2.13)$$

During RF construction, some data may be used more than once in the training of trees, while others might never be used. This causes the classifier more robust to slight variations in the input data and, at the same time, increases its accuracy (Rodriguez-Galiano et al., 2012). Samples that are not present in $\mathcal{D}_t$ are included as part of another subset called out-of-the-bag (OOB). For each subset $\mathcal{D}_t$ used for training the t -th tree, it exists an OOB_t . Thus, elements in OOB_t can be used to evaluate the performance of t -th tree. The ratio of misclassifications and the total number OOB elements provides an unbiased estimator of the generalization error of the model. The generalization error converges to a limit as the number of trees (n) in the forest becomes larger (Breiman, 2001).

Moreover, when a tree is constructed, a random subset of input attributes instead of using whole input attributes to split nodes. This can decrease the strength of each tree, but also reduces the correlation between the trees, which consequently decreases the generalization error (Rodriguez-Galiano et al., 2012).

Support Vector Machine

Support Vector Machine (SVM) method is based on statistical learning theory and aims to determine the location of decision boundaries that produce optimal separation of classes (Vapnik, 1998). In a two-class problem where classes are linearly separable, SVM find the hyperplane that leaves the greatest margin between the two classes ($\mathcal{Y} = \{-1, +1\}$). Thus, the hyperplane can be described as follow (Aggarwal, 2014):

$$\mathbf{w} \cdot \mathbf{x} + b = 0 \quad (2.14)$$

where $\mathbf{w}$ is a d -dimensional vector representing the coefficients of the hyperplane of separation, and b is a constant. Both parameters need to be learned from the training data $\mathcal{D}$. The margin is defined as the sum of the distances to the hyperplane from the closest points of the two

classes. In order to maximize the margin, the following minimization problem can be solved using standard Quadratic Programming optimization techniques (Aggarwal, 2014):

$$\min \frac{1}{2} \|\mathbf{w}\|^2 \quad \text{such that} \quad y^{(i)}(\mathbf{w} \cdot \mathbf{x}^{(i)} + b) - 1 \geq 0 \quad \forall i \quad (2.15)$$

The data points ($\mathbf{x}^{(i)}$) that are closest to the hyperplane are called “support vectors” and are used to measure the margin.

If the two classes are not linearly separable, SVM tries to find the hyperplane that maximizes the margin, while at the same time, minimizing a quantity proportional to the number of misclassification errors. A positive user-defined parameter C controls the trade-off between margin and misclassification error (Pal, 2005). SVM can also be extended to handle nonlinear decision surfaces by use of kernel functions, that allow to project the input data into a high-dimensional feature space. The most common kernel functions are the Gaussian radial basis function and polynomial kernels (Aggarwal, 2014). SVM was initially designed for two-class problems; however when dealing with multiple classes, techniques such as “one against one” and the “one against the rest” are used (Pal, 2005).

The main drawback of the aforementioned classification methods is the difficulty to deal with unbalanced training datasets which are very common in real-world problems such as RS applications. In this regard, RUSBoost is a hybrid boosting/sampling method proposed by Seiffert et al. (2010), which is a state-of-the-art method for learning from imbalanced datasets (Galar et al., 2012).

RUSBoost classifier

RUSBoost improves boosting algorithm by resampling training data in order to balance the class distribution. RUSBoost is based on SMOTEBoost (Chawla et al., 2003), which in turn is based on the AdaBoost.M2 algorithm (Freund and Schapire, 1996). Unlike other ensemble methods RUSBoost applies RUS as under-sampling strategy to randomly remove samples from the majority class, before the training of each weak learner algorithm that is part of the ensemble. This reduces the total time of training without performance loss (Seiffert et al., 2010).

RUSBoost combines an under-sampling technique, known as RUS, with AdaBoost.M2 algorithm. AdaBoost.M2 is an ensemble method for multi-class classification. Similar to the Random Forest method, described in Section 2.5.2, RUSBoost combines many weak classifiers h_t into a strong classifier H by linear combination. The final classifier is constructed as:

$$H(\mathbf{x}) = \arg \max_{y \in \mathcal{Y}} \sum_{t=1}^n h_t(\mathbf{x}, y) \log \frac{1}{\alpha_t} \quad (2.16)$$

where $h_t(\mathbf{x}, y)$ represents the output of classifier h_t , expressed as the posterior probability for the class $y \in \mathcal{Y}$ given the feature vector $\mathbf{x}$. The resulting class y , given the input feature vector

$\mathbf{x}$, is the one that gets the maximum value. The weak learners are added incrementally to H . In each iteration t , RUSBoost applies RUS to randomly subsample the majority class in training set $\mathcal{D}$ until a subset $\mathcal{D}_t$ with a desired class distribution is reached. For example, if the desired class ratio r is 50:50, then the majority class examples are randomly removed until the numbers of majority and minority class examples are equal. Hence a weight α_t is assigned to the weak learner according to the relation:

$$\alpha_t = \frac{\epsilon_t}{1 - \epsilon_t} \quad (2.17)$$

where ϵ_t represents the pseudo-loss based on the original training set $\mathcal{D}$ and it is calculated as:

$$\epsilon_t = \sum_{(i,y):y_i \neq y} D_t(i)(1 - h_t(\mathbf{x}_i, y_i) + h_t(\mathbf{x}_i, y : y \neq y_i)) \quad (2.18)$$

where $h_t(\mathbf{x}_i, y_i)$ expresses the posterior probability of the classifier h_t for a class y_i that is different from the real class. D is a weight distribution for all examples in $\mathcal{D}$, which weights (D_t) are updated after each iteration as follows:

$$D_{t+1}(i) = D_t(i)\alpha_t^{\frac{1}{2}(1+h_t(\mathbf{x},y_i))-h_t(\mathbf{x},y:y \neq y_i)} \quad (2.19)$$

and then D_{t+1} is normalized to 1. Initially, the weight of each example $D_1(i)$ is set to $1/n$ where n is the number of examples in the training set.

2.6 Object-based Image Analysis in Agriculture

As mentioned in the Introduction, RS is a cost-effective means of contributing significantly to provide accurate and timely information in the agricultural sector. Among other important aspects, remotely sensed information allows stakeholders (Atzberger et al., 2015):

- to better anticipate the effects of weather phenomena and evaluate its risk over large agricultural areas;
- to establish profitable insurance systems as well as government subsidies that stimulate the investment of smallholders;
- to understand possible effects of climate change and identify areas with higher yield potential;
- to detect stress situations and agricultural anomalies; and
- to improve crop yield and reduce costs through a precision farming (i.e., a variable rate management).

Images with a sub-metric spatial resolution provide a more detailed overview of agricultural lands, which make them suitable for tasks involving the internal analysis of agricultural plots such as precision farming. However, conventional pixel-based methods reveal difficulties during analysis of these images regarding factors such as phenological variability of crops, and

different forms of management (Atzberger et al., 2015). OBIA offers an alternative to deal with these problems by considering all pixels inside a agricultural plot as a single processing unit, which can reduce the effect of these factors during automatic analysis. Object-based studies in agricultural landscapes exploit the fact that plot boundaries are relatively stable, while the cropping pattern changed frequently (Liu et al., 2006), to improve the analysis by reducing the image variability. Thus, accurate field boundary delineation is a prerequisite for a successful analysis (Forster et al., 2010).

In agriculture, diverse OBIA studies have been performed using both per-field and object-based frameworks. Per-field approaches integrate multi-spectral images and vector field boundaries included in GIS systems to perform the image analysis (Turker and Ozdarici, 2011; Ruiz et al., 2011; Conrad et al., 2010; De Wit and Clevers, 2004; Wassenaar et al., 2002; Pedley and Curran, 1991). The main idea behind this approach is to overcome the problems of misclassification caused by the within-field internal variability by using the spatial context given by the vector. Depending on when the integration of these two datasets is performed, two approaches can be identified:

Before classification. The vector data is used to subdivide the image into plots, then each plot is characterized by a set of attributes such as texture and spectral values, and finally a classification is performed based on these attributes, which avoids the intra-class spectral variations of conventional pixel-based approaches (Weng et al., 2010). Some agricultural studies based on this approach have been performed by Ruiz et al. (2011), De Wit and Clevers (2004), and Wassenaar et al. (2002).

After classification. In this approach, the spatial context of the pixels is used after a first conventional pixel-based classification. Thus, the image is firstly classified on a per-pixel basis and the resulting map is combined with digital vector data to finally assign a label to each field based on the modal class computed within the field. Some works that follow this approach in the field of agriculture are Turker and Ozdarici (2011), Conrad et al. (2010), and Pedley and Curran (1991).

Although several countries have GIS systems with the cadastral information of agricultural lands, as pointed out in Section 2.3, in many cases this data is not freely available or it is not accurate and up-to-date. In addition, the process of creating and updating this cadastral information is tedious, time-consuming and incurs high costs. Figure 2.7 shows an example of the agricultural cadastre of the Curicó Province in the Chilean Central Valley generated in 2013 by the Centro de Información de Recursos Naturales (CIREN⁴), Chile. As can be observed, it exists a disagreement between the vector data and the raster data (image). On the one hand, some edges are displaced with respect to agricultural plots perceived in the image. Moreover, in several cases, the vectors do not consider intra-plot subdivisions. As might be expected, these factors can considerably reduce the performance of the OBIA studies.

⁴www.ciren.cl. Last access on June 2016.

Fig. 2.7.: Agricultural cadastre, provided by the Centro de Información de Recursos Naturales (CIREN), Chile, is superimposed on the image. The delineation of agricultural plots is shown in red.

On the other hand, agricultural OBIA studies that follow an object-based framework generate image-objects through image segmentation algorithms. Examples of this are the works of Weissteiner et al. (2016), Li et al. (2016), Ozelkan et al. (2016), O'Connell et al. (2015), White and Roy (2015), Peña et al. (2014), Grecchi et al. (2014), Galletti and Myint (2014), Peña-Barragán et al. (2011), Duro et al. (2012), Aksoy et al. (2010), and Tansey et al. (2009). In this regard, most of the reported work either uses eCognition software to segment agricultural images. After the creation of image-objects, morphological features such as shape, area, length, width, perimeter, and spectral and textural features are then used collectively to perform image analysis. Having a sometimes huge set of attributes available, it may be very difficult for a human to identify the optimum descriptive attributes of the objects for a successful analysis. Therefore, it is necessary the use of machine learning techniques for finding an adequate separation of classes. In this context, Decision Trees (DTs) have been preferred to perform the statistical analysis in agricultural images. The main reason is that they are structurally explicit, allowing for clear interpretation of the links between the dependent variable of class membership and the independent variables of the data under analysis (Atzberger et al., 2015).

Peña-Barragán et al. (2011) combined the MRS algorithm and DT method for identification of 13 major crops cultivated in an agricultural area of California (USA). They explored the use of several spectral vegetation indexes and textural features derived from spectral bands of ASTER satellite scenes acquired during three distinct growing season periods. Their multi-seasonal assessment of crop types and field status reported a high accuracy. Vieira et al. (2012) combined MRS and DT method to map sugarcane ready to harvest in Brazil. To obtain the binary map indicating the area of sugarcane, four Landsat images (TM-5 and ETM+) collected between September 2000 and March 2001 were used. A large number of attributes were extracted for each image-object, including spectral, spatial, and textural features. Other works on the same line have been made by Galletti and Myint (2014), and O'Connell et al. (2015).

Although found more effective than the pixel-based analysis (Duro et al., 2012), object-based approaches are apparently parametrically complex and computationally expensive as they require prior segmentation at different scales (Ursani et al., 2012). Since MRS algorithm and most region-growing segmentation algorithms for RS images need user-supplied parameters, one of the challenges for their use is selecting suitable parameters to ensure best quality results (Ursani et al., 2012).

As observed along this Section, is a critical step during the OBIA process. However, it should not be forgotten that segmentation is not the end but only a part of the process that involves analyzing an image.

2.7 Segmentation not as End-itself

Image segmentation is an ill-posed problem, since it has no a unique solution (Hay and Castilla, 2008). For example, two human photo-interpreters can not delineate in the same way the same image (Figure 2.8), the same can happen when at two different moments in time, a same interpreter performs a delineation of the same scene.

Fig. 2.8.: Segmentation is an ill-posed problem: two subjects can delineate differently the same image. Delineation made by subjects A (Figure 2.8a) and B (Figure 2.8b) are shown in red.

Image delineation can be carried out by manual and automatic means, which have particular problems. Although manual delineation can be very precise, it is highly time-consuming. On the other hand, automatic segmentation methods faces two common problems (Figure 2.9): (i) under-segmentation, and (ii) over-segmentation. Over-segmentation occurs when a geographical object is partitioned into multiple smaller image-objects, whereas under-segmentation occurs when different real-world objects are grouped into a large image-object (Liu et al., 2010; Castilla and Hay, 2008). In general, over-segmentation is less problematic than under-segmentation, since it is easier to merge the segments than split them (Liu et al., 2010). Also, it does not exist a straightforward relationship between similarity in the image and semantic

similarity, therefore it is preferable to err on the side of over-segmentation and later joint segments by relaxing the external contrast criterion (Castilla and Hay, 2008) or by means of a posterior analysis, e.g., classifying and merging neighboring segments that belong to the same class.

Fig. 2.9.: Under-segmentation and over-segmentation errors generated by the same algorithm but with different parameter settings. Borders are shown in red.

As for the analysis process, Liu et al. (2010) point out that under-segmentation introduces classification errors because all pixels in each mixed image-object have to be assigned to the same class. Consequently, the highest classification accuracy that can be achieved for each under-segmented image-object, no matter what classifier is used, is always less than 100% (Liu et al., 2010). This percentage is determined by the ratio of the primary class in the image-object. In this regard, over-segmentation does not necessarily introduce errors to the subsequent classification. On the other hand, features extracted from mis-segmented (both under-segmented or over-segmented) image-objects do not represent the properties of geographical objects such as area and shape, as a result, they can severely reduce the classification accuracy (Song et al., 2005). In this context, the final performance of object-based classification is determined by both positive and negative effects due to the use of image-objects as units of analysis and the set of features extracted from them during classification (Liu et al., 2010). In addition, because OBIA approach changes the nature of the data analyzed, the training set is also affected. By working with image-objects instead of pixels as processing units, the number of samples available to the classifier is dramatically reduced. This implies that in general it is unfeasible to select a sufficient number of samples to represent adequately all thematic classes in an image (Myburgh, 2012). Therefore, supervised methods that perform well under conditions of limited training set size are crucial in OBIA approaches (Myburgh, 2012).

Although most scientists associate the results of OBIA studies to optimal segmentation (Yang et al., 2015; Liu et al., 2012; Draguț et al., 2010), recent works have shown that obtaining

an optimal segmentation do not necessarily play a major role in OBIA research (Blaschke et al., 2015; Blaschke et al., 2014). Since obtaining such segmentation wherein a single object is associated with a unique segment is difficult to automate and is not required from the perspective of simply providing spatial context (Priya et al., 2015). In this background, some works (Priya et al., 2015; Garcia-Pedrero et al., 2015a) pose an alternative approach consisting of an intentional over-segmentation of the image, resulting in homogeneous regions called superpixels.

2.7.1 Superpixel Segmentation

In recent years, new types of image processing methods, known as superpixel methods (Neubert and Protzel, 2012), have been developed in the area of computer vision (Achanta et al., 2012; Van den Bergh et al., 2012; Liu et al., 2011; Li et al., 2012; Ren and Malik, 2003). A superpixel is a small, local, and coherent cluster that contains a statistically homogeneous image region according to certain criteria such as color, texture, among others (Ren and Malik, 2003). Superpixels are a form of image segmentation, but the focus lies more on an image over-segmentation, not on segmenting meaningful objects (Schick and Stiefelhagen, 2011). In this regard, superpixel processing is not seen as an end in itself but rather a preprocessing step in order to solve a major problem, in this case, the efficient analysis of a scene. Superpixel techniques enhance image analysis, *e.g.*, reducing the influence of noise and intra-class spectral variability, preserving most edges of images, and improving the computational speed of later steps such as classification, clustering, segmentation (Achanta et al., 2012). Moreover this type of processing is a pragmatic alternative when image objects have diverse scales or when the objects are not known in advance (Thompson et al., 2010).

Basically a superpixel method must fulfill the following requirements (Achanta et al., 2012; Liu et al., 2011):

- Superpixels should adhere well to image boundaries. In this context, object boundaries should be a subset of superpixel boundaries.
- Superpixels should be fast to compute, memory efficient, and simple to use. These features are essential to reduce the computational complexity of the latter image processing steps of a real-world application.
- Moving from pixels to superpixels should not reduce the attainable yield of the intended application, but increase the speed and improve the quality of its results.

Other desirable property is the control over the amount and compactness (spatial homogeneity criterion) of generated superpixels, which allows to adapt superpixels to specific requirements for real-world applications (Stutz, 2015). In particular, the control over these factors permits to generate superpixels that better capture spatial coherent information (Schick et al., 2014). However, these parameters must be explainable and easy to adjust.

Since the concept of superpixel was established by Ren and Malik (2003), various algorithms for superpixel segmentations have been proposed. They can be roughly grouped into graph-based and gradient-ascent methods. Algorithms in the first category usually treats pixels as nodes in a graph where edge weights are proportional to the similarity between neighboring pixels. Superpixels are created by minimizing a cost function defined over the graph. Examples in this group include: Normalized Cuts (NC; Mori, 2005; Shi and Malik, 2000), Felzenszwalb-Huttenlocher Segmentation (FHS; Felzenszwalb and Huttenlocher, 2004), Superpixel Lattices (SL; Moore et al., 2008), and Entropy Rate Superpixel Segmentation (ERS; Liu et al., 2011). On the other hand, gradient-ascent approaches to form superpixels usually start from a rough initial clustering of pixels and then iteratively refine the clusters until some convergence criterion is met. Examples in this group include: Mean Shift (MS; Comaniciu and Meer, 2002), Quickshift (QS; Vedaldi and Soatto, 2008), Turbopixel Segmentation (TP; Levinshtein et al., 2009), and Simple Linear Iterative Clustering (SLIC; Achanta et al., 2012).

Table 2.2 shows a general overview of the superpixel algorithms, considering factors such as their computational complexity, if they allow some control over the amount (CoA) and compactness (CoC, spatial homogeneity criterion) of superpixels, and finally, their main issues in accordance with the reviews of Achanta et al. (2012), Neubert and Protzel (2012), and Schick et al. (2014).

Tab. 2.2.: Overview of the superpixel algorithms.

Method	Computational Complexity*	CoA*	CoC°	Main issues
NC	$O(N^{2/3})$	Yes	No	It is slow and its boundary adherence is relatively poor.
MS	$O(N^2)$	No	No	It produces irregularly shaped superpixels of nonuniform size.
FHS	$O(N \log N)$	No	No	The resulting segments vary strongly in shape and size.
SL	$O(N^{2/3} \log N)$	Yes	No	It is slow and its boundary adherence is relatively poor.
QS	$O(dN^2)^\dagger$	No	No	It is faster than MS but yields inferior segmentation results.
TP	$O(N)^+$	Yes	No	It exhibits relatively poor boundary adherence.
ERS	$O(N \log N)$	Yes	No	It converts the RGB image to a grayscale representation.
SLIC	$O(N)$	Yes	Yes	It requires a post-processing step to enforce the connectivity of superpixels.

* N is the total number of pixels in the image.

* Control over the amount of superpixels.

° Control over the compactness of superpixels.

† d represents a small constant.

+ It is the complexity reported by their authors, however this method relies on algorithms of varying complexity (Achanta et al., 2012).

These methods are originally designed to work in a RGB spectral space, or any similar color space limited to three channels such as CIELab. It exists, however, some extensions that use depth as an additional channel to perform segmentation on RGB-D images (Weikersdorfer et al., 2012). From algorithms shown in Table 2.2, SLIC method has demonstrated to efficiently generate superpixels that show a good adherence to image boundaries, allowing to select the amount and compactness of superpixels. In this regard, SLIC algorithm is considered as a benchmark method for performing superpixel segmentation, and therefore it is explained below.

Simple Linear Iterative Clustering

Simple Linear Iterative Clustering (SLIC) algorithm, proposed by Achanta et al. (2010), is based on well-known *k-means* method to group pixels in RGB/CIE Lab color space. SLIC superpixels are generated according to two criteria: spectral similarity (limited to three channels) and spatial proximity.

The generation of superpixels is based on the assumption that limiting the search space to a region proportional to the desired superpixel size reduces considerably the calculation time. In fact, its computational complexity is linear in the number of pixels in the image (Achanta et al., 2012).

The first step of the segmentation framework begins with the sampling of K initial cluster centers on a regularly spaced grid of g pixels. In an RGB image, the initial centers are defined as:

$$C_k = [p^R, p^G, p^B, x, y]^T, \quad k = 1, 2, \dots, K \quad (2.20)$$

where p^b represents the spectral value in channel b – th of pixel p at position x and y . R , G , and B correspond to red, green, and blue spectral bands, respectively. To produce similar sized superpixels, the grid interval is defined as $g = N/k$, where N is the total number of pixels in the image. g determines the size of the superpixels, therefore, the greater value of g , the larger the superpixels. To avoid placing a center in a border or in a noisy pixel, initial cluster centers are moved to the pixel with the lowest magnitude gradient in a 3×3 neighborhood (Kim et al., 2013).

In the next step, each pixel p is associated to the nearest cluster center whose search space overlaps its location. The search region is enclosed to an area of $2g \times 2g$ pixels around each superpixel center. This approach not only reduces distance computations but also makes complexity independent of the number of superpixels. Then the cluster centers are updated to be the mean vector of all pixels belonging to the cluster. Both steps are repeated iteratively until a maximum of 10 iterations, since no further significant changes in superpixel quality could be observed (Achanta et al., 2010).

The clustering distance, based on an Euclidean distance, is a weighted relationship between spectral and spatial measures. The first measure ensures superpixel homogeneity, and the second one enforces compactness and regularity in superpixels shape. In order to work with multispectral images, the spectral square distance between pixels i and j is defined as follows:

$$d_c^2 = (p_i^R - p_j^R)^2 + (p_i^G - p_j^G)^2 + (p_i^B - p_j^B)^2 \quad (2.21)$$

The spatial square distance is calculated as:

$$d_s^2 = (x_i - x_j)^2 + (y_i - y_j)^2 \quad (2.22)$$

Fig. 2.10.: Search area of k-means and SLIC algorithms: in the conventional k-means algorithm, distances are computed from each cluster center to every pixel in the image (Figure 2.10a), whereas SLIC (Figure 2.10b) only computes distances from each cluster center to pixels within a $2g \times 2g$ region.

where x and y denote the position of the pixel. Finally, the clustering distance is calculated as:

$$D = d_c + \left(\frac{c}{g}\right) d_s \quad (2.23)$$

where c controls the compactness of the superpixels.

After performing the last iteration, like other algorithms based on clustering segmentation, SLIC needs a post-processing step in order to ensure that all superpixels fulfill the spatial proximity criterion. Thus isolated pixels and small segments are merged with its larger neighbor (Kim et al., 2013). An example of applying SLIC method to an RGB image is shown in Figure 2.11, as can be observed, superpixels adhere well to the boundaries of spectrally homogeneous regions.

Fig. 2.11.: Example of an RGB image segmented by SLIC. Borders of superpixels are shown in green.

A comprehensive review of the literature related to superpixels in the area of remote sensing from 2008 to 2015 has been realized. This review reveals a total of 64 publications comprising 27 articles in journals and 37 conference papers. As can be observed in Figure 2.12, the number of publications shows a growing trend, with 11 publications in JCR journals and 18 publications in peer-review conferences in 2015.

Fig. 2.12.: Publications per year using the concept of superpixels in remote sensing area.

Of the above publications, superpixel algorithms has been mainly used in RS to process Synthetic Aperture Radar (SAR, with 17 papers), hyper-spectral (17 papers), multi-spectral (17 papers, including both VHR and low resolutions), Aerial (9 papers) and Panchromatic (1 paper) images. In these works, superpixels are generally used as a preprocessing step in order to provide a basis to perform other complex tasks such as classification (25 papers), or even serving as a starting point of more complex segmentation algorithms (16 papers). Other applications are change detection (7 papers), object detection (7 papers), clustering (1 paper) and dimensionality reduction (1 paper). In this regard, superpixels appear to be a promising approach to object-level analysis of RS images. Moreover, all papers consulted for the development of this review have been summarized in Figures A.1 and A.2. These Figures show the papers, published in conferences and JCR journals, respectively, that use superpixels in RS applications.

2.8 Discussion

As seen along this Chapter, the object-based processing posed some difficulties for analyzing agricultural fragmented landscapes. In particular, these landscapes present a high spatial heterogeneity produced by the diversity in sizes, shapes, and crops of the different agricultural plots. Moreover, the plot boundaries are not always well defined, showing indistinct boundaries between neighboring fields, and between fields and natural vegetation. These factors cause segmentation methods in which objects-based approaches are based, fail to correctly segmenting agricultural fields. In this regard, most of these methods are highly dependent on a correct parameter selection that requires a prior knowledge about the scene or tuning by trial-error for correctly segmenting all objects.

Under these circumstances, it is assumed that segmentation is an ill-posed problem and therefore it should not be seen as an end in itself but simply as another part of the analysis process. In this context, the use of novel concepts coming from computer vision area are used to perform an object-based processing of RS imagery, in particular, the use of superpixels is explored in the following Chapters. To this end, a superpixel method for processing multi-spectral images as well as other RS data is proposed in Chapter 3.

In order to provide useful information on the spatial and temporal dynamics of agricultural fields raised as the objective of this thesis, the proposed superpixel method is exploited in different applications at three scales of interest: (i) to analyze agricultural fragmented landscapes, (ii) to delineate agricultural plots, and (iii) to capture the internal variability of agricultural plots. Figure 2.13 shows the different scales of analysis.

- **Analyzing agricultural fragmented landscapes.** Two methodologies are developed for the analysis of agricultural landscapes. In Chapter 4 it is proposed a framework to provide multi-scale representation of agricultural landscapes and, at the same time, a way to identify the best scale of analysis according to the spectral intra-variability and inter-heterogeneity of the regions obtained on each scale. Chapter 5 presents a framework that combines two scales corresponding to superpixels and edge-based segments for mapping diverse land-covers in multi-spectral images of fragmented agricultural landscapes.
- **Delineating agricultural plots.** The delineation of agricultural plots, established as the first specific objective of this dissertation, is addressed by two novel different approaches in Chapter 6. The first explores the use of a classification method for segmenting agricultural plots by agglomerating superpixels, posing an alternative to traditional segmentation methods that is based on learning how to segment images in a similar way as a human operator, as it does. The second one proposes an approach based on the consensus of diverse segmentation results for delineating of agricultural plots. Both approaches pursue to improve the boundaries of agricultural plots necessary for the study of the variability inside them.

- **Capturing the internal variability of agricultural plots.** The spatial and temporal analysis of the variability at plot level (specific objective 2) is addressed in Chapter 7. In this regard, a novel superpixel method that combines spatial and temporal components for analyzing multi-temporal imagery is presented.

Fig. 2.13.: Three scales of analysis of agricultural landscapes.

Part III

Supapixel-based Approaches for the
Spatio-Temporal Analysis of Agricultural
Landscapes

A Superpixel Method for Remote Sensing Applications

” *Two possibilities: making oneself infinitely small or being so. The second is perfection, that is to say, inactivity, the first is beginning, that is to say, action.*

— **Franz Kafka (1883–1924)**

The Zürau Aphorisms.

3.1 Introduction

Despite a widespread use in the area of computer vision, superpixels are still in process of being widely used in remote sensing applications (see Section 2.7.1). One of the main reasons limiting its adoption is that most superpixel methods are only able to handle images with three spectral channels (e.g., color images), whereas remote sensing applications require handling images with other characteristics, such as more information from electromagnetic spectrum recorded in images with more than three spectral channels. As established by Schiewe (2002), a segmentation algorithm must meet certain criteria to be suitable for remote sensing applications, such as handling various input data simultaneously. This input data includes multi-spectral data, which increases the redundancy of the data, as well as other data sources such as vector data (GIS data), which provide spatial information about image-objects through its delineation by points, lines, and polygon shapes.

In this Chapter, a superpixel method for analyzing remote sensing data is presented. The proposed method is an extension of the Simple Linear Iterative Clustering (SLIC) method described in Section 2.7.1. The proposed version of SLIC is adapted to Schiewe (2002) criteria, thus is able to handle multi-spectral (i.e., images with more than three spectral channels) as well as vector data. In addition, some improvements regarding the post-processing step have been realized. In the original version, isolated pixels and small segments are merged with its larger neighbor, however this can results in misclassification errors, because do not consider the spectral similarity between regions. Therefore, this process has been improved to merged this small regions with its most spectrally similar neighbor.

3.2 SLIC for Remote Sensing Applications

The first modification of SLIC algorithm extends the definition of spectral proximity provided by the original method to work with multi-spectral images of NB bands. Since the SLIC algorithm is based on the k -means method, which is a multi-dimensional algorithm grouping, to provide the feature of handling multispectral images to this new version two main changes are performed to the algorithm, particularly, the initialization step and the calculation of the spectral distances suffered slightly changes. The entire algorithm is summarized in Algorithm 3.1.

Algo. 3.1: SLIC algorithm for remote sensing data.

Input: A multi-spectral image.

Vector data containing the areas of interest ref (optional, if it is not specified, ref covers the entire image).

Output: A superpixel segmentation S .

/* Initialization */

Initialize cluster centers by sampling pixels at regular grid steps g covering ref area.

Move cluster centers to the lowest gradient position in a 3×3 neighborhood inside ref area.

Set label $S(i) \leftarrow -1$ for each pixel $p_i \in ref$.

Set distance $d(i) \leftarrow \infty$ for each pixel $p_i \in ref$.

/* Calculation of distances */

repeat

for each cluster center C_k **do**

for each pixel p_i in a $2g \times 2g$ region $\in ref$ around C_k **do**

 Compute the distance D between C_k and p_i .

if $D < d(i)$ **then**

 Set $d(i) \leftarrow D$.

 Set $S(i) \leftarrow k$.

end if

end for

end for

 Compute new cluster centers.

until the maximum of iterations is not reached

/* Post-processing */

$R \leftarrow$ segments which size $< th$

for each $r_i \in R$ **do**

$N \leftarrow$ the neighbor $\in ref$ with the minimum spectral distance respect segment r_i .

for each $r_n \in N$ **do**

 Compute the region r_n with the minimum spectral distance D' .

end for

 Set r_i label to its the label of the neighbor r_n with the minimum spectral distance D' .

end for

Relabel all pixels in S

return S

Starting from the original sampling of K initial cluster centers on a regularly spaced grid of g pixels. The initial centers defined in Equation 2.20 are now defined as follows:

$$C_k = [p^1, p^2, \dots, p^{NB}, x, y]^T, \quad k = 1, 2, \dots, K \quad (3.1)$$

where p^b represents the spectral value in band b^{th} of pixel p at position x and y .

The clustering distance that measures the spectral similarity of one pixel to the centroid (Equation 2.21), should be extended to multi-spectral pixels. Then, the spectral square distance between pixels p_i and p_j is calculated as follows:

$$d_c^2 = \sum_{b=1}^{NB} (p_i^b - p_j^b)^2 \quad (3.2)$$

On the other hand, the spatial distance (Equation 2.22) remains the same.

To handle vector data, a premise is adopted and it is that the vector data provides the spatial information about the area (*ref*) where the superpixels must be generated. Thus during initialization, calculation of distances, and post-processing steps only pixels within *ref* are processed. This improvement allows to generate superpixels in only certain regions rather than the whole scene, slightly reducing runtime in image processing. This is relevant to the study of specific areas of the image such as the inside of the agricultural parcels. Figure 3.1 shows an example of applying the proposed version of SLIC to a multi-spectral image but only to an specific area delineated by available vector data (shown in color green). It is important to note that if the vector data is not provided then the algorithm processes the entire image.

Fig. 3.1.: Example of applying superpixel method to an area of the multi-spectral delineated by vector data. This area (*ref*) is colored in green.

Finally, the last improvement corresponds to the post-processing step. The main modification with regard to the original version is that small regions are now merged with it most spectrally similar neighbor, which avoid the original misclassification errors. The spectral distance (D') is

expressed as the difference between the mean values (μ) of the small region i and its neighbor n , and is calculated as:

$$D' = \sqrt{\sum_b^{NB} (\mu_i^b - \mu_n^b)^2} \quad (3.3)$$

Thus a small region acquires the label of the neighboring region with the smallest spectral distance. In this work, regions with a size lower than a fifth of the desired superpixel size, has been considered small, however this threshold (th) can be manually set. In Figure 3.2, the results of the original and proposed post-processing approach are shown. As can be observed in the rectangle, the proposed approach have a better adherence to image borders than the original post-processing approach.

Fig. 3.2.: Example of applying post-processing techniques to an RGB image. Superpixel borders are colored in red, whereas enclosed in a rectangle, the difference between both approaches is shown.

3.3 Conclusions

In this Chapter, a modified version of the SLIC method for remote sensing data was presented. The new version is able to handle multi-spectral images as well as vector data. In the latter case, vector data provides the area in which SLIC processes are performed. In addition, some improvements regarding the post-processing step was performed in order to avoid possible misclassification errors. In this regard, the proposed post-processing approach obtained slightly better results regarding their adherence to image borders. This method has been extensively used along this dissertation, thus, the results of applying it to RS data are shown in next Chapters.

Optimal Scale of Analysis in a Hierarchical Superpixel-based Segmentation Method

” *From the landscape: a sense of scale. From the dead: a sense of scale.*

— **Richard Siken**
Detail Of The Woods.

4.1 Introduction

The concept of scale of an object can be defined as the level of aggregation and abstraction at which this object can be described (Benz et al., 2004). In this regard, Woodcock and Strahler (1987) expounded the importance of the factor of scale in the identification of image-objects present in remote sensing imagery, which depends greatly on the relationship between the spatial resolution of the images and the size of the image-object. Thus, the bigger the object is with respect to the image spatial resolution the more likely it is to be fragmented. On the other hand, if the size of the image-object is very small in relation to the spatial resolution it could happen that the object is not identified. Therefore, the availability of high resolution satellite images is essential to allow small objects to be identified on fine scales. Although it is a fact that satellite images with high spatial resolution exhibit high levels of detail; traditional image processing algorithms based on single pixel analysis are often not capable of extracting the entire information; this situation potentially leads to lower accuracy in the traditional remote sensing tasks such as classification, and change detection (Myint et al., 2011).

As mentioned in Section 2.3, the two main steps in the OBIA approach are an initial segmentation followed by an analysis process. Numerous researches have shown that the segmentation process is a critical step in the OBIA approach. This is due to the fact that the results of the classification step strongly depend on the initial segmentation (Gao and Mas, 2008; Yan et al., 2006).

Furthermore the accepted idea is that natural processes are scale dependent, it is also known that most of them are hierarchically structured. Therefore an appropriate understanding of these processes and interactions between them requires the determination of optimal study

scales (Woodcock and Strahler, 1987; Levin, 1992; Treitz and Howarth, 2000) and the spatial relationships between the regions at different scales (Burnett and Blaschke, 2003). These aspects are especially interesting for the analysis of crops, as their inherent complexity makes multi-scale analysis a requirement in dealing with the different sizes of the interest objects in the image (Baatz and Schäpe, 2000; Duveiller and Defourny, 2010).

Some studies have designed methodologies for estimating a single optimal scale for the segmentation for the whole scene (Espindola et al., 2006; Kim et al., 2009). Though other studies have implemented a multi-scale approach (Zhou and Troy, 2009; Trias-Sanz et al., 2008) to address this problem, in this cases the selection of the segmentation scales is typically based on an extensive knowledge of the study site (Johnson and Xie, 2011). Therefore, as far as our knowledge is concerned, the unsupervised definition of the most suitable scale for each land cover in a scene remains unsolved.

The concept of multi-scale/multi-resolution inherent to OBIA methodologies allows the scale problem to be dealt with (Burnett and Blaschke, 2003), by defining a hierarchy of objects at different scales through hierarchical segmentation algorithms.

The advent of the multi-resolution segmentation algorithm (Section 2.4.2) by Baatz and Schäpe (2000), available in eCognition software, was a milestone in research into multi-scale segmentation for RS images. The implemented multi-resolution segmentation technique uses homogeneity criteria based on color and shape, and a scale parameter in combination with local and global optimization techniques. The tuning of different parameters requires an expert knowledge, not always available; moreover, this task usually requires a lot of time. A fuzzy approach which automatically selects the parameters of the segmentation used in multi-resolution approach was proposed by Maxwell and Zhang (2005). In any case, the license cost of eCognition software prevents its use in many applications. Since then, other multi-resolution/hierarchical segmentation techniques have been proposed in the literature (Chen et al., 2003; Li et al., 2008).

A superpixel-based methodology to generate multi-scale segmentation as well as a mechanism to find an optimal analysis scale are explored in this Chapter. The proposed methodology is built on the superpixel approach proposed in Section 3. Two main contributions are achieved in this Chapter:

- An adapted multi-scale hierarchical segmentation algorithm to deal with objects that are scale dependent and hierarchically related in a satellite scene.
- A method for selecting the local optimal scale associated to different regions in the images. The selection of the optimal multi-scale segmentation is carried out by evaluating the intra-variability and inter-heterogeneity of the regions defined by the coarsest scale. This methodology is based on the evaluation method presented by Johnson and Xie (2011),

where the complete scene is used for evaluation, but adapting the one to deal with the hierarchical image segmentation process.

The methodology has been applied to a scene corresponding to an agricultural area, where different types of land covers, showing a high variability, are represented.

4.2 Materials and Methods

4.2.1 Data Set

To test the methodology proposed in this work a scene of an agricultural area with a high local variability has been selected. The study area is located in the central irrigated valley of Biobio Region, Chile (Figure 4.1). It corresponds to a rotation sugar beet crop of Sandrina, Magnolia, and Donella varieties. The plantation frame is 0.5 m between rows and 0.07 m between plants, with a center pivot irrigation system. The total area of the farm is around 450 ha, and the area of crop covered for the center pivot (study site) is around 17 ha. The corresponding multi-spectral satellite image was collected on January 18th, 2013 by the Pléiades satellite¹. The coordinates of the center of the scene are: 36°31'54" S; 72°08'03" O. The area of the scene is 420 hectares and corresponds to 1024x1024 pixels. The collected image has a spatial resolution of 2 meters and corresponds to an ORTHO-BASIC product, characterized by a basic radiometric normalization (for the calibration of the detector) and geometric correction (WGS84/UTM 19S projection). The spectral and spatial resolutions are presented in Table 4.1.

Tab. 4.1.: Spectral and spatial characteristics of the multi-spectral image taken by the Pléiades satellite.

Band Name	Band Width (<i>nm</i>)	Spatial Resolution (<i>m</i>)
blue	430-550	2
green	490-610	
red	600-720	
near infrared	750-950	

A color composition of the described multi-spectral scene is shown in Figure 4.2.

4.2.2 Methodology

The proposed methodology is made up of two main processing modules: (i) image segmentation, and (ii) segmentation evaluation. They are detailed below:

¹www.landinfo.com/pleiades-satellite-imagery.htm. Last access on June 2016.

Fig. 4.1.: The study area is located in Biobio Region, Chile. The black star (on the right map) shows the location of the study site.

Fig. 4.2.: Color composition (near infrared-green-blue) of the analyzed scene. The satellite image was collected on January 18th, 2013 by Pléiades Satellite, in the central irrigated valley of Chile ($36^{\circ}31'54''$ S; $72^{\circ}08'03''$ O).

Hierarchical image segmentation

Even though superpixels present some good properties mentioned before in Section 3, SLIC it is not able to generate a hierarchical segmentation. The proposed approach fills this gap

by proposing an adaptation to SLIC that makes the generation of a hierarchical multi-scale segmentation possible. Based on the fact that in SLIC, low values of k results in a coarser segmentation in comparison to finer segmentations obtained increasing the values of k , we propose to generate segmentations at different scales by varying the size of the superpixels and then combining the resulting segmentations. To establish the different segmentation scales, a vector $K = [k_1, k_2, \dots, k_n]$ containing different values for the size of the superpixels in ascending order is defined. Each value of the vector corresponds to a different scale (i). Consequently, segments corresponding to a segmentation carried out according to k_i will be larger than those of the segmentation according to k_{i+1} . A segment A^i in scale i is considered as the parent of a segment B^{i+1} in scale $i + 1$, if the area of B^{i+1} can be entirely covered by segment A^i . The parent-children relationship between segments is shown in Figure 4.3. In order to ensure that the resulting segmentation process is hierarchical, segments on scale k_{i+1} are partitioned by segments on scale k_i , obtaining a new segmentation that generally contains a higher number of segments than in the segmentation achieved in scale k_{i+1} . To match the number of superpixels in scale $i + 1$ to its original number of superpixels, smaller size segment are joined to the most similar neighboring segment belonging to the same parent. We use euclidean distance in the spectral space (color) to define the similarity between neighboring segments. A description of the proposed method is show in Algorithm 4.1.

Fig. 4.3.: Segment A^i is considered the parent of segments B^{i+1} and C^{i+1} , as both segments can be covered entirely by segment A^i . Each square of the grid represents an image pixel.

Determination of the optimal local scale

One important contribution of the proposed methodology is that it provides different optimal scales for different image regions, against the results shown in the work of Johnson and Xie (2011), where an optimal scale for the whole image is yielded. This fact is especially relevant in images where the inter-region variability is high, as it is the case of some crops in satellite images.

A general overview of the scheme proposed is shown in Figure 4.4. In order to determine the optimal scale for each image region generated by the hierarchical segmentation process, the following inter and intra-segment quality measures are used:

Algo. 4.1: Hierarchical SLIC

Input: A vector with the different superpixel sizes (K)

A global compactness factor (c)

An image (I)

Output: A hierarchical multi-scale image segmentation (S)

$S_1 \leftarrow \text{slic}(I, K_1, c)$ /* First level of segmentation */

for $i=2$ **to** $\text{length}(K)$ **do**

$L \leftarrow \text{slic}(I, K_i, c)$ /* i -th level of segmentation */

$S_i \leftarrow$ intersection between S_{i-1} and L

/* Enforcing connectivity of oversegmented regions */

$s \leftarrow$ segments in S_i sorted by size in ascending order

$j \leftarrow 1$

while Number of segments of S_i and L are not equal **do**

$N \leftarrow$ neighbors of s_j

for each region $r \in N$ **do**

Calculate the similarity between s_j and r

end for

$S_i \leftarrow$ merge region s_j with the r with minimum distance

end while

end for

return S

- the global Moran index (Fotheringham et al., 2000) is used to measure the inter-heterogeneity between segments;
- the weighted variance is used for intra-segment homogeneity.

Fig. 4.4.: General overview of the evaluation process.

The global Moran index measures the correlation between segments. It is used to evaluate the similarity of a segment with its neighborhood and is calculated as follows:

$$MI_i^p = \frac{n \sum_{u=1}^n \sum_{v=1}^n w_{uv} (\hat{y}_u^i - \hat{y}^p)(\hat{y}_v^i - \hat{y}^p)}{\sum_{u=1}^n (\hat{y}_u^p - \hat{y}^p)^2 \left(\sum_{u \neq v} \sum w_{uv} \right)} \quad (4.1)$$

where:

- n is the total number of regions on the analysis scale that are related to the parent segment p .
- w is an adjacency matrix among neighboring regions that are relatives of the segment p on scale 1 (parent), therefore w_{uv} is equal to 1 if regions u and v are adjacent, otherwise, $w_{uv} = 0$.
- $\hat{y}_u^i$ is the spectral average of a particular region (u) on scale i .
- $\hat{y}^p$ is the spectral average of a particular region on scale i .

In order to calculate the similarity on a particular scale level only those regions belonging to a certain parent segment p on scale 1 are considered. The Global Moran index ranges from the value -1 , that corresponds to the perfect dispersion, to 1 , that indicates a complete correlation. The 0 value corresponds to a random correlation.

The weighted variance is used for intra-segment quality measurement, and is calculated as follows:

$$wVar^i = \frac{\sum_{u=1}^n a_u^i v_u^i}{\sum_{u=1}^n a_u^i} \quad (4.2)$$

where:

- v_u^i is the variance of the pixels that compounds the segment u on scale i .
- a_u^i corresponds to the size of segment u in scale i .
- n is the total number of regions on the analysis scale that are related to the parent segment p .

Values of variance close to zero indicate homogeneity of a segment. Note that, when calculating this measurement, larger sized segments have a greater effect on the evaluation than those of a smaller size. The best scale is defined as the value of the objective function (F^i) closer to zero, F^i being calculated by adding the normalized (range $0 - 1$) of intra and inter-segments measurements. The objective function is calculated for each scale i as:

$$F^i = wVar^i + MI^i \quad (4.3)$$

where:

- the value of the objective function calculated using the parent segment (p), at the scale 1^{st} , and the each related segments at scale i .

Experiments

Two different experiments were carried out. In the first one, the number of superpixels associated with each scale was generated with a linear spacing. In the second one, the number of superpixels on each scale was generated with a dyadic spacing (2^n). The number of scales used in both experiments was 7, varying from 64 to 4,096 superpixels. Table 4.2 details the number of superpixels for each scale and for each approach. It can be noted that except for the first and the 7^{th} scale, the number of superpixels is always larger for the linear approach than for the dyadic one.

Tab. 4.2.: K -vectors values for linear and dyadic approaches.

	Scale						
	1	2	3	4	5	6	7
Linear	64	736	1048	2080	2752	3424	4096
Dyadic	64	128	256	512	1024	2048	4096

4.3 Results and discussion

In this work the four spectral bands described in Table 4.1 of the multi-spectral image (Figure 4.2) were used in the segmentation process and in the calculation of the metrics to determine the optimal scale for each parent superpixel.

The results are detailed in Figures 4.5, 4.6, and 4.7. The first column corresponds to the results obtained with the linear approach, whereas the second row shows the results obtained with the dyadic approach. The optimal scales provided for the hierarchical segmentation algorithm for each parent superpixel on the 1^{st} scale are represented in the first row (Figures 4.5a and 4.5b). Figures 4.6a and 4.6b represent the spatially distributed optimal scale values (from 1 to 7). Figures 4.7a and 4.7b display the boundaries of the segments for the best scales.

As above mentioned, the number of superpixels is always larger for the linear approach than for the dyadic one for each particular scale. Thus it can be seen in Figures 4.5a and 4.5b that in the linear approach most of the parent superpixels have assigned optimal scales from 4^{th} to 7^{th} , whereas coarser scales are assigned in the linear one. The higher number of superpixels on the intermediate scale for the linear approach (see Table 4.2) produces an abrupt change from coarsest to finest scale, giving rise to an over-segmentation in some homogeneous areas, as can be appreciated in Figures 4.6a, 4.6b, 4.7a and 4.7b. It can be observed that there are different well-recognized structures in Figure 4.6b but not in Figure 4.6a, such as: plots, roads and irrigation circular areas, known as pivots. This shows that the dyadic approach provides more realistic local optical scales than the linear one. In addition, in Figure 4.6b it is clear that

(a) Linear scale

(b) Dyadic scale

Fig. 4.5.: Optimal scales for each parent superpixel.

different scales are associated to different regions. Figure 4.8 shows the final segmentation result for the experiment in which the scales were generated through dyadic spacing. This image has been obtained from the Infrared band by assigning to each segment the average digital value of the pixels included in it. It can be observed that agricultural management units with high variability in the original scene has been identified as homogeneous areas, it is very important for generating information to improve the agricultural management.

4.4 Conclusions

The two main contributions of this work are: (i) an adapted multi-scale hierarchical segmentation algorithm for satellite images based on the superpixel algorithm proposed in Section 3 and (ii) a method to select the local optimal scale associated to different regions in the images. The determination of the local optimal scales is carried out by evaluating the intra-variability

Fig. 4.6.: Optimal scales spatially distributed. Red colors indicate small segments, whereas blue ones indicate larger segments.

Fig. 4.7.: Boundaries of each segment according to the optimal scale is colored in white.

and inter-heterogeneity of the regions obtained on each scale in respect to the parents-regions defined by the coarsest scale. An objective function that combines weighted variance and the global Moran index, was used to select the optimal scale. The number of regions on each scale has been determined in two different ways: linearly and dyadically spacing. From the results it can conclude that the method proposed allows users to benefit from detecting objects in an image at different scales. This provides the user with a better comprehension of the land-cover, its objects and phenomena. This fact is especially relevant in images where the inter-region variability is high, as is the case of agricultural lands in satellite images. It should be noted that the segmentation results are highly dependent on the approach used to determine the scales. The dyadic approach provides more realistic local optical scales, in the sense that, in the segmentation provided by this approach, different structures (plots, roads, and pivots)

Fig. 4.8.: Segmented image through the dyadic approach.

can be well recognized but not in the segmentation based on the linear approach, due to the over-segmentation produced in this case. Although this approach provides adequate scales from an inter and intra-segment homogeneity point of view, these scales do not correspond in all cases with the natural boundaries of the agricultural plots.

The results of this Chapter have been presented at one conference, and a extended version was later accepted for publication in a JCR journal (see Section 8.2).

A Dual-Scale Segmentation Methodology for Analyzing Fragmented Agricultural Landscapes

” *Human perception is by nature a dualistic phenomenon
–which makes the quest for enlightenment an uphill
struggle, to say the least.*

— **Douglas R. Hofstadter (1945–)**
Gödel, Escher, Bach: an Eternal Golden Braid.

5.1 Introduction

An automatic OBIA methodology for fragmented agricultural landscapes focusing on the generation of a thematic map of the available classes in the scene is proposed in this Chapter. The proposed methodology employs superpixels, instead of image pixels, as minimal processing units of analysis and provides a link between them and meaningful objects (agricultural plots) in order to perform the analysis of the scene. To capture the phenomena (small agricultural plots) occurring at coarser scales (bigger segments), an edge-based method is used to segment the image. The superpixel method proposed in Section 3 together with the information of based-edge segments (in form of vector data) are used to generate superpixels that exactly correspond with their surface. Superpixel and edge-based scales are used together, in a object-based classification approach, to generate a thematic map that combines both features.

5.2 Materials and Method

5.2.1 Data Set

The study area (Figure 5.1) corresponds to a Chilean central valley located in Valparaíso Region ($70^{\circ}40'7.17''$ W, $32^{\circ}48'11.28''$ S), and covers approximately 2660 ha. It is mostly characterized by small agricultural plots (5 ha on average) with crops of full canopy coverage and orchards. A WorldView-2 dataset acquired on December 3rd, 2011, is used for the analysis. The data corresponds to an Ortho-Ready Standard 2A product, and thus has already been radiometric and geometrically corrected. Available panchromatic (PAN) and multi-spectral (MS) bands are

also co-registered. The spectral and spatial properties of the WorldView-2 image are shown in Table 5.1. The full scene has dimensions of 10217 columns and 10397 rows in the PAN band.

Fig. 5.1.: The study area is located in Valparaíso Region, Chile. The black star (on the right map) shows the location of the study site.

Tab. 5.1.: Spectral properties of the WorldView-2 image.

Band	Spectral range (μm)	Resolution (m)
blue	0.45 – 0.52	2.44
green	0.52 – 0.60	
red	0.63 – 0.69	
near-infrared	0.76 – 0.90	0.61
PAN	0.45 – 0.90	

A field campaign was conducted in the study site on the same day of the acquired image to collect information about phenological status of the different land-covers, canopy height and type of planting (partial or full cover). The identified covers in the study area were: bare soil, water bodies, buildings and urban constructions, forest (including shrubs and other natural vegetated areas), and two different types of agricultural plots, crops and orchards. Crops correspond to different types of vegetables and pastures (e.g., alfalfa, maize). Orchards consisted of diverse fruit trees such as nectarines, table grapes, and walnuts, among others. This information was recorded in a GIS vector file and used as ground-truth data for validating the land-covers of interest and the classification results. The polygons included in the vector file are overlapped to the original image in Figure 5.2. The complete distribution of the available ground-truth examples for classification is shown in Table 5.2.

Fig. 5.2.: Real color composition (red, green and blue bands) of the WorldView-2 image. The polygons corresponding to the interest classes are overlapped on the image.

Tab. 5.2.: Class distribution of the available ground-truth examples. The percentage represents the area that covers a class regarding the total area of the ground-truth.

Class	Area (ha)	Percentage (%)
Bare soil	6.51	2.82
Urban constructions	26.70	11.55
Forest	25.20	10.90
Orchards	123.57	53.43
Crops	44.84	19.39
Water bodies	4.43	1.92

5.2.2 Methodology

In order to automatically analyze VHR images of fragmented agricultural landscapes our methodology combines superpixel processing and edge-based methods, which provide the basis to identify interest objects as agricultural plots and generate land-cover maps. A superpixel method produces a controlled over-segmentation of the image, reducing the intra-variability of the image pixels as well as the size of the analysis space, which allows performing an analysis similar to pixel-based ones. In addition, an edge-based method detects abrupt changes in pixel intensity which usually characterize the object boundaries in a scene. In the case of the images under study, edge-based methods can provide information about the edges corresponding to roads or the limits of agricultural plots, whereas regions within these edges can be obtained by means of low-level image processing operations (e.g., erosion, dilation). These regions are used as a reference for the superpixel method, proposed in Section 3, with the premise of the generated superpixels respect boundaries established *a priori*.

At superpixel level, a classification step is carried out using the RUSBoost algorithm (Section 2.5.2) as a classifier, whereas a simple rule-set (Section 5.2.3) is used at edge-based region level.

5.2.3 Image segmentation scheme

A general scheme of the segmentation process is shown in Figure 5.3. This segmentation scheme provides the following working scales: (i) superpixels (finer scale) provide spectral information so they are processed similarly to image pixels, and (ii) the edge-based regions (coarser scale) contain edge information in the better case related to agricultural plots boundaries, so they are treated as candidate agricultural plots.

Fig. 5.3.: Segmentation process of a multi-spectral image. Edge-based segmentation is applied to obtain an initial segmented image. These segments are later used as reference contour to generate superpixels.

Edge-based segmentation

Physical edges provide important visual information since they correspond to discontinuities in the physical properties of scene objects (Ziou and Tabbone, 1998). These edges are related to significant variations in the reflectance, illumination, orientation, and depth of scene surfaces. Starting from these edges, regions can be obtained by means of morphological operations such as flood-fill and dilation among others.

The procedure to obtain the regions is described in Algorithm 5.1. Edges are extracted for each band b in multi-spectral image (MS) using Canny edge detector (Canny, 1986) ($edge$). Then each edge is dilated ($\oplus$), using lines of 3 pixels length in angles of 0 and 90 degrees as structuring element objects ($strel$). Only the intersection of all edges is retained in E , after which all regions and holes of its complement $\tilde{E}$ are filled using an algorithm based on morphological reconstruction (Soille, 1999). The obtained regions R are dilated, and finally the connected components are labeled using a 4-neighbor connectivity, and resulting in segments S .

Algo. 5.1: Edge-based segmentation algorithm.

Input: A multi-spectral image MS .A morphological structure element $strel$ containing lines in different angles.**Output:** A segmented image S . $E \leftarrow \emptyset$ **for each** band $b \in MS$ **do** $e \leftarrow edge(b)$ $E \leftarrow E + (e \oplus strel)$ **end for** $\tilde{E} \leftarrow not(max(E))$ $R \leftarrow fill(\tilde{E})$ $R \leftarrow dilation(R)$ $S \leftarrow \text{connected-component-labeling } R$ **return** S

Analysis scheme

The OBIA methodology allows exploiting spatial, shapes, and semantic relations among regions together with their spectral characteristics. As in the works of Gianinetto et al. (2014), Petrou et al. (2012), and Jiao and Liu (2012) the features used for classification are the following:

Spectral. Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1974), Normalized Difference Water Index (NDWI) and Spectral Shape Index (SSI) (Chen et al., 2009), as defined in Equations 5.1–5.3.

$$NDVI = \frac{NIR - R}{NIR + R} \quad (5.1)$$

$$NDWI = \frac{G - NIR}{G + NIR} \quad (5.2)$$

$$SSI = |R + B + 2 \times G| \quad (5.3)$$

where R, B, G and NIR represent the red, blue, green and near-infrared bands, respectively.

Texture. Local entropy of an image (Gonzalez and Woods, 2007) computed in a neighborhood of 9x9 pixels, and calculated as in Equation 5.4.

$$E(z) = - \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i) \quad (5.4)$$

where z_i is a random variable indicating intensity, $p(z)$ is the histogram of the intensity levels in a region (usually set to a size of 9), L is the number of possible intensity levels.

Shape. Fractal dimension index (FRAC) (McGarigal et al., 2002), as in Equation 5.5.

$$FRAC = \frac{2 \cdot \ln(0.25 \cdot perimeter)}{\ln(area)} \quad (5.5)$$

where values greater than 1 indicate an increase in shape complexity. *FRAC* approximates 1 for shapes with very simple perimeters such as squares, and approximates 2 for shapes with highly convoluted, plane-filling perimeters.

Since two working scales (superpixels and edge-based) are available, different features are used at each analysis scale. Spectral and texture features are extracted from superpixels, whereas shape and texture features are used to characterize edge-based segments. The classification workflow is shown in Figure 5.4.

Fig. 5.4.: Workflow of the analysis process. Two classifications are carried out over superpixels and edge-based regions. The first one considers spectral and texture features. The second one employs shape and texture features, as well as the information provided by the previous classification to generate a final thematic map.

Each superpixel is characterized by the mean and standard deviation of the NDVI, NDWI, and SSI indexes, all the available spectral band values, and the local entropy; and classified according to the following classes: water bodies, buildings and urban constructions, bare soil, and vegetated areas, i.e., forest, crops, and orchards. Because the PAN band contains more information about texture, local entropy is calculated from it. Then a dataset is created using characterized superpixels as observations. These are labeled according to an overlap criteria: if at least 80% of a superpixel surface coincides with the surface of a ground-truth region, then the corresponding observation acquires the label of this region. Due to superpixels are similar in size and well adapted to image regions, the labeled subset follows the same distribution as the ground-truth dataset (see Table 5.2). This new dataset is used for classification. Since this dataset is imbalanced, a classifier that considers the distribution of the class is needed in order

to obtain satisfactory results. For this reason, the RUSBoost algorithm (Seiffert et al., 2010), described in Section 2.5.2, is used as classifier.

A cascade multi-staging classification of superpixels is conducted. In each stage, a classification is performed using RUSBoost classifier over only two classes, which are defined as the dichotomous relationship between an interest class and the rest of them. For example, when the class under study corresponds to water bodies, the other class includes everything (i.e., bare soil, urban constructions, and vegetated areas) that has not been previously classified. This classification scheme is inspired by land-cover taxonomies (e.g., CORINE Land-cover scheme), where dichotomous categorizations are followed by experts to generate standardized thematic maps.

At edge-based segments scale, fractal dimension index (Equation 5.5) is used to distinguish which segments could be plots from natural entities as forests in the whole segmented image. Therefore regions where $FRAC$ approximates 1 are considered as candidate plots, whereas segments with higher values are considered as forests. Because of their planting frame (the way how trees are planted), orchards present a high variability, whereas crops are more homogeneous respect to its gray values. To classify which plots are potentially crops or orchards, local entropy is calculated from the PAN band and each segment is then characterized and analyzed. This analysis is performed using a rule-based classification (Section 5.2.3).

From the analysis of superpixels and edge-based segments, maps at superpixel and plot scales are respectively generated. The former one is obtained from data-driven methods using ground-truth land-cover data and information provided by the edge-based regions of the agricultural plots. Superpixel map provides information about different interest classes and not about plots (interest objects) in the image. On the other hand, the plot map provides information about plot candidate segments using the rules based on shape and texture attributes. However, they do not supply information about land-cover classes on images.

An integration process is performed to merge both the information superpixels and plot maps. The composition is made analyzing each segment at a plot scale, by means of the classified superpixels which form part of it. A plot segment is labeled according to the vegetated region, on a superpixel map that covers most of its surface following the next condition: if the majority label agrees with the classification on the plot map, then the entire plot segment keeps its label. But if a discrepancy arises between both classifications, the segment is marked as incongruent. Two types of incongruent segments are considered:

Type 1 (T_1). These are vegetated regions that according to their shape were considered as non-plots. However, they are classified as orchards or crops at superpixel level.

Type 2 (T_2). These are segments whose classes in both maps are interchanged, e.g., on the plot map, a segment is identified as an orchard. However, the overlapping region is classified as crops on the superpixel map.

In order to provide an automatic solution to incongruent segments, two criteria are established to substitute T_1 and T_2 . In the case of T_1 segments, they are replaced by the overlapping areas on the superpixel map, due to the possibility that at a superpixel scale the region could be divided by different land-covers into smaller parts with a more regular shape. On the other hand, T_2 segments are labeled according to the majority vegetated class (orchards or crops) covering its surface at superpixel level. The information of the superpixel map, instead of the given by the plot map, is primarily used for labeling a T_2 segment because the rule used to labeled it could be far too restrictive for some plots in an incipient vegetative stage.

Rule-based classification

Information about regularity of the shape is used as a feature to distinguish which segments could be plots from natural entities as forests in the whole segmented image. Following the definition of fractal dimension index (Equation 5.5), regions with a shape value around 1 are considered regular whereas higher values are complex forms. Thus a simple rule based on segment shape is defined to distinguish between agricultural plots (orchards and crops) and non-plots:

$$segment = \begin{cases} \text{plot} & \text{if } FRAC > th_{frac1} \text{ and } FRAC < th_{frac2} \\ \text{non-plot} & \text{if } FRAC \geq th_{frac2} \end{cases} \quad (5.6)$$

Once candidate plots are obtained, they are analyzed by their textural characteristics. Because of the way the trees are planted (planting frame), orchards have high texture compared to crops. Therefore, a rule to differentiate these types of agricultural plots was formulated as follows:

$$plot = \begin{cases} \text{crop} & \text{if } entropy \leq th_{entropy} \\ \text{orchard} & \text{if } entropy > th_{entropy}, \end{cases} \quad (5.7)$$

5.3 Results and Discussion

From multi-spectral bands, two segmented images were obtained by using edge-based and SLIC methods. As shown in Figure 5.5a, edge-based segments are well suited to most of the small agricultural plots. From these segments, using a compactness factor of 80, a total of 297257 superpixels covering the whole image were automatically generated, where each superpixel is composed of 25 pixels on average. They represent 4% of observations to analyze respect to the entire number of image pixels under analysis. The inclusion of edge-based segments as *a priori* information for SLIC algorithm allowed segmentation inside of specific areas, which is useful for intra-plot analysis. As can be observed, superpixels respect the limits of the edge-based segments and adhere well to the boundaries of spectrally homogeneous regions.

Due to visualization issues, only a zooming-in image region containing a set of generated superpixels is shown in Figure 5.5b.

Fig. 5.5.: The edge-based segmentation (5.5a) provides regions whose edges (red lines) mark the location of discontinuities in color as the roads. The result of SLIC (5.5b) shows regions of slightly different shapes according to the spectral properties of the image, but following and preserving the edges previously obtained. Figure 5.5b corresponds to the area enclosed by the black dotted square marked in Fig. 5.5a.

During the superpixel multistage classification, at each stage, a total of 1000 decision trees were used as weak learners (h_t) to build a single RUSBoost classifier using a ratio of sampling r of 50:50. To validate classification results a 3-fold cross validation was performed by an image object assessment. That means that the statistics are calculated over the given classes to superpixels (image objects) corresponding to the test set in each fold. The obtained results showed mean accuracies (user and producer) greater than 88% with a small deviation standard (1.36 on average) for all classes except for the bare soil class, where the user accuracy is 58.37%. In addition, the results point out a commission error of 63.40% for the same class. This may occur because entities identified by experts as agricultural plots and forests are not completely pure, i.e., they can contain heterogeneities such as patches of bare soil. In spite of that fact, the obtained results can be considered competitive and stable, since they are high (averages of 88% and 91% for user and producer accuracies, respectively) and present a low variability (an average lower than 1.6 in both cases). The assessment of multi-staging classification at superpixel level is displayed in Table 5.3.

After the classification step, a superpixel map is generated integrating the superpixels classified by the classifier with the best accuracy score. Figure 5.6 shows the result of the classification at superpixel level. As can be observed, even though the generated superpixel map does not present adverse outcomes as a *salt-and-pepper* effect, not all agricultural plots are represented as homogeneous areas, since some of them mostly contain small bare soil regions. Likewise, the same occurs with bare soil regions where small vegetated regions are presented. This agrees

Tab. 5.3.: Accuracy assessment of the classification at superpixel level.

Classes	User Acc. (%)	Prod. Acc. (%)	Comm. (%)	Omm. (%)
Water bodies	96.56 ± 3.11	89.73 ± 1.11	3.20 ± 2.93	10.21 ± 1.11
Bare soil	58.37 ± 1.57	88.76 ± 2.22	63.40 ± 11.06	11.06 ± 2.22
Urban constructions	95.22 ± 0.72	90.70 ± 1.30	4.55 ± 0.76	9.22 ± 1.30
Forest	88.75 ± 1.11	89.75 ± 0.66	11.36 ± 1.25	10.23 ± 0.66
Crops	91.13 ± 1.39	90.20 ± 1.62	1.67 ± 0.27	1.4 ± 0.28
Orchards	98.32 ± 0.26	98.49 ± 0.28	8.79 ± 9.70	9.70 ± 1.62

with reality since plot homogeneity depends on various factors, such as management practices, soil properties, among others. Moreover, adjacent orchards separated only by dirt roads are joined together, making it impossible to distinguish one plot from another. In Table 5.4, the confusion matrix of the classification map is presented. Therein can be observed the heterogeneity phenomena among vegetated areas and bare soil, but also among forests, crops and orchards. To validate current and subsequently generated maps, an accuracy assessment based on random points was performed. An overall accuracy of 91.74% and a Cohen's kappa coefficient (κ) of 0.8728 are obtained by comparing the ground-truth data and the superpixel map. The above implies that information given to the classifier is not enough to completely separate the classes, and that more information needs to be added. For this purpose, edge-based segments are incorporated to the process.

Tab. 5.4.: Confusion matrix showing results of the classification at superpixel level. The values are expressed on percentages of overlapping area between superpixels and ground-truth regions.

		Classification					
		Water	Bare soil	Urban	Forest	Orchards	Crops
Ground-truth	Water	94.83%	0.72%	0.23%	3.61%	0.53%	0.09%
	Bare soil	0.00%	87.72%	0.70%	3.09%	3.16%	5.33%
	Urban	0.00%	1.88%	87.39%	4.70%	5.35%	0.69%
	Forest	0.00%	5.90%	8.52%	79.07%	6.23%	0.29%
	Orchards	0.01%	0.28%	0.17%	2.45%	95.97%	1.13%
	Crops	0.00%	1.74%	0.71%	0.04%	7.42%	90.09%

Edge-based segments provide a working scale closely related to the boundaries of the agricultural plots (interest objects). Therefore, they are used to generate additional information in order to improve the results of vegetated areas (i.e., forests, crops and orchards) at superpixel scale. It should be remarked that obtaining these segments only depends on low-level image processing operations (e.g., erosion, flood filling, dilation), so regions are generated without the necessity of fine-tuning any parameter. Inasmuch as edge-based regions are used as candidate plots, high level descriptions of them are used to determine which segments are more likely to be orchards or crops.

These segments were classified using the rules defined in Section 5.2.3. For Equation 5.6, thresholds were empirically selected to distinguish regular segments. In this work, a $FRAC$ value in a range of 1 ± 0.1 are considered as candidate parcels. Therefore, th_{frac1} and th_{frac2} were set to 0.9 and 1.1, respectively. To discriminate orchards and crops $th_{entropy}$ is set to 4, a value obtained through sampling of the texture image. During this analysis, only segments that

contain at least a surface covered 60% by vegetation (on the superpixel map) are considered to be analyzed.

The result of applying the rules at edge-based segments level is shown in Figure 5.7. The obtained map, called parcel map, shows vegetated areas likely to be parcels separated from others areas trough the use of a shape feature. Crops and orchards are differentiated using the texture characteristics of the candidate plots.

Fig. 5.6.: Superpixel map generated through the classification of superpixels according to the interest classes.

The integration of information of both superpixel and plot maps is displayed in Figure 5.8. The regions of T_1 and T_2 are represented in two gray-scale colors. As can be observed, identified parcels on the resulting map represent only one land-cover, as opposed to superpixel maps where the regions inside plots could contain more than one type of land-cover. This behavior is similar to the analysis done by human operators, where heterogeneities are often not considered when making the delimitation of land-covers. In general, T_1 regions mainly occur on segments that were incorrectly delimited at a parcel scale. Therefore, an improvement in the delimitation of the object is probably required. This is potentially useful for T_2 regions, which mainly arise because it is difficult to discriminate between classes in the parcels when vegetation is at an early phenological stage. In spite of that fact, information about the composition of a specific agricultural plot is useful for intra-plot analysis, so this information can be accessed through the superpixels that form part of it. Moreover, superpixels entirely provide the information

Fig. 5.7.: Map of the vegetated areas. Two types of agricultural plots are distinguished among them and from other vegetated areas using shape and texture characteristics.

about the non-vegetated areas (*e.g.*, water, bare soil), whereas incongruent segments (T_1 and T_2) give information about regions that require a disambiguation by an expert.

The result of applying the criteria to substitute T_1 and T_2 segments by a vegetative ground-truth class is displayed in Figure 5.9. The final map has a κ value of 0.8947 and an overall accuracy of 93.25%. These results show an improvement in classification compared to the obtained in superpixel maps. Moreover, in the confusion matrix (Table 5.5), a refinement of vegetative classes is observed, especially in the forests (shrubs and natural vegetated areas) class where an increment of accuracy, from 79.04% to 86.61%, is reached. This occurs because giving a unique label for entire parcel area reduces the intra-variability of the map. Therefore, some heterogeneities (*e.g.*, bare soil regions) are eliminated in better agreement with the ground-truth data in natural vegetated areas. Conversely, a slight decrease in accuracy, respect to the superpixel map, can be seen in the bare soil (-4.03%) and urban classes (-0.04%). This phenomenon mainly occurs at the edges of both classes that are adjacent to a vegetated area, which implies that there is a slight mismatch between some segments and the objects defined in the ground-truth data.

Fig. 5.8.: Combination of superpixel and parcel maps. In two different gray-levels, incongruent regions between parcel and superpixel maps are shown.

Fig. 5.9.: Final map generated with the disambiguation of T_1 and T_2 regions.

Tab. 5.5.: Confusion matrix showing results of the classification of the final map. The values are expressed on percentages of overlapping area between ground-truth regions and the final map.

		Classification					
		Water	Bare soil	Urban	Forest	Orchards	Crops
Ground-truth	Water	94.83%	0.72%	0.23%	3.61%	0.53%	0.09%
	Bare soil	0.00%	83.69%	0.00%	4.35%	4.45%	7.51%
	Urban	0.00%	1.88%	87.35%	4.64%	5.40%	0.74%
	Forest	0.00%	0.93%	4.35%	86.61%	7.93%	0.18%
	Orchards	0.00%	0.12%	0.17%	0.86%	96.15%	2.70%
	Crops	0.00%	0.01%	0.01%	0.02%	6.63%	93.32%

5.4 Conclusions

In this Chapter, a methodology based on superpixels for landscapes dominated by small agricultural plots was presented. The methodology incorporated two scales of analysis: superpixels and edge-based scale. They were made using the improved version of the SLIC algorithm (proposed in Section 3) and an edge-based procedure, for generating superpixels and segments closer to plots, respectively.

The analysis of both scales allowed the acquisition of different types of information about the landscape. On the one hand, superpixel analysis provided information about the scene at a scale more related to image pixels, but with the advantage of reducing the analysis space and image variability, and consequently, the processing time (supervised classification step). Moreover, the use of image edges as *a priori* information for the generation of superpixels allowed to obtain regions with well-defined borders. On the other hand, the edge-based segmentation method provided segments closer to real plots with no need of a time-demanding tuning process, since they are automatically generated. Even if some agricultural plots were not perfectly delineated, edge-based segments were useful for the analysis of forms. This allowed to determine which regions required a more detailed analysis and which ones are candidates for being plots. Both methods were easily automated, since only a few parameters were required, i.e., only two in the case of the superpixel method, which was useful for the automation of the image analysis.

The combination of the information of both scales provided an automatic way to obtain a thematic map with characteristics of superpixels and candidate plots. This map presented homogeneous appearance in vegetated areas, whereas heterogeneity was maintained through superpixels for non-vegetated areas. The methodology was tested using a scene with agricultural fragmented landscape. The use of two complementaries methods (SLIC and edge-based) achieved a classification accuracy better than when those methods were performed separately.

The results of this Chapter have been presented sent it to a JCR journal, after a peer-review process it was accepted for publication (see Section 8.2).

Delineating Agricultural Plots in VHR Remote Sensing Images

” *Fronteras que rigen los sumos lugares, fronteras tangibles y siempre intocables, lo mismo perpetuas que provisionales. Me envuelven fronteras por todas mis partes.*

— **Silvio Rodríguez Domínguez (1946–)**
Cuban musician.

6.1 Introduction

Accurate and up-to-date information about status, acreage and the type of agricultural lands is assumed as a valuable element for diverse agricultural-related agencies and diverse agricultural applications. This information allows stakeholders among other things to establish agricultural policies (Mirón Pérez, 2005; van Der Molen, 2002) to reduce greenhouse gas emissions, regulate water rights, and estimate subsidies; and to improve agricultural management through the use of precision agriculture techniques. In this regard, it is important to consider that about 75% of the world’s agricultural lands are small (less than 2 ha) and family-operated (Lowder et al., 2016). This implies highly fragmented agricultural landscapes with a high spatial heterogeneity produced by the diversity in sizes, shapes, and crops of the different agricultural plots. Therefore, in order to generate precise information about these agricultural lands a primary requirement is to have a correct delineation at plot level.

The delineation of agricultural lands has been addressed with different initiatives around the world for a long time. In 1980, the National Research Council (NRC) published the report: *Need for a Multipurpose Cadastre* (NRC, 1980), which has been updated by the study: *National Land Parcel Data: A Vision for the Future*, that examines the status of land parcel data in the United States and provides a set of recommendations that would foster a national system for land parcel (NRC, 2007). In the European scenario, it can be mentioned the Land Parcel Identification Systems (LPIs) promoted by the European Union in order to represent the activities of farmers on their lands (Leo and Lemoine, 2001). These initiatives commonly use very-high resolution remotely sensed imagery to perform a manual delineation of agricultural plot boundaries. However a non-trivial issue is how to process a huge data volume maintaining the accuracy and time requirements. Even though the manual delineation can be very precise,

it suffers from the subjectivity of operator and is highly time consuming. Moreover, the repeatability of the delineation is not insured even when the same operator performs it at two different times.

To address these problems, automatic and semiautomatic methods have been proposed in the remote sensing literature. Most of these methods are based on image segmentation. Mueller et al. (2004) proposed an object-based approach for extracting large, human-made objects, especially agricultural fields, from high resolution imagery. This approach combined edge detection models with region-based segmentation to extract regularly shaped objects. In the work of Da Costa et al. (2007), an algorithm to automatically delineate vine plots from very-high resolution images based on their textural properties was developed. From texture attributes, they applied a thresholding method to discriminate between vine and non-vine pixels. Tiwari et al. (2009) proposed a semi-automatic methodology for extracting field boundaries from satellite data.

A segmentation using tonal and textural gradients was performed and the generated regions were classified to derive preliminary field boundaries. Finally, snakes algorithm (Kass et al., 1988) was used to refine the geometry of these field boundaries. Turker and Kok (2013) used perceptual grouping for automatic extraction of dynamic sub-boundaries within existing agricultural fields from remote sensing imagery. To perform field-based analysis, the approach integrated field boundary data and satellite imagery. Canny edge detector was used to detect the edge pixels.

In general, approaches based on segmentation methods have the following drawbacks: (i) they are sensitive to intra-plot variability which can produce more segments than desired; (ii) most of these methods are highly dependent on a correct parameter selection (e.g. the similarity measured used to group image pixels) that requires a prior knowledge about the scene or tuning by trial-error. Moreover variability in sizes and shapes of the plots causes that a certain configuration parameters do not allow properly delineate all plots. Besides, approaches based on edge detection are used for delineating agricultural plots, however they tend to produce more of the desired edges, mainly due to plowed fields and image noise.

In this Chapter, two different methodologies based on superpixels for delineating agricultural plots are presented. The first methodology uses superpixels as minimum processing units, whereas a process of agglomeration of superpixels is used to obtain a final segmentation where the agricultural plots (objects of interest) are distinguished. In this regard, the determination if two adjacent superpixels should be merged is take it by a classification method, which is trained using part of a segmented scene under study. The second approach combines different segmentations at different scales performed by a superpixel method to obtain a single segmentation of the agricultural plots. Segmentations reduce noise in the image, as well as the dependence of the parameters selection, producing a better delimitation of the plots.

6.2 Materials and Methods

6.2.1 Data Set

The data set used in this study corresponds to the WorldView-2 multi-spectral image described in Section 5.2.1. Two regions containing mostly agricultural plots were selected and clipped from the WV-2 scene (as can be seen in Figure 6.1). Each clipping image has been fused to improve their spatial resolution using the well-known intensity–hue–saturation (IHS) algorithm (Carper, 1990), resulting in images with a size of 2088x2088 pixels. Agricultural plots in both images have been manually delineated obtaining two reference plot maps (Figures 6.2a and 6.2b).

Fig. 6.1.: Two regions of a WorldView-2 image are used in this study. The red squares show the location of both regions.

6.2.2 Approach based on Classification

To delineate agricultural plots, this methodology combines superpixel processing and a classification method, which provide the basis to decide when two adjacent superpixels should be merged. An overall overview of the proposed methodology is shown in Figure 6.3.

The methodology starts with an over-segmentation of the image obtained through the superpixel algorithm, described in Section 3. Then, from the generated superpixel representation, a dataset is created by extracting segment features of each pair of superpixels. Instances in dataset are

Fig. 6.2.: Study regions are shown in a real color composition. Borders of each agricultural plot are displayed in red.

Fig. 6.3.: Overview of the methodology based on machine learning.

labeled in two classes depending on if they belong to the same object (agricultural plot) or not, this information is given by previously generated reference plot map. Finally, dataset is used to generate a machine learning model, through training a classification method. This model is later used to determine from dataset features whether two superpixels should be merged, allowing to delineate agricultural plots.

For creating the dataset used for training, all available spectral bands (i.e., four bands, from blue to near-infrared) as well as three spectral indexes commonly used in remote sensing image analysis are used for feature extraction. The spectral indexes used in this study are: Normalized Difference Vegetation Index (*NDVI*), Normalized Difference Water Index (*NDWI*), and Spectral Shape Index (*SSI*), which were previously defined in Equations 5.1–5.3. In addition,

from the above features, a new set of texture-based features are computed calculating their local entropy (Equation 2.12), varying the size of the neighborhood in which entropy is measured. Thus, each superpixel is characterized by a feature vector, which components are defined by the average value of the pixel feature values that are part of it. Finally, the dataset F , used in the classification process, is created by characterizing each pair of superpixels (i and j) as follows:

$$F_{ij} = |f_i - f_j| \forall i \neq j \quad (6.1)$$

where f_i is the feature vector of the i^{th} superpixel, and $|\cdot|$ represents the absolute value. The main idea is to create a multidimensional feature space, in which a machine learning (classifier) method learns a function able to predict when a pair of superpixels belong to the same object. In this work, an absolute difference is used, however pairs of superpixels can be characterized in diverse ways (e.g., by applying different distance measures, or by concatenating both feature vectors), in this regard, the feature vector representing the relationship between those superpixels must be a vector with some properties such as symmetry (i.e., $F_{ij} = F_{ji}$), and it cannot be negative.

Since the aim of this work is to agglomerate superpixels, only those that are adjacent are considered. A label set $\mathbf{L}$ (target labels) is created using the information about the adjacency of superpixels, and the available reference plot map (ground-truth data). Thus, each instance of dataset $\mathbf{F}$ is labelled according to the following Equation:

$$L_{ij} = \begin{cases} +1, & \text{if } i \text{ and } j \text{ belong to the same plot and } i \neq j; \\ -1, & \text{otherwise.} \end{cases} \quad (6.2)$$

where i corresponds to the i^{th} superpixel, and L_{ij} represents the label of the instance F_{ij} . Here a positive label for L_{ij} indicates that superpixels i and j should be merged, whereas a negative one means that both superpixels belong to different objects, hence they should not be merged.

Due to the characteristics of the landscape under analysis, target labels are imbalanced (i.e., there are more positive labels than negative, or in other words more pairs of superpixels belong to the same object). Therefore, a classifier that considers this distribution of the classes is needed in order to obtain satisfactory results. For this reason, the RUSBoost algorithm (described in Section 2.5.2) was used as classifier.

6.2.3 Approach based on Consensus of Segmentations

This methodology is based on the generation of various segmentations, from which combination is possible to delineate agricultural plots. An overall overview of the proposed methodology is shown in Figure 6.4.

Fig. 6.4.: Overview of the methodology based on consensus of segmentations.

To facilitate plot delineation, the image is filtered reducing noise. The filter chosen for this purpose is the Rolling Guidance Filter, proposed by Zhang et al. (2014), which is based on the bilateral filter.

In this work, image segmentation is carried out by means of the SLIC algorithm (Section 3). Segmentation are generated fixing the value of the compactness parameter, but varying the size of the superpixels. Similar to Section 4, the number of superpixels is selected to follow a dyadic scale.

The consensus process consists of generating l segmentations, and then combine them through a voting scheme that determines which pixels tend to belong to the same region and which are part of the edges of objects in the image. Thus, for each pair of adjacent pixels i and j , the voting scheme E_{ij} takes into account if both belong to the same region. Then E_{ij} is defined as:

$$E_{ij} = \sum_{k=1}^l \Psi_{ij}^k \quad (6.3)$$

$$\Psi_{ij}^k = \begin{cases} S(r_a^k, r_b^k), & \text{if } i \in r_a^k, j \in r_b^k \text{ and } r_a^k \neq r_b^k; \\ 0, & \text{otherwise.} \end{cases} \quad (6.4)$$

where r_a^k represents the region a belonging to the k -th segmentation, and $S(r_a^k, r_b^k)$ is measure of the similarity between regions a and b . If i and j do not belong to the same region, they probably are part of the edges between two objects. The larger the value of E_{ij} , the stronger the separation between the objects. Finally, an edge map is generated with the information of voting scheme, which is normalized to a range between 0 and 1.

The similarity measure, analogous to the proposed by Uijlings et al. (2013), is defined as a combination between similarity measures of color and texture:

$$S(r_a^k, r_b^k) = S_{color}(r_a^k, r_b^k) + S_{texture}(r_a^k, r_b^k) \quad (6.5)$$

$S_{color}(r_a^k, r_b^k)$ measures the color similarity. For each region r_a^k , a color histogram is obtained using 25 bins for each spectral band. Then from color histograms, a feature vector $C_a = \{c_a^1, \dots, c_a^h\}$ with a length of $h = 25 \times B$ is generated for region a where B is the number of bands of the image. Feature vector is normalized using the norm L_1 . Finally, the similarity in color between two regions a and b is calculated using the χ^2 statistic (Rubner et al., 2000):

$$S_{color}(r_a^k, r_b^k) = \sum_{i=1}^h \frac{(c_a^i - (c_a^i + c_b^i)/2)^2}{(c_a^i + c_b^i)/2} \quad (6.6)$$

$S_{texture}(r_a^k, r_b^k)$ measures the texture similarity. Texture is calculated by means of steerable filters (Freeman and Adelson, 1991) using as basis Gaussian derivatives in eight directions and $\sigma = 1$. In this case, for each region r_a^k , a texture histogram is obtained using 10 bins for each spectral band and each filter direction. Then from texture histograms, a feature vector $T_a = \{t_a^1, \dots, t_a^h\}$ with a length $h = 80 \times B$ is generated for each region a . The similarity in texture between two regions a and b is calculated as follows:

$$S_{texture}(r_a^k, r_b^k) = \sum_{i=1}^h \frac{(t_a^i - (t_a^i + t_b^i)/2)^2}{(t_a^i + t_b^i)/2} \quad (6.7)$$

From the map of borders and using the method proposed by Arbelaez et al. (2009), an Ultrametric Contour Map (UCM) is calculated. An UCM is an edge map with the remarkable property that produces a set of closed curves when any threshold is set (Dollár and Zitnick, 2013). This does not necessarily occur with normal edge maps. In other words, a UCM is a soft representation of a segmentation that takes into account information from the edges of the image. Thus, the larger the threshold the greater the contrast of the edges of the generated segments.

6.3 Results and Discussion

6.3.1 Approach based on Classification

To prove the potential of the proposed approach based on a classification method, two experiments have been carried out. The first one uses the image A to generate the classification model which is subsequently tested on image B. The second experiment is similar to previous one but interchanging the images used to generate and test the model.

From each image, a total of 5450 superpixels were automatically generated through modified SLIC method (i.e., extended to multispectral images), where each superpixel is composed of 50 pixels on average. The number of pixels that constitute each superpixels was chosen experimentally to agree with the size of most of the treetops present in the images under analysis. They represent 2% of observations to analyse respect to the entire number of pixels under analysis per image. Due to visualisation issues, only enlarged regions containing a set of generated superpixels are shown in Figure 6.5. As can be observed, superpixels adhere well to the boundaries of spectrally homogeneous regions, in particular, borders of plots are well delineated.

Fig. 6.5.: Superpixel segmentation of two small areas. Borders of each superpixel are shown in red.

The number of instances generated for images A and B are 12,691 (77.44% positive and 22.56% negative) and 12,994 (78.75% positive and 21.24% negative), respectively. Each instance were characterized by a vector of 28 features: 7 corresponding to the four bands and the spectral indexes, and the remaining 21 to the local entropy computed on neighbourhoods of 3 different sizes (9, 17, 33) over the 7 first features. These instances were used to create two datasets, one for each image. During the classification step a total of 1000 decision trees were used as weak classifiers to build a single RUSBoost classifier using a ratio of sampling of 50:50.

To evaluate the results obtained by the classifier, a 10-fold cross validation was performed separately using both datasets. During each fold, non-overlapping testing and validation sets (which correspond to seen data) were generated by randomly selecting positive and negative instances maintaining their original class distribution. The obtained results show mean values of 94.01% and 92.85% for user and producer accuracies, respectively for image A and 89.58% and 92.97% for image B in the case of positive instances (merge). In the case of negative instances (do not merge), mean values of 75.13% and 78.63% for user and producer accuracies were obtained for image A and 73.66% and 76.57% for image B. This may occur because there are few negative (do not merge) instances for training the classifier, requiring more

negative instances to improve these results. The assessment of classification using the test set is displayed in Table 6.1. The low variability of the accuracies in 10 folds points out that the method is stable, therefore similar results are expected during the validation process.

Tab. 6.1.: Accuracy assessment of the classification in the validation set.

Image	Classes	User Accuracy (%)	Producer Accuracy (%)
A	+ (merge)	94.01 ± 1.09	92.85 ± 0.84
	- (do not merge)	75.13 ± 3.32	78.63 ± 2.85
B	+ (merge)	89.58 ± 0.87	92.97 ± 0.71
	- (do not merge)	73.66 ± 2.95	76.57 ± 2.99

To test the approach based on machine learning, the classifier of the fold with the best overall accuracy in the validation set was used to determine which superpixels should be merged in the test image. Thus, overall accuracies of 83.53% and 85.58% were obtained using as input the test images (unseen data) A and B, respectively. The results of applying these models to validation and test data are shown in Figure 6.6, as expected the results in validation set fit better with the ground-truth than the results obtained in test set. However in both cases, most of the superpixels were correctly merged, indicating that better results can be obtained by improving the methodology.

6.3.2 Approach based on Consensus

On the other hand, for testing the methodology based on the consensus of diverse segmentations, the following experiment was carried out. From the preprocessed image, SLIC algorithm is used to generate 17 different segmentations with a compactness factor of 0.04 times the maximum spectral value contained in the image. On the other hand, the number of superpixels goes from 4 to 272,484 pixels on average, following a dyadic scale. Examples of obtained segmentation, amplified for a better visualization, are shown in Figure 6.7. As shown, also in this case, the superpixels tend to adhere well to image-objects. Particularly, the smaller superpixel the better its adherence to the edges, as well as the homogeneity of the pixels that compose it (see Figure 6.7a). Moreover, in the segmentations with larger superpixels (Figure 6.7b), the segments are less homogeneous spectrally; however, they still respect the borders between regions with a high contrast.

The consensus of the different segmentations results in the edge map shown in Figure 6.8. This map combines different information such as: (i) different scales of segmentation through the different sizes of superpixels, (ii) the dissimilarity both in texture and color between neighboring regions, and (iii) the probability of a pixel to belong to an edge. The closer to 1 the value of a pixel, the greater the probability of it to belong to an edge in the multiple segmentations, and therefore, the contrast between regions separated by it in different scales is greater.

Fig. 6.6.: Results of applying the approach based on machine learning in the images under study. The edges, colored in black, have been thickened by morphological operations for better visualization.

The obtained edge map (Figure 6.8) provides useful information to identify the edges belonging to the plots; however, as mentioned above, it does not guarantee complete plots by applying a threshold, mainly due to non all edge-pixels have the same value. Thus, to obtain the plots was required the use of algorithms to complete edges. Therefore, a UCM is generated from the edge map. Unlike the edge map, the UCM method returns sharp curves, ensuring complete plots by applying the appropriate threshold. Figure 6.9 shows the edge map (Figure 6.9a) and the UCM (Figure 6.9b) of an enlarged region of the image.

(a) Superpixel size = 17,233 pixels on average (b) Superpixel size = 439 pixels on average

Fig. 6.7.: Superpixel segmentation of two small areas with different superpixel sizes. Borders of each superpixel are shown in red.

Fig. 6.8.: Consensus of segmentations. Higher values represent a high agreement between segmentations.

Different thresholds were tested to determine which allows better delineation of the agricultural plots. In this regard, the higher is the threshold (greater than 0.70) greater the registered error due mainly to outgrow few edges producing a sub-segmentation of the image. On the contrary, the less strict is the threshold, the error decreases while an increase in the number of segments occurs. In the framework of proven thresholds, the best result was obtained with 0.45.

Figure 6.10 shows a comparison between the ground-truth and the edges of the segments obtained with the threshold 0.45. As can be seen, most of the edges agree in both images, particularly in homogeneous areas with a high contrast between adjacent regions; however, it can be also distinguished discrepancies in plots with different tilled patterns, where the

Fig. 6.9.: Comparison between the edge map and the Ultrametric Contour Map.

appearance of inner edges is present, as in plots within which anomalies are perceived due to poor agricultural practices, land heterogeneity or crop diseases.

To compare numerically the results obtained by both approaches, the Boundary Displacement Error (BDE Freixenet et al., 2002; Yang et al., 2008) has been used as metric. BDE measures the average displacement error of boundary pixels between two segmented images. Particularly, it defines the error of one boundary pixel as the distance between the pixel and the closest pixel in the other boundary image. The value of BDE ranges between $[0, \infty)$ in the unit of pixel, lower is better (Yang et al., 2008). Each delineation obtained by the proposed approaches was compared with its corresponding ground-truth, in the case of the approach based on machine learning only the results applied to unseen data (testing set) were considered. Table 6.2 summarizes the results of this comparison. In addition, the results of applying an edge-detector method (reference method) to the images under study are also evaluated in order to server as a reference error. As can be observed, the approach based on the consensus of segmentations is slightly better than the one based on machine learning on both images. These results are consistent with what is visually appreciated. And both methods provide noteworthy better numerical results than the reference method.

Tab. 6.2.: Comparison between approaches for delineating agricultural plots. BDE is used to measure the error of these approaches.

	Image A	Image B
Approach based on classification	9.5740	11.0443
Approach based on consensus	7.9487	8.9699
Reference method	22.4555	23.8036

6.3.3 Conclusions

This Chapter has presented two approaches for the delineation of agricultural plots in high-resolution images (WorldView-2). The first approach combined superpixels and a supervised classification method to determine when adjacent superpixels should be merged. The obtained results shown that is possible to train a machine learning to delineate agricultural plots. In this regard, learning from data how agricultural plot are delineated poses an alternative

Fig. 6.10.: Comparison between the edges of the ground-truth (first column) and the edges obtained by the approach based on the consensus of segmentations with a threshold of 0.45 (second column). The edges, colored in black, have been thickened by morphological operations for better visualization.

to traditional segmentation algorithms, which could be exploited to imitate the labor of a human operator. The second approach proposed a consensus of diverse segmentations (superpixels) at different scales to determine the borders of the agricultural plots. The consensus of segmentations reduces image noise while facilitating edge detection and reducing the number of parameters to tune. In fact, only one threshold is necessary to obtain plot edges. Preliminary studies showed that a threshold about 0.45 is suitable to properly demarcate plots. A comparison between the two approaches showed that the latter provides slightly

more accurate results. However, both methods improve the results obtained by the reference method.

The results of this Chapter have been presented at two conferences, and submitted to a JCR journal (see Section 8.2), currently it is under a peer-review process.

A superpixel-based approach for temporal analysis of intra-plot variability

“... we may have knowledge of the past but cannot control it; we may control the future but have no knowledge of it.

— Claude E. Shannon (1916–2001)

Coding theorems for a discrete source with a fidelity criterion.

7.1 Introduction

In remote sensing literature, temporal analysis of a scene usually involves the detection of changes in it. The study of these changes in agricultural lands provides useful information about their internal variability. This information is important because of its practical uses in various applications, in particular in precision farming, for detecting agricultural anomalies, and realizing field-level improvements, such as drainage, land leveling, and irrigation. Formally, change detection (CD) is defined as (Singh, 1989): “the process of identifying differences in the state of an object or phenomenon by observing it at different times”.

CD methodologies analyze qualitatively the temporal effects of phenomena and quantify the changes using multi-temporal remote sensing datasets (Hussain et al., 2013). In this regard, remote sensing data provides spatially distributed information about alterations in the spectral behavior (reflectance value or local texture) produced by changes in the object of interest. A broad range of CD applications using multi-spectral images has been developed in the remote sensing literature, some examples are methods for analyzing difference images (Coppin and Bauer, 1996; Quarmby and Cushnie, 1989), classification-based approaches (Ghosh et al., 2011; Lunetta et al., 2006; Ji et al., 2006; Yuan et al., 2005), and kernel-based methods such as principal component analysis (Deng et al., 2008) or multivariate alteration detection (Nielsen and Vestergaard, 2013). These approaches compare corresponding image pixels of different acquisition times, and therefore they have the same limitations of the image analysis based on pixels, especially with VHR images.

In recent years, OBIA approaches for change detection have been developed. In contrast to pixel-based approaches, OBIA offers the possibility of using features from image-objects such

as spectral, textural, spatial, and topological, which are useful to model landscape in a more “realistic” way. From the point of view of the object-based analysis, Chen et al. (2012) define the CD as “the process of identifying differences in geographic objects at different moments using object-based image analysis”. In the remote sensing literature, three types of object-based approaches for change detection (OBCD) can be distinguished (Hussain et al., 2013):

Direct Object change detection. This approach performs a direct comparison between the image-objects from different dates. Thus, changes are detected either by comparing the geometrical properties (e.g., width and area) or comparing spectral information (e.g., mean spectral values) and/or extracted features from image-objects. The main disadvantage of these approaches is that change is linked to only the objects extracted from first image and therefore do not provide new objects that might be created in the second image because of change. Examples of these type of approaches are shown in the works of Lefebvre et al. (2008); Hall and Hay (2003); Tomowski et al. (2011).

Classified Objects change detection. This approach performs object-based analysis multi-temporal images to extract image-objects and independently classify them. In this regard, the performance of these approaches, similar to pixel-based approaches, is related to the accuracy of the classification algorithm. The classification accuracy is in turn related to the performance of an image segmentation technique which produces image-objects that can vary of shape and sizes depending on different segmentation parameters. Examples of these approaches can be found in the works of De Chant and Kelly (2009); Li and Narayanan (2003); Hazel (2001).

Multi-temporal-object change detection. In this approach, object-based analysis is applied directly to stacked multi-temporal images producing image-objects that correspond spatially, from which changes are analyzed. The stacked images may comprise co-registered panchromatic, multi-spectral, spectral indexes, or any spatially distributed variable in form of multi-temporal imagery. The result is a single segmentation of all the stacked images where image-objects are consistent in size, shape and location coordinate over time. Some works following this approach have been developed by Desclée et al. (2006); Bontemps et al. (2008); Conchedda et al. (2008).

Superpixels have been also used for change detection. Wu et al. (2012a) proposed a multi-temporal-object change detection based on superpixels for studying the effect of the tsunami produced in Indonesia in December 2004 through the comparison of two Quickbird images acquired before and after the tsunami. In their approach, PCA method is applied to reduce stacked images to three principal components and then segmented by SLIC method. From generated superpixels, change vectors are created using spectral, textural, and structural features. Then, a thresholding algorithm is applied to provide changed, unchanged, and undefined labels. Finally, a SVM classifier is trained with changed and unchanged samples in order to classify undefined areas. The main drawback of this method do to consider that changed pixels are not normally distributed, which can cause some problems due to the thresholding method. In general, these methods perform comparisons only between two

different dates, limiting the analysis of phenomena to this temporal window. Thus, with a temporal behavior that can change in more than two dates, these methods require performing more than one analysis.

In this Chapter, the study of the variability inside plots is explored. To this end, a novel method based on the superpixel algorithm, presented in Section 3, is proposed. In this regard, the proposed method extends the superpixel algorithm to deal with multi-temporal datasets. In this work, each layer of the stack corresponds to a multi-spectral image for a particular date. In particular, from a multi-temporal stack superpixels are created, similar to multi-temporal-object change detection approaches but instead of generating a single segmentation for the whole stack, as many segmentations as layers has the stack are generated. Superpixels may extend along all layers of the stack, thus, a superpixel may or may not appear in all dates. Thus, these temporal changes are considered in this work as an indicative of spatial variability. In this context, if the same superpixel appears on all dates are described as a sign of stability, whereas frequent changes are interpreted as a persistent variability.

7.2 Dataset and Methods

7.2.1 Data Set

The data set used in this study correspond to multi-temporal imagery of the study area described in Section 4.2.1. Three images captured by the sensors on board Pléiades satellite and subsequently pansharpened are available for the study. Pansharpening was performed by means of the WAT $\otimes$ FRAC algorithm (Lillo-Saavedra et al., 2011) but calculating the local entropy instead of the fractal dimension for weighting the contribution of spatial and spectral information. The resulting size of each image is 2527×3860 pixels, where each pixel represents an area of 0.5×0.5 m. The time series comprises a 35-day period for the late spring and early summer. Between each image, there is a gap of approximately 18 days. The images used in this study are shown in Figure 7.1, the dates under analysis are December 4 and 28, 2012, and January 18, 2013. Agricultural plots were delineated manually to establish the areas in which the analysis is performed.

7.2.2 Methodology

An overview of the analysis scheme is shown in Figure 7.2. The first stage of the proposed approach is the pre-processing of the images forming the stack.

Thus, in this stage images are atmospherically corrected via the the dark-object subtraction (DOS) technique (Chavez, 1988), in this way, effects of the atmosphere on the reflectance values are removed allowing the comparison between images of different dates. DOS technique

Fig. 7.1.: Multi-temporal multi-spectral images. A real-color composite (red, green, blue) is used to show images, borders of the agricultural plots are shown in red.

adjusts the brightness values of a given scene based on information in that particular image, i.e., the brightness values of the dark objects. The algorithm takes optical characteristics of the atmosphere in the image directly from the data in the image, which facilitates its execution as it requires no external parameters.

Then, to reduce image noise due to the spectral variability of the agricultural plots. In this work, image filtering is performed with the Rolling Guidance Filter, proposed by Zhang et al. (2014), which is based on the bilateral filter, that allows reducing image noise while maintaining main image borders. In the second stage, a proposed temporal superpixel method is applied to segment the stack of images.

In the last stage, simple rules (see Section 7.2.2) based on the behavior of generated superpixels are followed to determine where changes in the image have occurred.

Fig. 7.2.: Overview of the temporal analysis of agricultural plots.

Temporal segmentation

The temporal version of the SLIC segmentation method, here proposed, extends the method described in Section 3, to deal with multi-temporal datasets. This extension is based on the method proposed by Lucchi et al. (2012) for segmenting 3D images. In the proposed method, the three dimensions are represented by the two coordinates x and y for the spatial location and a z coordinate for the temporal component. This mainly affects the definition of the initial centers, and how the clustering distance between two pixels is calculated. The initial centers are now defined as:

$$C_k = [p^1, p^2, \dots, p^{NB}, x, y, z]^T, \quad k = 1, 2, \dots, K \quad (7.1)$$

On the other hand, the original spatial distance (Equation 2.22) now also takes into account the temporal coordinate:

$$d_s^2 = (x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2 \quad (7.2)$$

where subscripts i and j represent pixels p_i and p_j , respectively. Unlike spatial coordinates, the temporal coordinate is not restricted to a small space, but encompasses the full stack of images. The initialization is also modified to allow sampling seeds in the z coordinate. This sampling is performed by randomly selecting a date, while x and y coordinates are sampling as usual.

Variability Analysis Rules

Since the extended superpixel method performs a clustering in the feature space taking into account spatial and temporal components, generated superpixels represent spectrally similar small regions that can change in shape from a date to another, or simply can appear or disappear on any date. In this regard, when applied to agricultural plots, superpixels are able to capture their internal variability both spatial and temporal. Thus, the temporal evolution of a superpixel (i.e., its appearance or disappearance on any date) can be interpreted as change, and consequently, the persistence of this behavior is considered as an indicative of the variability of the regions in the image. The temporal evolution of a superpixel is shown in Figure 7.3, a change is produced when a superpixel do not appear on all dates (Figure 7.3a), in other case, there is no change (Figure 7.3b).

On this basis, simple rules can be formulated to establish the variability on a particular date (reference date). In this work, the grade of variability is determined according to the temporal evolution (i.e., changes) of superpixels as follows:

No variability. It is defined as the appearance of the same superpixel on all dates, i.e., the same phenomenon remains in all images.

Low-persistence variability. It is explained as the appearance of the same superpixel on the reference date and on consecutive dates before.

Fig. 7.3.: Temporal evolution of superpixels.

High-persistence variability. It is defined as the appearance of a superpixel on only the reference date, while in the rest of dates, it emerged other superpixels. In other words, it represents the fleeting appearance of a phenomenon in the image.

Thus from aforementioned rules, a map of variability can be created for each date in the multi-temporal stack.

7.3 Results and Discussion

In this work the four spectral bands (blue, green, red and near infrared) of the three pansharp-ened multi-spectral images available (Figure 7.1) were combined to create the multi-temporal stack used during the analysis. After the pre-processing methods are applied to each image in the stack, it is segmented using the proposed superpixel processing to obtain superpixels with a size of 500 pixels on average. In Figure 7.4, superpixels are shown in an enlarged region of study area. In general, superpixels are well adapted to the homogeneous regions inside the agricultural plots. As can be seen, superpixels may vary from a date to another, for example, the superpixels comprising the region covered by a cloud, shown in Figure 7.4a, only emerge in that date, while superpixels in the other dates (Figures 7.4b and 7.4c) correspond to other image-objects.

Figure 7.5 shows the map of changes obtained from applying the rule described in Section 7.2.2 to the superpixels generated in the third date (01-18-2013). In this map, areas inside agricultural plots that suffered any change are shown in yellow. As can be observed, most agricultural plots show no change in spectral level, this may be caused by the little difference in days between images. A total of 9.66 ha suffered any change along all dates, which represents the 10.60% of the surface under analysis, where as a 89.40% remains the same. Only a 18.74% of the areas that suffered changes has a high-persistence variability (i.e., these regions, represented by red color, are areas enclosed by superpixels that only appear on this date),

Fig. 7.4.: Superpixels generated for the different dates. A false-color composite (near infrared, green and blue) is used to highlight the vegetation in the images. Superpixel edges are shown in black.

while the remaining 81.26% represents areas with a low-persistence variability (i.e., regions that change only respect to one date). In this regard, this map provides an overview of the variability inside plots, considering three dates instead of two as other approaches commonly do.

Fig. 7.5.: Map of spatio-temporal dynamics of variability.

Figure 7.6 shows an enlarged area corresponding to the regions with a low-persistence variability according to Figure 7.5. As can be observed, regions of the second and third dates

(12-28-2012 and 12-28-2012) are spectrally similar between them, whereas there is a high difference between both dates and the first date (12-04-2012) because the appearance of a cloud covering some plots. Since the change is only observed regarding the first date, this change is considered as a low-persistence variability of these regions.

On the other hand, Figure 7.7 shows the regions with a high-persistence variability corresponding to the same area shown in Figure 7.6. As can be observed in this case, the regions are spectrally dissimilar on the three dates. Moreover, between the first and the second date textural changes can be also observed. In this context, the obtained results visually agree with reality.

Fig. 7.6.: Regions with a low-persistence variability according to the date 01-18-2013. Only these regions are shown inside the agricultural plots.

Fig. 7.7.: Regions with a high-persistence variability according to the date 01-18-2013. Only these regions are shown inside the agricultural plots.

7.4 Conclusions

This Chapter presented an approach for the temporal analysis of the variability inside agricultural plots using very high resolution multi-spectral temporal stack. This approach is based on an extended version of the superpixel algorithm (described in Chapter 3), which adds a new temporal component to the originally defined spatial components allowing to deal

with multi-temporal stacks of multi-spectral images. Thus, generated superpixels represented spectrally similar small regions that can change from a date to another. This temporal behavior was analyzed to provide a map of the variability inside agricultural plots. In this work, the appearance a disappearance of a superpixel was used as criterion to determine the variability map. Thus, three types of variability were considered: (i) no variability, (ii) low-persistence variability, and (iii) high-persistence variability. The first one described when a superpixel emerges on all dates, the second one represented the case in which a superpixel emerge on a few dates, and finally, the third case in which a superpixel only appears on a single date. Using this method, changes were detected in 9.66 ha equivalent to a 10.60% of the surface of the agricultural plots under analysis. From them, a 18.74% presented a high-persistence variability. Despite the lack of ground truth data that hinders the objective evaluation of the obtained results, they are visually promising; however some considerations such as an analysis regarding the changes in shapes would be made in order to improve the results.

Part IV

Conclusions and future work

Conclusions and future work

” *In five billion years, the sun will expand and engulf our orbit as the charred ember that was once Earth vaporizes. Have a nice day.*

— Neil deGrasse Tyson (1958–)

American astrophysicist, cosmologist, author, and science communicator.

Making use of very-high resolution remote sensing images, this thesis was introduced within the context of the analysis of fragmented agricultural landscapes. The focus was on generating information about spatial and temporal region dynamics of the agricultural fields following an object-based approach centered on the agricultural plots as objects of interest. This dissertation assumed that segmentation, in which object-based methodologies rely, is an ill-posed problem and therefore it should not be seen as an end in itself but simply as another part of the analysis process. Thus, this research was conducted by exploring novel image processing methods coming from the computer vision area and known as superpixel processing. In this regard, a superpixel method for remote sensing data (Chapter 3) was proposed in order to serve as a link between single pixels and meaningful objects in multi-spectral images, helping to reach positive results and representing an important milestone in the achievement of the objective addressed by this dissertation.

In this regard, to achieve this objective two previous studies centered on the entire scene of fragmented agricultural landscapes were proposed. The first methodology, described in Chapter 4, provided a framework for multi-scale segmentation and, at the same time, a way to identify the best scale of analysis according to the spectral intra-variability and inter-heterogeneity of the regions obtained on each scale. The proposed method for multi-scale segmentation was designed combining diverse segmentation generated using the aforementioned superpixel method with different superpixel sizes. Each superpixel size is associated to a scale of analysis. Two alternatives for generating scale were followed: (i) linear and (ii) dyadic scales. Final segmentation results exhibited a high dependence on how the scales were determined. The dyadic scales provided more realistic local visual scales, in the sense that different structures (agricultural plots, roads, and irrigation pivots) were well recognized but not in the segmentation based on the linear scales, due to the over-segmentation produced in this case. In this context, the methodology proved to adequately provide regions spectrally similar in all scales, which can be useful for agricultural applications such as land-cover mapping. Despite the above, the high variability inside agricultural plots is difficult to properly set

the limits thereof. The second methodology (Chapter 7) provided a framework that combines two scales corresponding to superpixels with edge-based segments for mapping land-covers in multi-spectral images of fragmented agricultural landscapes. The edge-based scale provided segments with edges closer to plot boundaries with no need of a time-demanding tuning process, whereas superpixel scale provided information about the scene at a scale between image pixels and the edge-based scale, reducing the analysis space and image variability, and consequently, the processing time during classification step. Object-based classification was performed by combining the results of a supervised classifier for superpixels and set of rules for classifying edge-based segments. The use of two complementary methods to generate the scales achieved a classification accuracy better than when those methods were performed separately. Although not all agricultural plots were well delineated, a good accuracy was achieved. In this regard, it can be concluded that the development of better method for delineating agricultural plots could reduce misclassification errors between these plots and neighboring land covers.

In particular, their use was explored during the research of the two specific objectives formulated in the Chapter 1 that contribute to the overall objective of this thesis.

- **Develop an automatic object-based method for delineating agricultural plots from very-high resolution satellite images.**

Two novel approaches were proposed for delineating agricultural plots in Chapter 6. These approaches exploited superpixels in two different ways in order to segmenting the plots and consequently obtained their boundaries. The first approach used a classification method for agglomerating superpixels, transforming the segmentation into a classification problem. Thus the classifier was trained to learn the relationships, in a multi-dimensional feature space, both the superpixels that belong to the same plot and those that are part of other objects. Unlike most traditional segmentation methods that require a priori knowledge about the scene for selecting optimal segmentation parameters, the proposed approach, through a supervised classification, was able to learn from previously segmented images (annotated data) how to perform the plot delineation. Although this represents advantage, the lack of enough annotated data for training could be seen as its main drawback limiting its application in most real-world applications. In this regard, the second proposed approach addressed a different direction. Based on the premise that different segmentation performed using different parameters can capture diverse phenomena, this approach combined different segmentation results in order to obtain a single result in which plots were segmented. This approach used diverse superpixel segmentation generated varying the superpixel size and weighting their borders according to the similarity between them. It was found that superpixels preserve most important plot boundaries in most of segmentation, which allowed through a segmentation based on boundaries to reach a consensus about the delineation of the agricultural plots. In both cases, obtained results outperformed the results generated by a reference method during their application to two very-high resolution images, proving their effectiveness.

- **Develop an automatic object-based method for the joint spatial and temporal analysis of the variability at plot level.**

For this objective, a novel approach for detecting the variability inside agricultural plots was developed and presented in Chapter 7. The proposed approach relied on a superpixel method that is able to deal with spatial and temporal components of multi-date imagery to find homogeneous regions inside plots. Thus temporal behavior of superpixels expressed as the changes in their shapes, as well as their appearance and disappearance through time have served as an indicative of landscape variability. Further analysis of this behavior through the different dates provided the information about the spatio-temporal variability inside the plots. In this regard, three behaviors were considered to determine the variability: (i) no variability, (ii) low-persistence variability, and (iii) high-persistence variability. The first one occurred when a superpixel emerges on all dates; the second one, described a superpixel that appears on almost all dates; finally, the third one, defined a superpixel that only emerges on a date. Above considerations allowed to create a map of spatio-temporal dynamics that provides an overview of the variability inside plots considering more than two images at once. However, a main drawback of the approach is the lack of visualization tools that facilitated the analysis of the generated information.

This work has allowed to satisfactorily verify the hypothesis; however, some problems have been encountered that open new lines of research which would improve the results obtained so far.

8.1 Future Work

Along the development of this dissertation, some constraints regarding the availability of annotated data were encountered, limiting the improve and in some cases the validation of the proposed approaches, moreover, the use of cutting-edge techniques such as deep learning methods that require large databases with annotations, is not possible in this circumstance. In this regard, there is a clear need to generate annotated information to improve the models used. Annotated data is an expensive asset in remote sensing area. The generation of this data commonly requires expensive field campaigns, and a vast experience in handling GIS software, as well as many hours of works. A future line of research can be oriented to generate tools that provide collaborative frameworks to annotate remote sensing images. In fact some preliminary works in this respect are currently being developed.

Other line of research opened during this dissertation is related to the visualization of the results of the study of variability inside plots raised in Chapter 7. In this regard, the development of a tool that allows stakeholders to easily interact and analyze this information could be of great helpful for the management of plots. Moreover, the use of additional information such as physical or nonphysical variables (e.g., evapotranspiration, ground maps) could be explored

in order to relate variability information with agricultural anomalies, for example, irrigation problems could be found by analyzing the variability in the evapotranspiration of the plot.

Moreover, an open-source software that includes all developed approaches in this dissertation could be implemented, in order to facilitate their distribution and use for the scientific community.

8.2 Selected Publications

To conclude this chapter, the most relevant publications, derived from this research and developed in previous Chapters, are referenced here.

JCR Journals:

- Garcia-Pedrero, A., Gonzalo-Martin, C., Fonseca-Luengo, D., and Lillo-Saavedra, M. (2015). A GEOBIA Methodology for Fragmented Agricultural Landscapes. Remote Sensing, 7(1), 767–787. Impact factor: 3.036 (Q1). Related to Chapter 5.
- Gonzalo-Martín, C., Lillo-Saavedra, M., Menasalvas, E., Fonseca-Luengo, D., García-Pedrero, A., and Costumero, R. (2015). Local optimal scale in a hierarchical segmentation method for satellite images: An OBIA approach for the agricultural landscape. Journal of Intelligent Information Systems, 1–13. Impact factor: 1.000 (Q3). Related to Chapter 4.
- García-Pedrero, A., Gonzalo-Martín, C., and Lillo-Saavedra, M. (Under review after revision). Machine Learning Approach for Agriculture Parcel Delineation Through Agglomerative Segmentation. Remote Sensing Letters. Impact factor: 1.487 (Q2). Related to Chapter 6.

International conferences:

- García-Pedrero, A., Gonzalo-Martín, C., Lillo-Saavedra, M., and Rodríguez-Esparragón, D. (2013). A strategy to improve quality and speed of multi-scale high-resolution image analysis through superpixels. In *Proceedings of the 33rd EARSeL Symposium Conference: Towards Horizon 2020: Earth Observation and Social Perspectives*, Matera, Italy (pp. 753–760). Related to Chapter 3.
- Fonseca-Luengo, D., García-Pedrero, A., Lillo-Saavedra, M., Costumero, R., Menasalvas, E., and Gonzalo-Martín, C. (2014). Optimal Scale in a Hierarchical Segmentation Method for Satellite Images. In *Rough Sets and Intelligent Systems Paradigms* (pp. 351–358). Springer International Publishing. Related to Chapter 4.

- García-Pedrero, A., Gonzalo-Martín, C., and Lillo-Saavedra, M. (2015). A machine learning approach for agricultural parcel delimitation through agglomerative segmentation. In *Proceedings of the 35th EARSeL Symposium Conference: European Remote Sensing – Progress, Challenges and Opportunities*, Stockholm, Sweden (pp. 68–73). Related to Chapter 6.

National conferences:

- García-Pedrero, A., Gonzalo-Martín, C., and Lillo-Saavedra, M. (2013). Una metodología OBIA basada en superpíxeles. In *Proceedings of the XV Congreso de la Asociación Española de Teledetección: Sistemas operacionales de observación de la Tierra*, Madrid, Spain (pp. 160–163). Related to Chapter 3.
- García-Pedrero, A., Gonzalo-Martín, C., Rodríguez-Esparragón, D. and Lillo-Saavedra, M. (2015). Delimitación de parcelas agrícolas mediante consenso de diversas segmentaciones. In *Proceedings of the XVI Congreso de la Asociación Española de Teledetección: Teledetección humedales y espacios protegidos*, Seville, Spain (pp. 60–63). Related to Chapter 6.

In addition, contributions indirectly related to the work performed during this dissertation are listed below.

JCR Journals:

- Ortiz Toro, C.A., Gonzalo Martín, C., García Pedrero, A., and Menasalvas Ruiz, E. (2015). Superpixel-Based Roughness Measure for Multispectral Satellite Image Segmentation. *Remote Sensing*, 7(11), 14620–14645. Impact factor: 3.036 (Q1).
- Rivera, D., Godoy-Faúndez, A., Lillo, M., Alvez, A., Delgado, V., Gonzalo-Martín, C., Menasalvas, E., Costumero, R., and García-Pedrero, A. (2016). Legal disputes as a proxy for regional conflicts over water rights in Chile. *Journal of Hydrology*, 535, 36–45. Impact factor: 3.043 (Q1).
- Rodríguez-Esparragón, D., Marcello, J., Eugenio-González, F., García-Pedrero, A., and Gonzalo-Martín, C. (2016). Object-based quality evaluation procedure for fused remote sensing imagery. Paper to appear in *Neurocomputing*. Impact factor: 2.392 (Q1).
- Ibarrola-Ulzurrun, E., Gonzalo-Martín, C., Marcello-Ruiz, J., García-Pedrero, A., and Rodríguez-Esparragón, D. (Under review). Fusion of High Resolution Multispectral Images in Coastal and Land Vulnerable Ecosystems. *Remote Sensing of Environment*. Impact factor: 5.881 (Q1).

- Costumero, R., Sánchez, J., García-Pedrero, A., Rivera, D., Lillo, M., Gonzalo-Martín, C. and Menasalvas, E. (Under review after revision). Geography of water legal disputes in Chile. Journal of maps. Impact factor: 1.435 (Q3).
- Gonzalo-Martín C., Lillo-Saavedra, M., García-Pedrero, A., Fonseca-Luengo, D., Lagos, O., and Menasalvas, E. (Under review after revision). Daily Evapotranspiration Mapping using Regression Random Forest Models. Remote Sensing. Impact factor: 3.036 (Q1).
- Lillo-Saavedra, M., Gonzalo-Martín C., García-Pedrero, A., and Lagos, O. (Under review). Scale-Aware Pansharpening Algorithm for Agricultural Fragmented Landscapes. Remote Sensing. Impact factor: 3.036 (Q1).

International conferences:

- Garcia-Pedrero, A., Rodríguez-Esparragón, D., Gonzalo-Martín, C., Ibarrola, E., Lillo-Saavedra, M., and Marcello, J. (2015). Automatic identification of shrub vegetation of the Teide National Park. In 4th International Work Conference on Bioinspired Intelligence (IWOBI), 2015 (pp. 55–60). IEEE.
- García-Pedrero, A., Gonzalo-Martín, C., Lillo-Saavedra, M., and Ortiz-Toro, C. (2016). A web-based platform for remote sensing image annotation. To be presented in the GEOBIA 2016: solutions and synergies, 6th GEOBIA Conference, in Enschede, The Netherlands, from 14-16 September 2016.
- Gonzalo-Martín, C., García-Pedrero, A., Lillo-Saavedra, M., and Menasalvas, E. (2016). Deep Learning for superpixel-based classification of remote sensing images. To be presented in the GEOBIA 2016: solutions and synergies, 6th GEOBIA Conference, in Enschede, The Netherlands, from 14-16 September 2016.

Appendixes

A.1 Superpixels in Remote Sensing

Fig. A.1.: Publications in conferences that use superpixels related to remote sensing area.

Fig. A.2.: Publications in JCR Jorunals that use superpixels related to remote sensing area.

Bibliography

- Achanta, R, A Shaji, et al. (2010). “Slic superpixels”. In: *École Polytechnique Fédéral de Laussanne (EPFL), Tech. Rep 149300* (cit. on p. 40).
- (2012). “SLIC Superpixels Compared to State-of-the-Art Superpixel Methods”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 34.11, pp. 2274–2282 (cit. on pp. 38–40).
- Acuña, Rafael Guillermo Gonzalez, Junli Tao, et al. (2015). “Robust Segmentation of Aerial Image Data Recorded for Landscape Ecology Studies”. In: *Pacific-Rim Symposium on Image and Video Technology*. Springer, pp. 61–72 (cit. on pp. 110, 111).
- Adams, R. and L. Bischof (1994). “Seeded region growing”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16.6, pp. 641–647 (cit. on p. 23).
- Aggarwal, Charu C (2014). “An Introduction to Data Classification”. In: *Data Classification: Algorithms and Applications*. CRC Press, pp. 1–36 (cit. on pp. 28, 30–32).
- Aiazzi, Bruno, Luciano Alparone, Stefano Baronti, rea Garzelli, and Massimo Selva (2012). “Twenty-Five Years of Pansharpening”. In: *Signal and Image Processing for Remote Sensing, Second Edition*. CRC Press, pp. 533–548 (cit. on p. 21).
- Aichert, André (2008). “Feature extraction techniques”. In: *Camp medical seminar ws0708* (cit. on p. 27).
- Akbarizadeh, Gholamreza and Marjan Aleghafour (2015). “Unsupervised hierarchical SAR image segmentation using lossy data compression”. In: *Information and Knowledge Technology (IKT), 2015 7th Conference on*. IEEE, pp. 1–4 (cit. on pp. 110, 111).
- Aksoy, Selim, H Gökhan Akçay, and Tom Wassenaar (2010). “Automatic mapping of linear woody vegetation features in agricultural landscapes using very high resolution imagery”. In: *IEEE Transactions on Geoscience and Remote Sensing* 48.1, pp. 511–522 (cit. on p. 35).
- Alcantara, Camilo, Tobias Kuemmerle, Alexander V Prishchepov, and Volker C Radeloff (2012). “Mapping abandoned agriculture with multi-temporal MODIS satellite data”. In: *Remote Sensing of Environment* 124, pp. 334–347 (cit. on p. 4).
- Alhichri, Haikel, Yacoub Bazi, Naif Alajlan, and Nassim Ammour (2015). “A hierarchical learning paradigm for semi-supervised classification of remote sensing images”. In: *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*. IEEE, pp. 4388–4391 (cit. on pp. 110, 111).
- Altman, Naomi S (1992). “An introduction to kernel and nearest-neighbor nonparametric regression”. In: *The American Statistician* 46.3, pp. 175–185 (cit. on p. 27).
- Aplin, P. and P. M. Atkinson (2001). “Sub-pixel land cover mapping for per-field classification”. In: *International Journal of Remote Sensing* 22.14, pp. 2853–2858 (cit. on p. 18).

- Arbelaez, Pablo, Michael Maire, Charless Fowlkes, and Jitendra Malik (2009). "From contours to regions: An empirical evaluation". In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, pp. 2294–2301 (cit. on p. 83).
- Arthur, David and Sergei Vassilvitskii (2007). "K-means++: The Advantages of Careful Seeding". In: *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA '07. New Orleans, Louisiana: Society for Industrial and Applied Mathematics, pp. 1027–1035 (cit. on p. 27).
- Atkinson, Peter M and Nicholas J Tate (2000). "Spatial scale problems and geostatistical solutions: a review". In: *The Professional Geographer* 52.4, pp. 607–623 (cit. on p. 16).
- Atzberger, Clement (2013). "Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs". In: *Remote Sensing* 5.2, pp. 949–981 (cit. on p. 4).
- Atzberger, Clement, Francesco Vuolo, et al. (2015). "Agriculture". In: *Remote Sensing Handbook: Land Resources Monitoring, Modeling, and Mapping with Remote Sensing* 2, pp. 113–128 (cit. on pp. 33–35).
- Baatz, Martin and Arno Schäpe (2000). *Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation* (cit. on pp. 25, 52).
- Bauer, Marvin E and Jan E Cipra (1973). "Identification of agricultural crops by computer processing of ERTS MSS data". In: (cit. on p. 4).
- Bégué, Agnès, Damien Arvor, Camille Lelong, Elodie Vintrou, and Margareth Simoes (2015). "Agricultural Systems Studies using Remote Sensing". In: *Remote Sensing Handbook: Land Resources Monitoring, Modeling, and Mapping with Remote Sensing* 2, pp. 113–128 (cit. on p. 5).
- Benz, Ursula C, Peter Hofmann, Gregor Willhauck, Iris Lingenfelder, and Markus Heynen (2004). "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information". In: *ISPRS-J. Photogramm. Remote Sens.* 58.3, pp. 239–258 (cit. on p. 51).
- Berberoglu, Süha and Paul J Curran (2004). "Merging spectral and textural information for classifying remotely sensed images". In: *Remote sensing image analysis: including the spatial domain*. Springer, pp. 113–136 (cit. on p. 18).
- Besbes, Olfa, Nozha Boujemaa, and Ziad Belhadj (2009). "Cue Integration for Urban Area Extraction in Remote Sensing Images". In: *Image Analysis and Recognition*. Springer, pp. 248–257 (cit. on pp. 110, 111).
- Bhanu, Bir and Sungkee Lee (2012). *Genetic learning for adaptive image segmentation*. Vol. 287. Springer Science & Business Media (cit. on p. 22).
- Bisoi, Ajay Kumar and Jibitesh Mishra (2001). "On calculation of fractal dimension of images". In: *Pattern Recognition Letters* 22.6–7, pp. 631–637 (cit. on p. 29).
- Blaschke, T (2010). "Object based image analysis for remote sensing". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 65.1, pp. 2–16 (cit. on p. 14).
- Blaschke, T, GJ Hay, et al. (2014). "Geographic Object-Based Image Analysis – Towards a new paradigm". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 87, pp. 180–191 (cit. on pp. 7, 8, 13–17, 38).
- Blaschke, T, S Lang, and G Hay, eds. (2008). *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*. Lecture Notes in Geoinformation and Cartography. Springer (cit. on pp. 14, 17).

- Blaschke, T and J Strobl (2001). "What's wrong with pixels? Some recent developments interfacing remote sensing and GIS". In: *GeoBIT/GIS* 6.01, pp. 12–17 (cit. on pp. 8, 14).
- Blaschke, Thomas, Charles Burnett, and Anssi Pekkarinen (2004). "Image segmentation methods for object-based analysis and classification". In: *Remote sensing image analysis: Including the spatial domain*. Springer, pp. 211–236 (cit. on pp. 22–24).
- Blaschke, Thomas, Maggi Kelly, and Helena Merschdorf (2015). "Object-Based Image Analysis: Evolution, History, State of the Art, and Future Vision". In: *Remotely Sensed Data Characterization, Classification, and Accuracies*. CRC Press, pp. 277–294 (cit. on pp. 13, 38).
- Bontemps, Sophie, Patrick Bogaert, Nicolas Titeux, and Pierre Defourny (2008). "An object-based change detection method accounting for temporal dependences in time series with medium to coarse spatial resolution". In: *Remote Sensing of Environment* 112.6, pp. 3181–3191 (cit. on p. 92).
- Boykov, Yuri Y and Marie-Pierre Jolly (2001). "Interactive graph cuts for optimal boundary & region segmentation of objects in ND images". In: *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*. Vol. 1. IEEE, pp. 105–112 (cit. on p. 24).
- Breiman, Leo (2001). "Random forests". In: *Machine learning* 45.1, pp. 5–32 (cit. on pp. 27, 31).
- Burnett, Carolyn and Thomas Blaschke (2003). "A multi-scale segmentation/object relationship modelling methodology for landscape analysis". In: *Ecological modelling* 168.3, pp. 233–249 (cit. on p. 52).
- Canny, John (1986). "A Computational Approach to Edge Detection". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI*-8.6, pp. 679–698 (cit. on p. 66).
- Canty, Morton J (2011). *Image Analysis, Classification, and Change Detection in Remote Sensing : With Algorithms for ENVI/IDL, Second Edition*. 2nd ed. CRC Press (cit. on p. 23).
- Carper, W Joseph (1990). "The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data". In: *Photogramm. Eng. Remote Sens.* 56.4, pp. 457–467 (cit. on p. 79).
- Castilla, G and GJ Hay (2008). "Image objects and geographic objects". In: *Object-based image analysis*. Springer, pp. 91–110 (cit. on pp. 14–16, 36, 37).
- Chang, Yian-Leng and Xiaobo Li (1994). "Adaptive image region-growing". In: *Image Processing, IEEE Transactions on* 3.6, pp. 868–872 (cit. on p. 23).
- Chavez, Pat S (1988). "An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data". In: *Remote sensing of environment* 24.3, pp. 459–479 (cit. on p. 93).
- Chawla, Nitesh V, Aleksandar Lazarevic, Lawrence O Hall, and Kevin W Bowyer (2003). "SMOTEBoost: Improving Prediction of the Minority Class in Boosting". In: *Knowledge Discovery in Databases: PKDD 2003*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 107–119 (cit. on p. 32).
- Chen, Gang, Geoffrey J. Hay, Luis M. T. Carvalho, and Michael A. Wulder (2012). "Object-based change detection". In: *International Journal of Remote Sensing* 33.14, pp. 4434–4457 (cit. on p. 92).
- Chen, Qiu-Xiao, Jian-Cheng Luo, Cheng-hu Zhou, and Tao Pei (2003). "A hybrid multi-scale segmentation approach for remotely sensed imagery". In: *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International*. Vol. 6, pp. 3416–3419 (cit. on p. 52).
- Chen, Yunhao, Wei Su, Jing Li, and Zhongping Sun (2009). "Hierarchical object oriented classification using very high resolution imagery and LIDAR data over urban areas". In: *Advances in Space Research* 43.7, pp. 1101–1110 (cit. on p. 67).

- Chen, Zhao and Bin Wang (2014). "An improved spectral-spatial classification framework for hyperspectral remote sensing images". In: *Audio, Language and Image Processing (ICALIP), 2014 International Conference on*. IEEE, pp. 532–536 (cit. on pp. 110, 111).
- Cheng, H.D., X.H. Jiang, Y. Sun, and Jingli Wang (2001). "Color image segmentation: advances and prospects". In: *Pattern Recognition* 34.12, pp. 2259–2281 (cit. on p. 24).
- Comaniciu, Dorin and Peter Meer (2002). "Mean shift: A robust approach toward feature space analysis". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 24.5, pp. 603–619 (cit. on pp. 27, 39).
- Conchedda, Giulia, Laurent Durieux, and Philippe Mayaux (2008). "An object-based method for mapping and change analysis in mangrove ecosystems". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 63.5, pp. 578–589 (cit. on p. 92).
- Conrad, Christopher, Sebastian Fritsch, Julian Zeidler, Gerd Rücker, and Stefan Dech (2010). "Per-Field Irrigated Crop Classification in Arid Central Asia Using SPOT and ASTER Data". In: *Remote Sensing* 2.4, p. 1035 (cit. on pp. 18, 34).
- Coppin, Pol R and Marvin E Bauer (1996). "Digital change detection in forest ecosystems with remote sensing imagery". In: *Remote sensing reviews* 13.3-4, pp. 207–234 (cit. on p. 91).
- Cui, Xiaoguang, Yuan Tian, Lubin Weng, and Yiping Yang (2014). "Anomaly detection in hyperspectral imagery based on low-rank and sparse decomposition". In: *Fifth International Conference on Graphic and Image Processing*. International Society for Optics and Photonics, 90690R–90690R (cit. on pp. 110, 111).
- Da Costa, Jean Pierre, Franck Michelet, Christian Germain, Olivier Lavialle, and Gilbert Grenier (2007). "Delineation of vine parcels by segmentation of high resolution remote sensed images". In: *Precision Agriculture* 8.1-2, pp. 95–110 (cit. on p. 78).
- De Chant, Tim and Maggi Kelly (2009). "Individual object change detection for monitoring the impact of a forest pathogen on a hardwood forest". In: *Photogrammetric Engineering & Remote Sensing* 75.8, pp. 1005–1013 (cit. on p. 92).
- De Wit, AJW and JGPW Clevers (2004). "Efficiency and accuracy of per-field classification for operational crop mapping". In: *International journal of remote sensing* 25.20, pp. 4091–4112 (cit. on p. 34).
- Debats, Stephanie R., Dee Luo, Lyndon D. Estes, Thomas J. Fuchs, and Kelly K. Caylor (2016). "A generalized computer vision approach to mapping crop fields in heterogeneous agricultural landscapes". In: *Remote Sensing of Environment* 179, pp. 210–221 (cit. on p. 9).
- Definiens, AG (2009). "Definiens Developer 7 user guide". In: *München: Definiens AG* (cit. on pp. 28, 29).
- Deng, Chunhua, Zhiguo Cao, Zhiwen Fang, and Zhenghong Yu (2013). "Ship detection from optical satellite image using optical flow and saliency". In: *Eighth International Symposium on Multispectral Image Processing and Pattern Recognition*. International Society for Optics and Photonics, 89210F–89210F (cit. on pp. 110, 111).
- Deng, JS, K Wang, YH Deng, and GJ Qi (2008). "PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data". In: *International Journal of Remote Sensing* 29.16, pp. 4823–4838 (cit. on p. 91).
- Desclée, Baudouin, Patrick Bogaert, and Pierre Defourny (2006). "Forest change detection by statistical object-based method". In: *Remote Sensing of Environment* 102.1, pp. 1–11 (cit. on p. 92).

- DigitalGlobe (2015). *Remote Sensing Technology Trends and Agriculture*. <https://dg-cms-uploads-production.s3.amazonaws.com/uploads/document/file/31/DG-RemoteSensing-WP.pdf>. [Online; accessed 27-April-2016] (cit. on p. 6).
- Dollár, Piotr and C Lawrence Zitnick (2013). “Structured forests for fast edge detection”. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1841–1848 (cit. on p. 83).
- Draguț, Lucian, Dirk Tiede, and Shaun R Levick (2010). “ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data”. In: *International Journal of Geographical Information Science* 24.6, pp. 859–871 (cit. on p. 37).
- Duan, Wuhui, Shutao Li, and Leyuan Fang (2015). “Superpixel-based composite kernel for hyperspectral image classification”. In: *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*. IEEE, pp. 1698–1701 (cit. on pp. 110, 111).
- Duro, Dennis C, Steven E Franklin, and Monique G Dubé (2012). “A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery”. In: *Remote Sensing of Environment* 118, pp. 259–272 (cit. on pp. 35, 36).
- Duro, Dennis Correa (2012). “Remote sensing and GIS in support of sustainable agricultural development”. PhD thesis. Saskatoon, Canada: University of Saskatchewan (cit. on p. 18).
- Duveiller, Grégory and Pierre Defourny (2010). “A conceptual framework to define the spatial resolution requirements for agricultural monitoring using remote sensing”. In: *Remote Sensing of Environment* 114.11, pp. 2637–2650 (cit. on p. 52).
- Edlinger, Jana, Christopher Conrad, John Lamers, Gulchekhra Khasankhanova, and Thomas Koellner (2012). “Reconstructing the spatio-temporal development of irrigation systems in Uzbekistan using Landsat time series”. In: *Remote Sensing* 4.12, pp. 3972–3994 (cit. on p. 4).
- Ehlers, Manfred (2008). “Geoinformatics and digital earth initiatives: a German perspective”. In: *International journal of digital earth* 1.1, pp. 17–30 (cit. on p. 15).
- Espindola, GM, Gilberto Câmara, IA Reis, LS Bins, and AM Monteiro (2006). “Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation”. In: *International Journal of Remote Sensing* 27.14, pp. 3035–3040 (cit. on p. 52).
- Ester, Martin, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu (1996). “A density-based algorithm for discovering clusters in large spatial databases with noise.” In: *Kdd*. Vol. 96. 34, pp. 226–231 (cit. on p. 27).
- Felzenszwalb, Pedro F and Daniel P Huttenlocher (2004). “Efficient graph-based image segmentation”. In: *International Journal of Computer Vision* 59.2, pp. 167–181 (cit. on pp. 24, 39).
- Fonseca-Luengo, David, Angel García-Pedrero, et al. (2014). “Optimal scale in a hierarchical segmentation method for satellite images”. In: *Rough Sets and Intelligent Systems Paradigms*. Springer, pp. 351–358 (cit. on pp. 110, 111).
- Food and Agriculture Organization (2009). “How to Feed the World in 2050”. In: *Rome: High-Level Expert Forum* (cit. on p. 3).
- Forster, Dionys, Tobias Walter Kellenberger, Yves Buehler, and Bernd Lennartz (2010). “Mapping diversified peri-urban agriculture–potential of object-based versus per-field land cover/land use classification”. In: *Geocarto International* 25.3, pp. 171–186 (cit. on pp. 9, 18, 34).
- Fotheringham, A Stewart, Chris Brunsdon, and Martin Charlton (2000). *Quantitative geography: perspectives on spatial data analysis*. Sage (cit. on p. 56).

- Freeman, William T and Edward H Adelson (1991). "The design and use of steerable filters". In: *IEEE Transactions on Pattern analysis and machine intelligence* 13.9, pp. 891–906 (cit. on p. 83).
- Freixenet, Jordi, Xavier Muñoz, David Raba, Joan Martí, and Xavier Cufí (2002). "Yet another survey on image segmentation: Region and boundary information integration". In: *European Conference on Computer Vision*. Springer, pp. 408–422 (cit. on p. 88).
- Freund, Y and R E Schapire (1996). "Experiments with a new boosting algorithm". In: *ICML* (cit. on p. 32).
- Fu, King-Sun and JK Mui (1981). "A survey on image segmentation". In: *Pattern recognition* 13.1, pp. 3–16 (cit. on p. 24).
- Galar, M, A Fernández, E Barrenechea, H Bustince, and F Herrera (2012). "A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches". In: *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 42.4, pp. 463–484 (cit. on p. 32).
- Galford, Gillian L, John F Mustard, et al. (2008). "Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil". In: *Remote sensing of environment* 112.2, pp. 576–587 (cit. on p. 4).
- Galletti, Christopher S and Soe W Myint (2014). "Land-use mapping in a mixed urban-agricultural arid landscape using object-based image analysis: A case study from Maricopa, Arizona". In: *Remote Sensing* 6.7, pp. 6089–6110 (cit. on p. 35).
- Gao, Yan and Jean Francois Mas (2008). "A comparison of the performance of pixel-based and object-based classifications over images with various spatial resolutions". In: *Online Journal of Earth Sciences* 2.1, pp. 27–35 (cit. on p. 51).
- Gao, Zhiqiang, Wei Gao, and Ni-Bin Chang (2011). "Integrating temperature vegetation dryness index (TVDI) and regional water stress index (RWSI) for drought assessment with the aid of LANDSAT TM/ETM+ images". In: *International Journal of Applied Earth Observation and Geoinformation* 13.3, pp. 495–503 (cit. on p. 4).
- Garcia-Pedrero, Angel, Consuelo Gonzalo-Martin, David Fonseca-Luengo, and Mario Lillo-Saavedra (2015a). "A GEOBIA methodology for fragmented agricultural landscapes". In: *Remote Sensing* 7.1, pp. 767–787 (cit. on p. 38).
- García-Pedrero, Angel, Consuelo Gonzalo-Martín, Mario Lillo-Saavedra, and Dionisio Rodríguez-Esparragón (2013). "A strategy to improve quality and speed of multi-scale high-resolution image analysis through superpixels". In: *Towards Horizon 2020*, pp. 753–760 (cit. on pp. 110, 111).
- Garcia-Pedrero, Angel, Dionisio Rodriguez-Esparragon, et al. (2015b). "Automatic identification of shrub vegetation of the Teide National Park". In: *Bioinspired Intelligence (IWOB), 2015 4th International Work Conference on*. IEEE, pp. 55–60 (cit. on pp. 110, 111).
- Gascon, Ferran, E Cadau, et al. (2014). "Copernicus Sentinel-2 mission: products, algorithms and Cal/Val". In: *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, 92181E–92181E (cit. on p. 4).
- Ghosh, Ashish, Niladri Shekhar Mishra, and Susmita Ghosh (2011). "Fuzzy clustering algorithms for unsupervised change detection in remote sensing images". In: *Information Sciences* 181.4, pp. 699–715 (cit. on p. 91).

- Gianinetto, Marco, Marco Rusmini, et al. (2014). "Hierarchical classification of complex landscape with VHR pan-sharpened satellite data and OBIA techniques". In: *European Journal of Remote Sensing* 47, pp. 229–250 (cit. on p. 67).
- Gonzalez, Rafael C. and Richard E. Woods (2007). *Digital Image Processing*. 3rd Edition. Pearson (cit. on pp. 20, 23, 28, 29, 67).
- Gonzalez, Rafael C., Richard E. Woods, and Steven L. Eddins (2009). *Digital Image Processing Using MATLAB*. 2nd edition. Gatesmark Publishing (cit. on p. 29).
- González-Sanpedro, MC, Thuy Le Toan, J Moreno, L Kergoat, and E Rubio (2008). "Seasonal variations of leaf area index of agricultural fields retrieved from Landsat data". In: *Remote Sensing of Environment* 112.3, pp. 810–824 (cit. on p. 4).
- Gonzalo, Consuelo and Mario Lillo-Saavedra (2008). "A directed search algorithm for setting the spectral-spatial quality trade-off of fused images by the wavelet à trous method". In: *Canadian Journal of Remote Sensing* 34.4, pp. 367–375 (cit. on p. 21).
- Govaerts, Bram, Nele Verhulst, et al. (2007). "Evaluating spatial within plot crop variability for different management practices with an optical sensor?" In: *Plant and Soil* 299.1-2, pp. 29–42 (cit. on p. 7).
- Grecchi, Rosana Cristina, Q Hugh J Gwyn, Goze Bertin Benie, Antonio Roberto Formaggio, and Fernando Cesar Fahl (2014). "Land use and land cover changes in the Brazilian Cerrado: A multidisciplinary approach to assess the impacts of agricultural expansion". In: *Applied Geography* 55, pp. 300–312 (cit. on p. 35).
- Guyon, Isabelle and André Elisseeff (2006). "An introduction to feature extraction". In: *Feature extraction*. Springer, pp. 1–25 (cit. on p. 27).
- Hall, Ola and Geoffrey J Hay (2003). "A multiscale object-specific approach to digital change detection". In: *International Journal of Applied Earth Observation and Geoinformation* 4.4, pp. 311–327 (cit. on p. 92).
- Haralick, Robert M (1979). "Statistical and structural approaches to texture". In: *Proceedings of the IEEE* 67.5, pp. 786–804 (cit. on p. 29).
- Haralick, Robert M and Linda G Shapiro (1985). "Image segmentation techniques". In: *Computer vision, graphics, and image processing* 29.1, pp. 100–132 (cit. on p. 24).
- Harris, Ray and Ingo Baumann (2015). "Open data policies and satellite Earth observation". In: *Space Policy* 32, pp. 44–53 (cit. on p. 4).
- Hay, Geoffrey J, Thomas Blaschke, Danielle J Marceau, and André Bouchard (2003). "A comparison of three image-object methods for the multiscale analysis of landscape structure". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 57.5, pp. 327–345 (cit. on p. 16).
- Hay, Geoffrey J and G Castilla (2008). "Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline". In: *Object-based image analysis*. Springer, pp. 75–89 (cit. on pp. 17, 28, 36).
- Hay, Geoffrey J, B Hemachandran, et al. (2014). *GEOBIA – Evolving Beyond Segmentation*. <http://geobia2014.web.auth.gr/geobia14/sites/default/files/pictures/hay.pdf>. [Online; accessed 28-November-2015] (cit. on p. 15).
- Hay, GJ, KO Niemann, and GF McLean (1996). "An object-specific image-texture analysis of H-resolution forest imagery". In: *Remote Sensing of Environment* 55.2, pp. 108–122 (cit. on pp. 7, 13).

- Hazel, Geoffrey G (2001). "Object-level change detection in spectral imagery". In: *IEEE Transactions on Geoscience and Remote Sensing* 39.3, pp. 553–561 (cit. on p. 92).
- He, Wenju, Marc Jäger, and Olaf Hellwich (2008). "Comparison of three unsupervised segmentation algorithms for SAR data in urban areas". In: *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*. Vol. 1. IEEE, pp. I-241 (cit. on pp. 110, 111).
- Heikkilä, Marko, Matti Pietikäinen, and Cordelia Schmid (2009). "Description of interest regions with local binary patterns". In: *Pattern Recognition* 42.3, pp. 425–436 (cit. on p. 29).
- Henebry, Geoffrey M and Kirsten M de Beurs (2013). "Remote sensing of land surface phenology: a prospectus". In: *Phenology: An Integrative Environmental Science*. Springer, pp. 385–411 (cit. on pp. 4, 5).
- Hernando, A, D Tiede, F Albrecht, and S Lang (2012). "Spatial and thematic assessment of object-based forest stand delineation using an OFA-matrix". In: *Int. J. Applied Earth Observation and Geoinformation* 19, pp. 214–225 (cit. on p. 17).
- Heumann, Benjamin W. (2011). "An Object-Based Classification of Mangroves Using a Hybrid Decision Tree—Support Vector Machine Approach". In: *Remote Sensing* 3.11, pp. 2440–2460 (cit. on p. 30).
- Huang, Xiaojing, Wen Yang, Gui-Song Xia, and Mingsheng Liao (2015). "Superpixel-based change detection in high resolution sar images using region covariance features". In: *Analysis of Multitemporal Remote Sensing Images (Multi-Temp)*, 2015 8th International Workshop on the. IEEE, pp. 1–4 (cit. on pp. 110, 111).
- Hussain, Masroor, Dongmei Chen, Angela Cheng, Hui Wei, and David Stanley (2013). "Change detection from remotely sensed images: From pixel-based to object-based approaches". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 80, pp. 91–106 (cit. on pp. 91, 92).
- International Fund for Agricultural Development and United Nations Environment Programme (2013). "Smallholders, food security and the environment". In: *Rome: International Fund for Agricultural Development* (cit. on p. 5).
- Jackson, Thomas J, Daoyi Chen, et al. (2004). "Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans". In: *Remote Sensing of Environment* 92.4, pp. 475–482 (cit. on p. 4).
- Jayne, TS, David Mather, and Elliot Mghenyi (2010). "Principal challenges confronting smallholder agriculture in sub-Saharan Africa". In: *World development* 38.10, pp. 1384–1398 (cit. on p. 5).
- Ji, Wei, Jia Ma, Rima Wahab Twibell, and Karen Underhill (2006). "Characterizing urban sprawl using multi-stage remote sensing images and landscape metrics". In: *Computers, Environment and Urban Systems* 30.6, pp. 861–879 (cit. on p. 91).
- Jiang, Qiling, Liujuan Cao, Ming Cheng, Cheng Wang, and Jonathan Li (2015). "Deep neural networks-based vehicle detection in satellite images". In: *Bioelectronics and Bioinformatics (ISBB)*, 2015 International Symposium on. IEEE, pp. 184–187 (cit. on pp. 110, 111).
- Jiao, Limin and Yaolin Liu (2012). "Analyzing the shape characteristics of land use classes in remote sensing imagery". In: *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci*, pp. 135–140 (cit. on p. 67).
- Johnson, Brian and Zhixiao Xie (2011). "Unsupervised image segmentation evaluation and refinement using a multi-scale approach". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 66.4, pp. 473–483 (cit. on pp. 52, 55).
- Jones, Hamlyn G and Robin A Vaughan (2010). *Remote sensing of vegetation: principles, techniques, and applications*. Oxford university press (cit. on p. 3).

- Kass, Michael, Andrew Witkin, and Demetri Terzopoulos (1988). "Snakes: Active contour models". In: *International journal of computer vision* 1.4, pp. 321–331 (cit. on p. 78).
- Ke, Li, You Xiong, and Wan Gang (2015). "Remote Sensing Image Classification Method Based on Superpixel Segmentation and Adaptive Weighting K-Means". In: *2015 International Conference on Virtual Reality and Visualization (ICVRV)*. IEEE, pp. 40–45 (cit. on pp. 110, 111).
- Kim, K. S., D. Zhang, M. C. Kang, and S. J. Ko (2013). "Improved simple linear iterative clustering superpixels". In: *2013 IEEE International Symposium on Consumer Electronics (ISCE)*, pp. 259–260 (cit. on pp. 40, 41).
- Kim, M, M Madden, and T Warner (2008). "Estimation of optimal image object size for the segmentation of forest stands with multispectral IKONOS imagery". In: *Object-based image analysis*. Springer, pp. 291–307 (cit. on pp. 7, 8, 13).
- Kim, Minho, Marguerite Madden, and Timothy A Warner (2009). "Forest type mapping using object-specific texture measures from multispectral Ikonos imagery". In: *Photogrammetric Engineering & Remote Sensing* 75.7, pp. 819–829 (cit. on p. 52).
- Lang, S (2008). "Object-based image analysis for remote sensing applications: modeling reality–dealing with complexity". In: *Object-based image analysis*. Springer, pp. 3–27 (cit. on p. 15).
- Lang, Stefan and Dirk Tiede (2015). "Geospatial Data Integration in OBIA: Implications of Accuracy and Validity". In: *Remotely Sensed Data Characterization, Classification, and Accuracies*. CRC Press, pp. 295–316 (cit. on p. 15).
- Lefebvre, Antoine, Thomas Corpetti, and Laurence Hubert-Moy (2008). "Object-oriented approach and texture analysis for change detection in very high resolution images". In: *IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium*. Vol. 4. IEEE, pp. IV–663 (cit. on p. 92).
- Leo, O and G Lemoine (2001). *Land Parcel Identification Systems in the Frame of Regulation (EC) 1593/2000 Version 1.4*. Tech. rep. Institute for Environment and Sustainability: Ispra, Italia (cit. on p. 77).
- Levin, S.A. (1992). "The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture". In: *Ecology* 73.6, pp. 1943–1967 (cit. on p. 52).
- Levinstein, Alex, Adrian Stere, et al. (2009). "Turbopixels: Fast superpixels using geometric flows". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 31.12, pp. 2290–2297 (cit. on p. 39).
- Li, HT, HY Gu, YS Han, and JH Yang (2008). "An efficient multi-scale segmentation for high-resolution remote sensing imagery based on statistical region merging and minimum heterogeneity rule". In: *Earth Observation and Remote Sensing Applications, 2008. EORSA 2008. International Workshop on*. IEEE, pp. 1–6 (cit. on p. 52).
- Li, Huali, Shutao Li, and Liangpei Zhang (2015). "Adaptive endmember extraction based sparse nonnegative matrix factorization with spatial local information". In: *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, pp. 1753–1756 (cit. on pp. 110, 111).
- Li, Jiang and Ram M Narayanan (2003). "A shape-based approach to change detection of lakes using time series remote sensing images". In: *IEEE Transactions on Geoscience and Remote Sensing* 41.11, pp. 2466–2477 (cit. on p. 92).
- Li, Manchun, Lei Ma, Thomas Blaschke, Liang Cheng, and Dirk Tiede (2016). "A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments". In: *International Journal of Applied Earth Observation and Geoinformation* 49, pp. 87–98 (cit. on pp. 30, 35).

- Li, Shanshan, Xiuping Jia, and Bing Zhang (2013). "Superpixel-based Markov random field for classification of hyperspectral images". In: *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*. IEEE, pp. 3491–3494 (cit. on pp. 110, 111).
- Li, Zhenguo, Xiao-Ming Wu, and Shih-Fu Chang (2012). "Segmentation using superpixels: A bipartite graph partitioning approach". In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, pp. 789–796 (cit. on p. 38).
- Liang, Shunlin (2008). *Advances in Land Remote Sensing*. Ed. by Shunlin Liang. System, Modeling, Inversion and Application. Dordrecht: Springer Science & Business Media (cit. on p. 3).
- Lillo-Saavedra, Mario, Consuelo Gonzalo, and Octavio Lagos (2011). "Toward reduction of artifacts in fused images". In: *International Journal of Applied Earth Observation and Geoinformation* 13.3, pp. 368–375 (cit. on pp. 21, 93).
- Liu, Jianguo, Elizabeth Pattey, et al. (2010). "Estimating crop stresses, aboveground dry biomass and yield of corn using multi-temporal optical data combined with a radiation use efficiency model". In: *Remote Sensing of Environment* 114.6, pp. 1167–1177 (cit. on pp. 4, 36, 37).
- Liu, Ming-Yu, O Tuzel, S Ramalingam, and R Chellappa (2011). "Entropy rate superpixel segmentation". In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, pp. 2097–2104 (cit. on pp. 38, 39).
- Liu, Yong, Ling Bian, et al. (2012). "Discrepancy measures for selecting optimal combination of parameter values in object-based image analysis". In: *ISPRS journal of photogrammetry and remote sensing* 68, pp. 144–156 (cit. on p. 37).
- Liu, Yongxue, Manchun Li, Liang Mao, Feifei Xu, and Shuo Huang (2006). "Review of remotely sensed imagery classification patterns based on object-oriented image analysis". In: *Chinese Geographical Science* 16.3, pp. 282–288 (cit. on p. 34).
- Loh, Wei-Yin (2011). "Classification and regression trees". In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 1.1, pp. 14–23 (cit. on p. 27).
- Lowder, Sarah K., Jakob Skoet, and Terri Raney (2016). "The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide". In: *World Development*. In Press, Corrected Proof (cit. on p. 77).
- Lucchi, Aurélien, Kevin Smith, Radhakrishna Achanta, Graham Knott, and Pascal Fua (2012). "Supervoxel-based segmentation of mitochondria in em image stacks with learned shape features". In: *IEEE transactions on medical imaging* 31.2, pp. 474–486 (cit. on p. 95).
- Lunetta, Ross S, Joseph F Knight, Jayantha Ediriwickrema, John G Lyon, and L Dorsey Worthy (2006). "Land-cover change detection using multi-temporal MODIS NDVI data". In: *Remote sensing of environment* 105.2, pp. 142–154 (cit. on p. 91).
- Maccherone, Brandon and Shannell Frazier (2010). *MODIS: Specifications*. <http://modis.gsfc.nasa.gov/about/specifications.php>. [Online; accessed 20-February-2016] (cit. on p. 4).
- Mallinis, Georgios, Nikos Koutsias, Maria Tsakiri-Strati, and Michael Karteris (2008). "Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site". In: *{ISPRS} Journal of Photogrammetry and Remote Sensing* 63.2, pp. 237–250 (cit. on p. 30).
- Mandrake, Lukas, David R Thompson, Martha Gilmore, Rebecca Castano, and Eldar ZN Dobrea (2010). "Automated Neutral Region selection using superpixels". In: *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2010 2nd Workshop on*. IEEE, pp. 1–4 (cit. on pp. 110, 111).

- Marpu, Prashanth Reddy (2009). "Geographic Object-based Image Analysis". PhD thesis. Freiberg, Saxony, Germany: Technische Universität Bergakademie Freiberg (cit. on pp. 22, 24, 25, 28, 29).
- Masi, Giuseppe, Raffaele Gaetano, Giovanni Poggi, and Giuseppe Scarpa (2015). "Superpixel-based segmentation of remote sensing images through correlation clustering". In: *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*. IEEE, pp. 1028–1031 (cit. on pp. 110, 111).
- Maxwell, AE, TA Warner, MP Strager, JF Conley, and AL Sharp (2015). "Assessing machine-learning algorithms and image-and lidar-derived variables for GEOBIA classification of mining and mine reclamation". In: *International Journal of Remote Sensing* 36.4, pp. 954–978 (cit. on p. 30).
- Maxwell, T. and Y. Zhang (2005). "A fuzzy logic approach to optimization of segmentation of object-oriented classification". In: *Proceedings of SPIE 50th Annual Meeting - Optics & Photonics*. Vol. 5909, pp. 1–11 (cit. on p. 52).
- McGarigal, Kevin, Sam A Cushman, Maile C Neel, and Eduard Ene (2002). "FRAGSTATS: spatial pattern analysis program for categorical maps". In: (cit. on p. 68).
- Meinel, G and M Neubert (2004). "A comparison of segmentation programs for high resolution remote sensing data". In: *International Archives of Photogrammetry and Remote Sensing* 35.Part B, pp. 1097–1105 (cit. on p. 14).
- Miller, Holly M, Leslie Richardson, Stephen R Koontz, John Loomis, and Lynne Koontz (2013). "Users, uses, and value of Landsat satellite imagery—Results from the 2012 survey of users". In: *US Geological Survey Open-File Report* 1269, p. 51 (cit. on p. 4).
- Mirón Pérez, Jerónimo (2005). *Cadastre and the Reform of European Union's Common Agricultural Policy. Implementation of the SIGPAC*. http://www.catastro.meh.es/documentos/publicaciones/ct/ct54/01-catastro_54_ing.pdf (cit. on p. 77).
- Moore, A. P., S. J. D. Prince, J. Warrell, U. Mohammed, and G. Jones (2008). "Superpixel lattices". In: *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pp. 1–8 (cit. on p. 39).
- Mori, Greg (2005). "Guiding model search using segmentation". In: *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*. Vol. 2. IEEE, pp. 1417–1423 (cit. on p. 39).
- Mueller, Marina, Karl Segl, and Hermann Kaufmann (2004). "Edge- and region-based segmentation technique for the extraction of large, man-made objects in high-resolution satellite imagery". In: *Pattern Recognition* 37.8, pp. 1619–1628 (cit. on p. 78).
- Mulla, David J (2013). "Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps". In: *Special Issue: Sensing Technologies for Sustainable Agriculture* 114.4, pp. 358–371 (cit. on pp. 4–6).
- Myburgh, Gerhard (2012). "The impact of training set size and feature dimensionality on supervised object-based classification: a comparison of three classifiers". PhD thesis. Stellenbosch: Stellenbosch University (cit. on p. 37).
- Myint, S W, P Gober, A Brazel, S Grossman-Clarke, and Q Weng (2011). "Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery". In: *Remote Sensing of Environment* 115.5, pp. 1145–1161 (cit. on p. 51).
- National Research Council (1980). *Need for a Multipurpose Cadastre*. National Academy Press, Washington, DC (cit. on p. 77).
- (2007). *National Land Parcel Data: A Vision for the Future*. The National Academies Press, Washington, DC (cit. on p. 77).

- Neubert, P and P Protzel (2012). "Superpixel Benchmark and Comparison". In: *Proc. Forum Bildverarbeitung* (cit. on pp. 13, 38, 39).
- Nielsen, A. A. and J. S. Vestergaard (2013). "A kernel version of multivariate alteration detection". In: *2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS*, pp. 3451–3454 (cit. on p. 91).
- Nussbaum, Sven and Gunter Menz (2008). "Object-based image analysis and treaty verification". In: *Object-Based Image Analysis and Treaty Verification*, Edited by S. Nussbaum and G. Menz. Berlin: Springer, 2008. ISBN: 978-1-4020-6960-4 1 (cit. on p. 25).
- O'Connell, Jerome, Ute Bradter, and Tim G Benton (2015). "Wide-area mapping of small-scale features in agricultural landscapes using airborne remote sensing". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 109, pp. 165–177 (cit. on p. 35).
- Odeh, Inakwu OA and Alex B McBratney (2000). "Using AVHRR images for spatial prediction of clay content in the lower Namoi Valley of eastern Australia". In: *Geoderma* 97.3, pp. 237–254 (cit. on p. 4).
- Otsu, Nobuyuki (1975). "A threshold selection method from gray-level histograms". In: *Automatica* 11.285-296, pp. 23–27 (cit. on p. 22).
- Özdoğan, Mutlu (2015). "Image Classification Methods in Land Cover and Land Use". In: *Remotely Sensed Data Characterization, Classification, and Accuracies*. CRC Press, pp. 231–246 (cit. on pp. 15, 26).
- Ozelkan, Emre, Gang Chen, and Burak Berk Ustundag (2016). "Multiscale object-based drought monitoring and comparison in rainfed and irrigated agriculture from Landsat 8 OLI imagery". In: *International Journal of Applied Earth Observation and Geoinformation* 44, pp. 159–170 (cit. on p. 35).
- Ozturk, Safak, Yunus Emre Esin, and Yusuf Artan (2015). "Object detection in rural areas using hyperspectral imaging". In: *SPIE Remote Sensing*. International Society for Optics and Photonics, pp. 96432M–96432M (cit. on pp. 110, 111).
- Pal, M (2005). "Random forest classifier for remote sensing classification". In: *International Journal of Remote Sensing* 26.1, pp. 217–222 (cit. on p. 32).
- Pal, Nikhil R and Sankar K Pal (1993). "A review on image segmentation techniques". In: *Pattern Recognition* 26.9, pp. 1277–1294 (cit. on p. 24).
- Pedley, Mike I and Paul J Curran (1991). "Per-field classification: an example using SPOT HRV imagery". In: *Remote Sensing* 12.11, pp. 2181–2192 (cit. on p. 34).
- Peña, José M, Pedro A Gutiérrez, et al. (2014). "Object-based image classification of summer crops with machine learning methods". In: *Remote Sensing* 6.6, pp. 5019–5041 (cit. on p. 35).
- Peña-Barragán, José M, Moffatt K Ngugi, Richard E Plant, and Johan Six (2011). "Object-based crop identification using multiple vegetation indices, textural features and crop phenology". In: *Remote Sensing of Environment* 115.6, pp. 1301–1316 (cit. on p. 35).
- Petrou, ZI, C Tarantino, M Adamo, P Blonda, and M Petrou (2012). "Estimation of vegetation height through satellite image texture analysis". In: *Proceedings of the International archives of the photogrammetry, remote sensing and spatial information sciences, XXII ISPRS Congress, Melbourne*. Vol. 25, pp. 321–326 (cit. on p. 67).
- Priya, Tanu, Saurabh Prasad, and Hao Wu (2015). "Superpixels for spatially reinforced bayesian classification of hyperspectral images". In: *Geoscience and Remote Sensing Letters, IEEE* 12.5, pp. 1071–1075 (cit. on p. 38).

- Quarmby, NA and JL Cushnie (1989). "Monitoring urban land cover changes at the urban fringe from SPOT HRV imagery in south-east England". In: *International Journal of Remote Sensing* 10.6, pp. 953–963 (cit. on p. 91).
- Rembold, Felix, Clement Atzberger, Igor Savin, and Oscar Rojas (2013). "Using Low Resolution Satellite Imagery for Yield Prediction and Yield Anomaly Detection". In: *Remote Sensing* 5.4, pp. 1704–1733 (cit. on p. 5).
- Ren, X and J Malik (2003). "Learning a classification model for segmentation". In: *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. IEEE, 10–17 vol.1 (cit. on pp. 8, 13, 38, 39).
- Rodriguez-Galiano, Victor Francisco, Bardan Ghimire, John Rogan, Mario Chica-Olmo, and Juan Pedro Rigol-Sanchez (2012). "An assessment of the effectiveness of a random forest classifier for land-cover classification". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 67, pp. 93–104 (cit. on p. 31).
- Roscher, Ribana and Bjorn Waske (2014). "Superpixel-based classification of hyperspectral data using sparse representation and conditional random fields". In: *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*. IEEE, pp. 3674–3677 (cit. on pp. 110, 111).
- Rouse Jr., J. W., R. H. Haas, J. A. Schell, and D. W. Deering (1974). "Monitoring Vegetation Systems in the Great Plains with Ert's". In: *NASA Special Publication* 351, p. 309 (cit. on p. 67).
- Roy, David P, MA Wulder, et al. (2014). "Landsat-8: Science and product vision for terrestrial global change research". In: *Remote Sensing of Environment* 145, pp. 154–172 (cit. on p. 4).
- Ru, Hui, Pingping Huang, Xun Sun, and Yan Liu (2015). "Detecting changes in high resolution remote sensing images using superpixels". In: *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*. IEEE, pp. 1682–1685 (cit. on pp. 110, 111).
- Ru, Hui, Xiangli Yang, Dongqing Peng, and Pingping Huang (2014). "Superparsing based change detection in high resolution remote sensing imagery". In: *Signal Processing (ICSP), 2014 12th International Conference on*. IEEE, pp. 996–999 (cit. on pp. 110, 111).
- Rubner, Yossi, Carlo Tomasi, and Leonidas J Guibas (2000). "The earth mover's distance as a metric for image retrieval". In: *International journal of computer vision* 40.2, pp. 99–121 (cit. on p. 83).
- Ruiz, L A, J A Recio, A Fernández-Sarriá, and T Hermosilla (2011). "A feature extraction software tool for agricultural object-based image analysis". In: *Computers and Electronics in Agriculture* 76.2, pp. 284–296 (cit. on pp. 18, 34).
- Ruiz, Luis A, Jorge A Recio, Txomin Hermosilla, and A F Sarria (2009). "Identification of agricultural and land cover database changes using object-oriented classification techniques". In: *33rd International Symposium on Remote Sensing of Environment, May*, pp. 4–8 (cit. on p. 18).
- Schick, A and R Stiefelhagen (2011). "Evaluating image segments by applying the description length to sets of superpixels". In: *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*. IEEE, pp. 1394–1401 (cit. on p. 38).
- Schick, Alexander, Mika Fischer, and Rainer Stiefelhagen (2014). "An evaluation of the compactness of superpixels". In: *Pattern Recognition Letters* 43, pp. 71–80 (cit. on pp. 38, 39).
- Schiewe, J (2002). "Segmentation of high-resolution remotely sensed data-concepts, applications and problems". In: *International Archives of Photogrammetry Remote Sensing And Spatial Information Sciences* 34.4, pp. 380–385 (cit. on pp. 14, 21, 22, 24, 47).

- Schöpfer, E, S Lang, and F Albrecht (2008). "Object-fate analysis-spatial relationships for the assessment of object transition and correspondence". In: *Object-based image analysis*. Springer, pp. 785–801 (cit. on p. 19).
- Schöpfer, Elisabeth, Stefan Lang, and Josef Strobl (2010). "Segmentation and object-based image analysis". In: *Remote Sensing of Urban and Suburban Areas*. Springer, pp. 181–192 (cit. on pp. 8, 13, 14).
- Schowengerdt, Robert A (2006). *Remote sensing: models and methods for image processing*. Academic press (cit. on pp. 20, 21, 23, 27, 29, 30).
- Seelan, Santhosh K, Soizik Laguet, Grant M Casady, and George A Seielstad (2003). "Remote sensing applications for precision agriculture: A learning community approach". In: *Remote Sensing of Environment* 88.1, pp. 157–169 (cit. on p. 3).
- Seiffert, C, T M Khoshgoftaar, J Van Hulse, and A Napolitano (2010). "RUSBoost: A Hybrid Approach to Alleviating Class Imbalance". In: *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* 40.1, pp. 185–197 (cit. on pp. 32, 69).
- Sethi, Manu, Yupeng Yan, Anand Rangarajan, Ranga Raju Vatsavai, and Sanjay Ranka (2015). "Scalable Machine Learning Approaches for Neighborhood Classification Using Very High Resolution Remote Sensing Imagery". In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, pp. 2069–2078 (cit. on pp. 110, 111).
- Sethi, Manu, Yupeng Yan, Anand Rangarajan, Ranga Raju Vatsavai, and Sanjay Ranka (2014). "An efficient computational framework for labeling large scale spatiotemporal remote sensing datasets". In: *Contemporary Computing (IC3), 2014 Seventh International Conference on*. IEEE, pp. 635–640 (cit. on pp. 110, 111).
- Sezgin, Mehmet et al. (2004). "Survey over image thresholding techniques and quantitative performance evaluation". In: *Journal of Electronic imaging* 13.1, pp. 146–168 (cit. on p. 22).
- Shi, Jianbo and Jitendra Malik (2000). "Normalized cuts and image segmentation". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 22.8, pp. 888–905 (cit. on pp. 24, 39).
- Shi, Xiaoliang, Yingying Xu, Guixu Zhang, and Chaomin Shen (2015). "Change Detection Using L₀ Smoothing and Superpixel Techniques". In: *International Conference on Knowledge Science, Engineering and Management*. Springer, pp. 600–611 (cit. on pp. 110, 111).
- Shi, Yun, Shun-ping Ji, et al. (2016). "Framework of SAGI Agriculture Remote Sensing and Its Perspectives in Supporting National Food Security". In: *Journal of Integrative Agriculture* 13.7, pp. 1443–1450 (cit. on p. 5).
- Shih, Frank Y. and Shouxian Cheng (2005). "Automatic seeded region growing for color image segmentation". In: *Image and Vision Computing* 23.10, pp. 877–886 (cit. on p. 23).
- Singh, Ashbindu (1989). "Review article digital change detection techniques using remotely-sensed data". In: *International journal of remote sensing* 10.6, pp. 989–1003 (cit. on p. 91).
- Siva, Parthipan and Alexander Wong (2014). "URC: Unsupervised regional clustering of remote sensing imagery". In: *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*. IEEE, pp. 4938–4941 (cit. on pp. 110, 111).
- Soille, Pierre (1999). *Morphological image analysis: principles and applications*. Springer-Verlag New York, Inc. (cit. on p. 66).
- Song, M., D. L. Civco, and J. D. Hurd (2005). "A competitive pixel-object approach for land cover classification". In: *International Journal of Remote Sensing* 26.22, pp. 4981–4997 (cit. on p. 37).

- Spann, Michael and Roland Wilson (1985). "A quad-tree approach to image segmentation which combines statistical and spatial information". In: *Pattern Recognition* 18.3, pp. 257–269 (cit. on p. 24).
- Stathaki, Tania (2011). *Image fusion: algorithms and applications*. Academic Press (cit. on p. 21).
- Stumpf, André and Norman Kerle (2011). "Object-oriented mapping of landslides using Random Forests". In: *Remote Sensing of Environment* 115.10, pp. 2564–2577 (cit. on p. 30).
- Stutz, David (2015). "Superpixel Segmentation: An Evaluation". In: *Pattern Recognition*. Springer, pp. 555–562 (cit. on p. 38).
- Tang, Jiliang, Salem Alelyani, and Huan Liu (2014). "Feature selection for classification: A review". In: *Data Classification: Algorithms and Applications*. CRC Press, pp. 37–60 (cit. on p. 28).
- Tansey, Kevin, Ian Chambers, Andrew Anstee, Anthony Denniss, and Alistair Lamb (2009). "Object-oriented classification of very high resolution airborne imagery for the extraction of hedgerows and field margin cover in agricultural areas". In: *Applied geography* 29.2, pp. 145–157 (cit. on p. 35).
- Thompson, D R, L Mandrake, M S Gilmore, and R Castano (2010). "Superpixel Endmember Detection". In: *Geoscience and Remote Sensing, IEEE Transactions on* 48.11, pp. 4023–4033 (cit. on p. 38).
- Tiwari, Poonam S, Hina Pande, Minakshi Kumar, and Vinay Kumar Dadhwal (2009). "Potential of IRS P-6 LISS IV for agriculture field boundary delineation". In: *Journal of Applied Remote Sensing* 3.1, pp. 033528–033528 (cit. on p. 78).
- Tomowski, Daniel, Manfred Ehlers, and Sascha Klonus (2011). "Colour and texture based change detection for urban disaster analysis". In: *2011 Joint Urban Remote Sensing Event*. IEEE, pp. 329–332 (cit. on p. 92).
- Treitz, Paul and Philip Howarth (2000). "High spatial resolution remote sensing data for forest ecosystem classification: An examination of spatial scale". In: *Remote sensing of Environment* 72.3, pp. 268–289 (cit. on p. 52).
- Trias-Sanz, Roger, Georges Stamon, and Jean Louchet (2008). "Using colour, texture, and hierarchial segmentation for high-resolution remote sensing". In: *ISPRS Journal of Photogrammetry and remote sensing* 63.2, pp. 156–168 (cit. on p. 52).
- Turker, Mustafa and Emre Hamit Kok (2013). "Field-based sub-boundary extraction from remote sensing imagery using perceptual grouping". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 79, pp. 106–121 (cit. on p. 78).
- Turker, Mustafa and Asli Ozdarici (2011). "Field-based crop classification using SPOT4, SPOT5, IKONOS and QuickBird imagery for agricultural areas: a comparison study". In: *International Journal of Remote Sensing* 32.24, pp. 9735–9768 (cit. on p. 34).
- Uijlings, Jasper RR, Koen EA van de Sande, Theo Gevers, and Arnold WM Smeulders (2013). "Selective search for object recognition". In: *International journal of computer vision* 104.2, pp. 154–171 (cit. on p. 83).
- Ursani, Ahsan Ahmad, Kidiyo Kpalma, Camille CD Lelong, and Joseph Ronsin (2012). "Fusion of textural and spectral information for tree crop and other agricultural cover mapping with very-high resolution satellite images". In: *IEEE Journal of selected topics in applied earth observations and remote sensing* 5.1, pp. 225–235 (cit. on p. 36).
- Van den Bergh, Michael, Xavier Boix, Gemma Roig, Benjamin de Capitani, and Luc J Van Gool (2012). "SEEDS: Superpixels Extracted via Energy-Driven Sampling". In: *ECCV (7)*, pp. 13–26 (cit. on p. 38).

- van Der Molen, Paul (2002). *The use of the Cadastre among the Members States: Property rights, land registration and Cadastre in the European Union*. <http://www.catastro.meh.es/documentos/publicaciones/ct/ct45/02ingles.pdf> (cit. on p. 77).
- Vapnik, Vladimir N. (1998). *Statistical Learning Theory*. Wiley-Interscience (cit. on pp. 27, 31).
- Vargas, JE, AX Falcao, et al. (2015). "Contextual superpixel description for remote sensing image classification". In: *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*. IEEE, pp. 1132–1135 (cit. on pp. 110, 111).
- Vargas, John E, Priscila Saito, Alexandre X Falcao, Pedro J De Rezende, and Jefersson A Dos Santos (2014). "Superpixel-based interactive classification of very high resolution images". In: *Graphics, Patterns and Images (SIBGRAPI), 2014 27th SIBGRAPI Conference on*. IEEE, pp. 173–179 (cit. on pp. 110, 111).
- Vedaldi, Andrea and Stefano Soatto (2008). "Quick shift and kernel methods for mode seeking". In: *Computer vision—ECCV 2008*. Springer, pp. 705–718 (cit. on p. 39).
- Veljanovski, Tatjana, Urša Kanjir, and Krištof Oštir (2011). "Object-based image analysis of remote sensing data". In: *Geodetski vestnik* 55.4, pp. 678–688 (cit. on pp. 18, 19).
- Verbesselt, Jan, Rob Hyndman, Glenn Newnham, and Darius Culvenor (2010). "Detecting trend and seasonal changes in satellite image time series". In: *Remote sensing of Environment* 114.1, pp. 106–115 (cit. on p. 4).
- Vieira, Matheus Alves, Antonio Roberto Formaggio, et al. (2012). "Object based image analysis and data mining applied to a remotely sensed Landsat time-series to map sugarcane over large areas". In: *Remote Sensing of Environment* 123, pp. 553–562 (cit. on pp. 19, 35).
- Vivone, G., L. Alparone, et al. (2015). "A Critical Comparison Among Pansharpening Algorithms". In: *IEEE Transactions on Geoscience and Remote Sensing* 53.5, pp. 2565–2586 (cit. on p. 21).
- Walter, Volker (2004). "Object-based classification of remote sensing data for change detection". In: *ISPRS Journal of photogrammetry and remote sensing* 58.3, pp. 225–238 (cit. on p. 18).
- Wardlow, Brian D and Stephen L Egbert (2008). "Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the US Central Great Plains". In: *Remote sensing of environment* 112.3, pp. 1096–1116 (cit. on p. 4).
- Wassenaar, T, J-M Robbez-Masson, P Andrieux, and F Baret (2002). "Vineyard identification and description of spatial crop structure by per-field frequency analysis". In: *International Journal of Remote Sensing* 23.17, pp. 3311–3325 (cit. on p. 34).
- Wei, Liu, Shaozi Li, et al. (2013). "Spectral-spatial classification of hyperspectral imagery based on Random Forests". In: *Proceedings of the Fifth International Conference on Internet Multimedia Computing and Service*. ACM, pp. 163–168 (cit. on pp. 110, 111).
- Weikersdorfer, David, David Gossow, and Michael Beetz (2012). "Depth-adaptive superpixels". In: *Pattern Recognition (ICPR), 2012 21st International Conference on*. IEEE, pp. 2087–2090 (cit. on p. 39).
- Weisstener, Christof J, Celia García-Feced, and Maria Luisa Paracchini (2016). "A new view on EU agricultural landscapes: Quantifying patchiness to assess farmland heterogeneity". In: *Ecological Indicators* 61, pp. 317–327 (cit. on p. 35).
- Weng, Qihao et al. (2010). *Remote sensing and GIS integration: theories, methods, and applications*. McGraw-Hill New York (cit. on p. 34).

- White, Emma V and David P Roy (2015). "A contemporary decennial examination of changing agricultural field sizes using Landsat time series data". In: *Geo: Geography and Environment* 2.1, pp. 33–54 (cit. on p. 35).
- Wiegand, CL, AJ Richardson, and ET Kanemasu (1979). "Leaf area index estimates for wheat from Landsat and their implications for evapotranspiration and crop modeling". In: *Agronomy Journal* 71.2, pp. 336–342 (cit. on p. 4).
- Woodcock, Curtis E and Alan H Strahler (1987). "The factor of scale in remote sensing". In: *Remote Sensing of Environment* 21.3, pp. 311–332 (cit. on pp. 51, 52).
- Wu, Jianguo (2016). *Ecology: Patch dynamics*. <http://global.britannica.com/science/patch-dynamics>. [Online; accessed 25-March-2016] (cit. on p. 17).
- Wu, Zhaocong, Zhongwen Hu, and Qian Fan (2012a). "Superpixel-based unsupervised change detection using multi-dimensional change vector analysis and SVM-based classification". In: *Proceedings of the Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXII ISPRS Congress, Melbourne, Australia*. Vol. 25 (cit. on p. 92).
- (2012b). "Superpixel-based unsupervised change detection using multi-dimensional change vector analysis and SVM-based classification". In: *Proceedings of the Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXII ISPRS Congress, Melbourne, Australia*. Vol. 25 (cit. on pp. 110, 111).
- Wu, Zhengzhe, Ville Heikkinen, Markku Hauta-Kasari, Jussi Parkkinen, and Timo Tokola (2014). "ALS data based forest stand delineation with a coarse-to-fine segmentation approach". In: *Image and Signal Processing (CISP), 2014 7th International Congress on*. IEEE, pp. 547–552 (cit. on pp. 110, 111).
- Xie, Lei, Hong Zhang, Chao Wang, Meng Liu, and Bo Zhang (2015). "Superpixel-based PolSAR images change detection". In: *Synthetic Aperture Radar (AP SAR), 2015 IEEE 5th Asia-Pacific Conference on*. IEEE, pp. 792–796 (cit. on pp. 110, 111).
- Xu, Chi, Yutong Li, et al. (2012). "Evaluating the difference between the normalized difference vegetation index and net primary productivity as the indicators of vegetation vigor assessment at landscape scale". In: *Environmental monitoring and assessment* 184.3, pp. 1275–1286 (cit. on p. 4).
- Xu, Q., B. Li, Y. Zhang, and L. Ding (2014). "High-Fidelity Component Substitution Pansharpening by the Fitting of Substitution Data". In: *IEEE Transactions on Geoscience and Remote Sensing* 52.11, pp. 7380–7392 (cit. on p. 21).
- Yan, Gao, J-F Mas, BHP Maathuis, Zhang Xiangmin, and PM Van Dijk (2006). "Comparison of pixel-based and object-oriented image classification approaches—a case study in a coal fire area, Wuda, Inner Mongolia, China". In: *International Journal of Remote Sensing* 27.18, pp. 4039–4055 (cit. on p. 51).
- Yang, Allen Y, John Wright, Yi Ma, and S Shankar Sastry (2008). "Unsupervised segmentation of natural images via lossy data compression". In: *Computer Vision and Image Understanding* 110.2, pp. 212–225 (cit. on p. 88).
- Yang, Chenghai, James H Everitt, Qian Du, Bin Luo, and Jocelyn Chanussot (2013). "Using high-resolution airborne and satellite imagery to assess crop growth and yield variability for precision agriculture". In: *Proceedings of the IEEE* 101.3, pp. 582–592 (cit. on pp. 5–7).
- Yang, Jian, Yuhong He, and Qihao Weng (2015). "An automated method to parameterize segmentation scale by enhancing intrasegment homogeneity and intersegment heterogeneity". In: *IEEE Geoscience and Remote Sensing Letters* 12.6, pp. 1282–1286 (cit. on p. 37).

- Yuan, Fei, Kali E Sawaya, Brian C Loeffelholz, and Marvin E Bauer (2005). "Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal Landsat remote sensing". In: *Remote sensing of Environment* 98.2, pp. 317–328 (cit. on p. 91).
- Yuan, Jiangye and Anil M Cheriyaat (2013). "Road segmentation in aerial images by exploiting road vector data". In: *Computing for Geospatial Research and Application (COM. Geo), 2013 Fourth International Conference on*. IEEE, pp. 16–23 (cit. on pp. 110, 111).
- Zhan, Qingming (2003). "A hierarchical object-based approach for urban land-use classification from remote sensing data". PhD thesis. Wageningen, Netherlands: Wagening University (cit. on p. 18).
- Zhan, Tianming, Yang Xu, Le Sun, Zebin Wu, and Yongzhao Zhan (2015). "Hyperspectral image classification using multilayer superpixel graph and loopy belief propagation". In: *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, pp. 1690–1693 (cit. on pp. 110, 111).
- Zhang, Caiyun (2015). "Applying data fusion techniques for benthic habitat mapping and monitoring in a coral reef ecosystem". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 104, pp. 213–223 (cit. on p. 14).
- Zhang, Yu-Jin (2006). "An overview of image and video segmentation in the last 40 years". In: *Advances in Image and Video Segmentation*, pp. 1–15 (cit. on p. 22).
- (2009). "Image Segmentation in the Last 40 Years". In: *Encyclopedia of Information Science and Technology*, pp. 1818–1823 (cit. on p. 22).
- Zhang, Qi, Xiaoyong Shen, Li Xu, and Jiaya Jia (2014). "Rolling guidance filter". In: *European Conference on Computer Vision*. Springer, pp. 815–830 (cit. on pp. 20, 82, 94).
- Zhang, Xuewen, Selene E Chew, Zhenlin Xu, and Nathan D Cahill (2015). "SLIC superpixels for efficient graph-based dimensionality reduction of hyperspectral imagery". In: *SPIE Defense + Security*. International Society for Optics and Photonics, pp. 947209–947209 (cit. on pp. 110, 111).
- Zhou, Weiqi and Austin Troy (2009). "Development of an object-based framework for classifying and inventorying human-dominated forest ecosystems". In: *International Journal of Remote Sensing* 30.23, pp. 6343–6360 (cit. on p. 52).
- Ziou, Djemel and Salvatore Tabbone (1998). "Edge detection techniques—an overview". In: *Pattern Recognition And Image Analysis C/C Of Raspoznavaniye Obrazov I Analiz Izobrazhenii* 8, pp. 537–559 (cit. on p. 66).

List of Figures

1.1	Farming system maps of the developing regions of the world.	5
1.2	An overview of agricultural plots using different image resolutions.	6
2.1	Objects and sensor resolutions.	13
2.2	Schema of image interpretation process.	15
2.3	Scale representation.	17
2.4	Flowchart illustrating the main stages that are part of the OBIA approach.	19
2.5	Scheme of a hierarchical segmentation.	24
2.6	General scheme of the analysis using supervised methods.	28
2.7	Agricultural cadastre provided the CIREN, Chile.	35
2.8	Segmentation is an ill-posed problem.	36
2.9	Under-segmentation and over-segmentation errors.	37
2.10	Search area of k-means and SLIC algorithms.	41
2.11	Example of an RGB image segmented by SLIC.	41
2.12	Publications per year using the concept of superpixels.	42
2.13	Three scales of analysis of agricultural landscapes.	44
3.1	Example of applying superpixel method to an area of the multi-spectral delineated by vector data.	49
3.2	Example of applying post-processing techniques.	50
4.1	The study area is located in Biobio Region, Chile.	54
4.2	Color composition (near infrared-green-blue) of the analyzed scene.	54
4.3	Hierarchical SLIC: Segments hierarchy.	55
4.4	General overview of the evaluation process.	56
4.5	Optimal scales for each parent superpixel.	59
4.6	Optimal scales spatially distributed.	60
4.7	Boundaries of each segment according to the optimal scale.	60
4.8	Segmented image through the dyadic approach.	61
5.1	The study area is located in Valparaiso Region, Chile.	64
5.2	Real color composition of the WorldView-2 image.	65
5.3	Segmentation process of a multi-spectral image.	66
5.4	Workflow of the analysis process.	68
5.5	Results of the edge-based and SLIC segmentations.	71
5.6	Superpixel map.	73
5.7	Map of the vegetated areas.	74
5.8	Combination of superpixel and parcel maps.	75
5.9	Final map generated with the disambiguation of regions.	75

6.1	Regions under study.	79
6.2	Study regions and reference plot maps.	80
6.3	Overview of the methodology based on machine learning.	80
6.4	Overview of the methodology based on consensus of segmentations.	82
6.5	Superpixel segmentation of two small areas.	84
6.6	Results of applying the approach based on machine learning in the images under study.	86
6.7	Superpixel segmentation of two small areas with different superpixel sizes. . . .	87
6.8	Consensus of segmentations.	87
6.9	Comparison between the edge map and the UCM.	88
6.10	Comparison between the edges of the ground-truth and the edges obtained by the approach based on the consensus of segmentations.	89
7.1	Multi-temporal multi-spectral images.	94
7.2	Overview of the temporal analysis of agricultural plots.	94
7.3	Temporal evolution of superpixels.	96
7.4	Superpixels generated for the different dates.	97
7.5	Map of spatio-temporal dynamics of variability.	97
7.6	Regions with a low-persistence variability.	98
7.7	Regions with a high-persistence variability.	98
A.1	Publications in conferences that use superpixels related to remote sensing area . .	110
A.2	Publications in JCR Jorunals that use superpixels related to remote sensing area.	111

List of Tables

1.1	Farming system description.	5
1.2	Suitability of satellite images for studying agricultural plots.	7
2.1	Definitions of certain terms related to OBIA of Remote Sensing data.	18
2.2	Overview of the superpixel algorithms.	39
4.1	Spectral and spatial characteristics of the multi-spectral image taken by the Pléiades satellite	53
4.2	K -vectors values for linear and dyadic approaches.	58
5.1	Spectral properties of the WorldView-2 image.	64
5.2	Class distribution of the available ground-truth examples.	65
5.3	Accuracy assessment of the classification at superpixel level.	72
5.4	Confusion matrix showing results of the classification at superpixel level.	72
5.5	Confusion matrix showing results of the classification of the final map.	76
6.1	Accuracy assessment of the classification in the validation set.	85
6.2	Comparison between approaches for delineating agricultural plots.	88

List of Algorithms

3.1	SLIC algorithm for remote sensing data.	48
4.1	Hierarchical SLIC	56
5.1	Edge-based segmentation algorithm.	67

Colophon

This thesis was typeset with $\text{\LaTeX}$ 2_ε. It uses the *Clean Thesis* style developed by Ricardo Langner. The design of the *Clean Thesis* style is inspired by user guide documents from Apple Inc.

Download the *Clean Thesis* style at <http://cleanthesis.der-ric.de/>.