Ca' Foscari
University
of Venice

PhD Degree

In computer science

Final Thesis

Hyperspectral based industrial inspection systems

Supervisor
Ch. Prof. Andrea Albarelli

Graduand
Marco Boschetti
Matricolation number 956373

Academic Year
2020 / 2021

Abstract

Industrial processes always require more the exact and precise knowledge of the material being processed, in order to refine, automate and optimize the production flow, to maximize the final product's value and quality, and to maximize the resource utilization in a sustainable way.

Several vendors produce and sell optical scanners made of a combination of cameras and light sources, and there is a push to enhance their performances by using more advanced technologies like faster cameras, higher resolution sensors or by sensing other informative wavelengths (infrared, UV, x-ray).

This PhD project is focused on the analysis and development of a hyperspectral measurement system, working in the visible, near infrared and short wave infrared range (VIS-NIR-SWIR, wavelengths spanning from 400nm to 1700nm). Such system has been validated in different real industrial cases, in particular for the sorting of wood boards (for example the classification of heart- and sap-wood in Eucalyptus hardwood) and of fruits (for example the estimation of dry matter in avocados and sugar content in oranges).

The work was divided in several phases: selection of main components and development of the hardware parts, development of the software pipeline for modelling the data and evaluate the results, selection of use cases with thorough testing both in laboratory and in-field, and finally the engineering of a final product to use in a real industrial environment. The result is the combination of novel hardware and software tools that has been adopted to create actual products developed by the company I work for.

Index

List of figure.

List of tables.

1 Introduction

The original idea of a project on hyperspectral imaging started from colleagues of our daughter company in Sweden in 2015, who showed me a complete solution made of a hyperspectral system and a closed sources software. The technology was promising, but the system was bulky, slow and too expensive to be practically used. Moreover, its working principle was a complete black box.

In addition I attended one trade show where were shown camera sensors based on InGaAs technology and sensitive to the SWIR wavelength range, a region where water has a high absorption of light, thus making it possible to these sensors to "see the water", which seemed useful to measure properties of organic materials, like wood and food.

These two facts seemed a good starting point for an internal development project for the company I work for, Microtec, and an interesting subject for an industrial PhD project.

1.1 Hyperspectral technology

The basic idea of hyperspectral imaging is to combining spectroscopy, the ability to measure the continuous spectrum of the light reflected by an object in order to study some of its properties, with imaging, the ability of doing this with a high spatial definition. In other words, this technology expands the perception of colours, well beyond the common three components of colour images (red, green, blue) to some hundreds. In this way, as moving from grayscale to colour images increase the ability of understanding a scene, the availability of many more components extends dramatically the possibility to measure particular properties of the objects to analyse.

In addition to this, the extension of the used wavelength ranges from the visible (as perceived by the human eye) to the near infrared and short wave infrared wavelength ranges, augment the discriminative power of such devices compared to what a person can do.

Figure 1.1 top shows a grayscale image, a colour image and a pseudo-colour image obtained by means of hyperspectral techniques of rice with some insects; while in the first two images the insects are not or just barely visible, in the third image they are clearly distinguishable.

Moreover, Figure 1.1 bottom shows a picture of water and sugar and salt, which appear respectively transparent or having the same white colour in the visible wavelength range, while are visible or pretty different if observed in the shortwave infrared range.

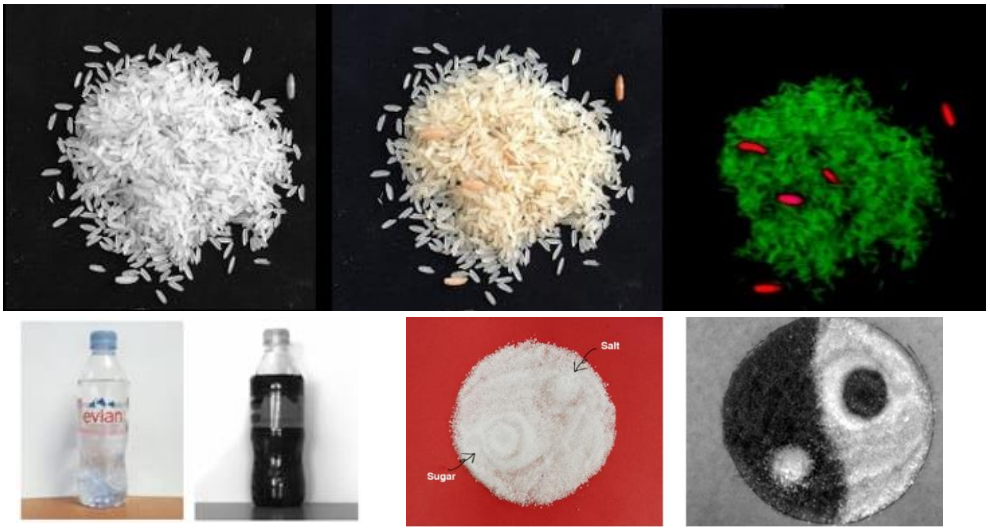


Figure 1.1. Top: Same scene as grayscale image, color image and processed hyperspectral image. Bottom: bottle of water and sugar and salt seen in the visible spectrum and in the short-wave infrared spectrum.

1.2 Application of hyperspectral technology in industrial applications

Microtec's mission is to be leader in the production of scanners for the measurement of quality in the wood processing and food industry. The performances of these scanners are linked with their ability to measure a continuous increasing number of properties, from purely aesthetic to structural, which can be visible from the external surface or can require the examination of the internal part of the object. Being an industrial process, the requirement is to analyse the totality of the production in a non-destructive manner, coping with the speed and duration of the industry.

One important concept in the realization of these scanners has been the fusion of data coming from different sensors, in order to have a richer signal to analyse. Figure 1.2 shows the concept of a *multisensor approach*, where multiple areas of the electromagnetic spectrum are used, starting with the visible range but using also the x-ray, microwave and radio wave portions; the figure shows this concept more in practice, where different images obtained from sensor sensitive to different wavelengths (visible, x-ray) and using different principles of measurement (measure of reflectivity, triangulation, fiber deviation) are superimposed to create a multi-dimensional image that is clearly much richer of information.

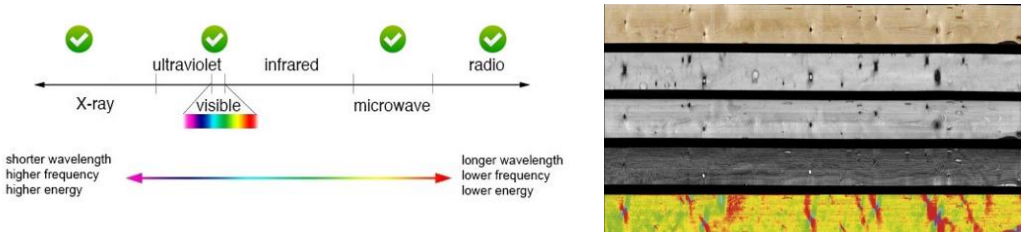


Figure 1.2. Microtec's multisensor approach.

In these terms, hyperspectral technology is an extension to this concept, since it considers:

- the use of a continuous spectrum compared to a few bands (like the three color components of color sensors);
- the use of a different range in the wavelength spectrum, comprehending the near-infrared and in particular the shortwave-infrared, namely the range from 700nm to 1700nm, which is novel for our company.

The utilization in industry processes imposes, in addition to the validation of the increased performances, other practical constraints related to speed and cost.

These sensors need to be fast, since the measuring speeds can reach the 15 meters/second in some wood board scanning application, with the requirement of a measurement every few millimetres.

These sensors cannot have a too high cost, since this would not allow the use in commercial application in some sectors, like the food industry, which are very price sensitive.

1.3 Structure of the thesis

The content of this thesis reflects my work on the internal developments at Microtec of the hyperspectral technology.

The project is built around four main activities:

- Acquisition device (hardware)
- Workflow from data to model (software and process)
- Applications
- Ecosystem around this project

The acquisition device is an important aspect, since the final goal of this project is to design new functional and robust measuring devices that must work in the harsh conditions of industrial processes. Since the existing devices were not fully satisfactory (in terms of acquisition speed, dimensions, cost, stability), a lot of work was put to evaluate the available technologies, select the best components and design and produce

prototypes. In particular I worked on the camera and on the illumination, covering different aspects concerning flexibility of use, acquisition speed, stability over long temporal period, possibility of combination with other existing sensors, and finally cost. This part was a continuous effort during the whole project, and it was propaedeutic to the final installations. Chapter 3 in the thesis covers these activities.

The workflow process is the combination of the methods used to extract useful information from the raw data, i.e. the software to calibrate and use the acquisition devices, to store, manage and label the acquired data, the algorithms to create and evaluate models and finally to deploy them in the final environment. The development of the workflow was a continuous effort during the project, like the work on hardware, receiving feedbacks and the push for improvement from the application use cases. Chapter 4 covers this part.

The applications are the concrete use cases where the developed technology and algorithms can be applied in a useful manner. The general applications were in the automatic sorting of wood, fruit and vegetable products, but in order to discover which use cases were effectively solvable by means of hyperspectral imaging and were commercially interesting application as well, it was necessary to combine the suggestions from the sales department in Microtec, with the needs of customers and with the advices of research partners. Many tests were done in order to quickly assess the feasibility of the applications or discard the dead ends, and the most promising ones were then further analysed, studied, and a few brought to concrete in field applications. These concrete cases were also the pulling force for improvements concerning the acquisition devices and the workflow software. The results of this part are presented in chapter 5.

Finally, I tried to expand this work outside the company Microtec as well, thus a substantial effort was put in the creation of additional projects and activities, in cooperation with universities and research centres as well, beyond the contacts within the University of Ca' Foscari. This component was useful to increase the exchange in ideas and knowhow, in terms of algorithms (for example with the University of Bolzano) and of knowledge on biological aspects (for example with the Fondazione Edmund Mach in Trento, the University of Bolzano and the Laimburg Research Centre in Bolzano). In addition we exchanged raw material (pieces of wood, fruits) and acquired data, in order to share the results and the facility for destructive and non-destructive measurements. A list of these additional projects is described in chapter 6, together with the concrete outcome of my work.

To summarize, the structure of this thesis is as follows:

Chapter 2 will present an overview of spectroscopy and hyperspectral techniques, in terms of technologies, algorithms and literature regarding its application.

Chapter 3 will show the work that has been done from the hardware point of view. In particular two key aspects will be described, the work that has been done on the development of two camera systems and of a lighting setup to overcome the limits of classic halogen bulbs.

In chapter 4 the workflow of the chosen hyperspectral techniques, consisting in the data collection, processing, model creation, and deployment will be examined.

Chapter 5 will show some of the most promising applications, in the wood and food (fruit) sector. I will present results for the classification of some wood properties and for the estimation of internal properties of avocado and citrus fruits. Two of these applications were first tested in laboratory conditions and then applied in real field production. Finally a few additional applications will be shortly presented.

Chapter 6 shows the different outcomes of this thesis, in terms of publications and patent (because of the industrial nature of the PhD, the amount of scientific publication is limited due to secrecy reasons; in addition some patents have been obtained or are in the process of evaluation), and in terms of other projects and cooperations that I started in connection with this subject.

Finally in chapter 7 some possible next steps and conclusions will be presented.

The work of this project has spawned over a period of about four years. I was in charge of the scientific direction of the entire project, the experiments, the methodology and hardware selection; I also worked on the development of the acquisition software, camera firmware and mechanic, modelling software and data evaluation. However, the results would have not been possible without the contribution of many colleagues who worked with me, thus the content of this thesis is based also on the work, internal technical reports and activities made in cooperation with Simone Faccini (sections 3.3, 5.1, 5.2, 5.3), Matteo Caffini (sections 3.3, 5.1, 5.2, 5.3), Ilario Chini (section 5.3), Andrea Gottardo (section 3.2), Andrea Azzalin (section 3.2), Alberto Celin (section 3.2), Johann Thaler (section 3.2), Andrea Ciresa (sections 6.1.2, 6.1.4).

2 State of the art

Hyperspectral imaging (HSI) is a recent technology that combines image processing and spectroscopy techniques to obtain both spatial and spectral information from an object.

In general terms the goal of hyperspectral imaging is to obtain the "spectrum" for "each pixel" in the image of a scene and to “study” it. The term hyperspectral means that we are dealing with narrow spectral bands over a continuous spectral range, in respect to multispectral, where only a few discrete bands are considered. Moreover, the term imaging means that we are considering spatial resolution in addition to spectral resolution, in respect to a spectrometer that measures the spectrum from an incoming fibre optic from a single spot. Finally, these data are analysed by means of spectroscopy techniques, which is the study of the interaction between matter and electromagnetic radiation, in particular the visible and near infrared range.

2.1 Imaging methods

Hyperspectral imaging is a type of spectral imaging technology that integrates imaging and spectroscopy to obtain 3D data cubes, called hypercubes, that contains 2-D spatial and 1-D spectral information from objects. It is related to another technology called multispectral imaging, the main difference being the number of wavelengths and spectral resolution of the acquired data: the boundary between these two techniques is generally set so that below 10-20 discrete wavelengths or wavebands is termed multispectral, while above this is termed hyperspectral [1].

In order to obtain the 3D hypercube there are four main configurations: point or whiskbroom scanning, line or push broom scanning, area or staring scanning and single shot, as depicted in Figure 2.1.

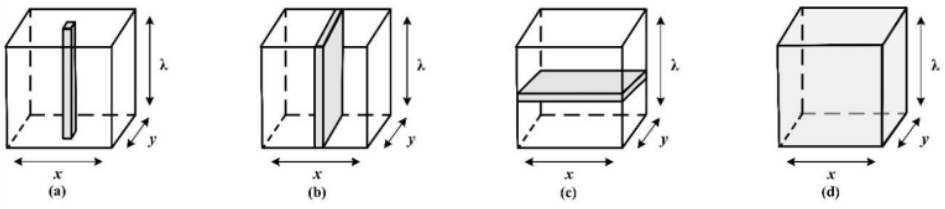


Figure 2.1. The four approaches to acquire a three-dimensional hyperspectral image cube: (a) point scanning, (b) line scanning, (c) area scanning and (d) single shot. X and Y represent the spatial dimensions, Lambda is the wavelength.

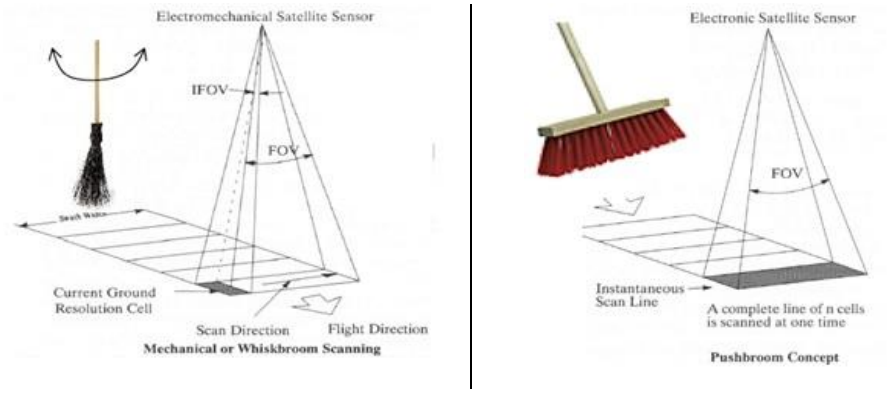


Figure 2.2. Left: basic principle of Whiskbroom scanning. Right: Pushbroom scanning.

2.1.1 Point scanning (Whiskbroom) imaging

With this method (Figure 2.2 left) a single pixel detector captures the full spectrum of a spot, and the scene or sample is scanned over the two spatial directions. This method was used in many orbital satellites, where the measuring spot sweeps in a direction perpendicular to the flight path, collecting one pixel at a time. For example, all the Landsat 1 to 7 satellites (launched between 1972 and 1999) used this kind of design.

The sensor used in this design can be a standard spectrometer, which can be therefore easily be used as hyperspectral device with the addition of a two-axis mechanical stage.

The main benefit of this technique is the ability to obtain high spectral resolution, but the spatial resolution can be low or the acquisition time can be very long.

2.1.2 Line scanning (Pushbroom) imaging

With this method (Figure 2.2 right) a line of spatial information with a full spectrum for each pixel is captured, and the scene or sample is scanned over one spatial direction. This technique is widely used in remote satellite sensing [2]; as an example, the OLI (Operational Land Imager) instrument on Landsat 8 launched in 2013 uses a push broom design.

This approach is well suited for conveyor belt systems in the production line for online inspection because of its good compromise between spatial and spectral resolution. This design is the most used configuration in the literature and for practical applications in industrial environment as well.

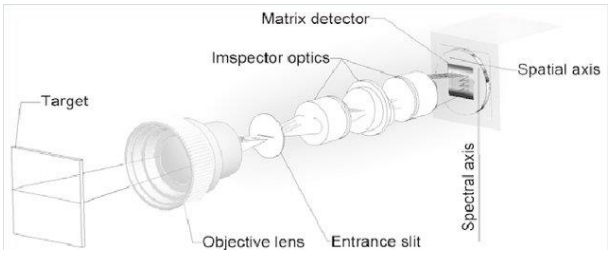


Figure 2.3. Structure of a spectrograph, an optical element commonly used in push broom imaging.

2.1.3 Area scanning (Staring) imaging

With this method, a frame of spatial "monochromatic" information is taken for every exposure, and the scan is performed over the spectral range. The spectral scan is performed either by exchanging a set of filters or by using a tuneable filter like LCTD (liquid crystal tuneable passband filter) as depicted in Figure 2.4 left.

This design can be used when the sample or scene is stationary during the measurement; when this is not possible, as in the case of aerial remote sensing, the different scans at different wavelengths can be realigned by using spatial features in each image.

A particular case of this design, that uses a combination of a standard industrial camera with a linear variable filter, has been implemented by the company Glana [3]. With this design, every column in a frame contains a different wavelength (from blue to NIR), and in order to get a full hypercube it is required to sweep the scene and to realign different images, as depicted in Figure 2.4 right.

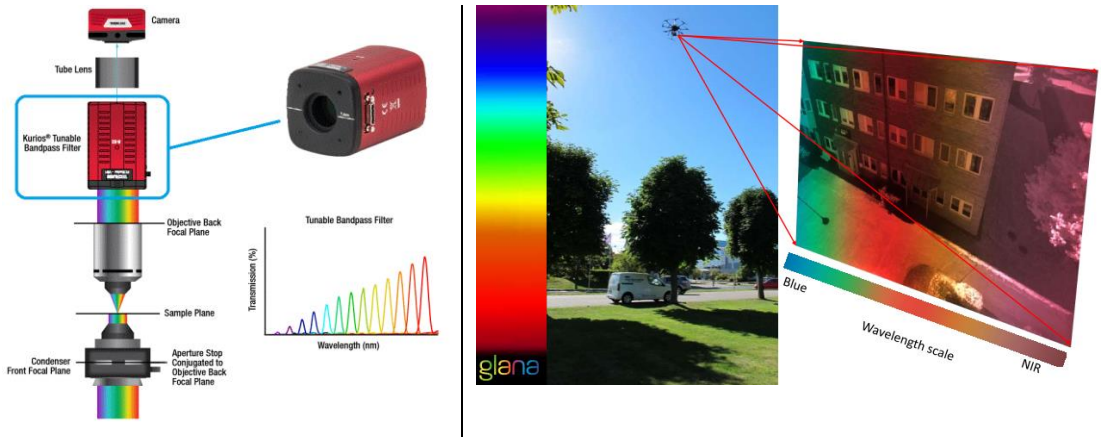


Figure 2.4. Right: configuration with tunable bandpass filter (Thorlabs). Right: configuration with linear variable filter (Glana).

2.1.4 Single shot (snapshot) imaging

This method uses a spatial detector able to capture with a single exposure the complete 3D image cube, without scanning in either spectral or spatial domain. Different techniques have been proposed to build such a device.

Compressive sensing exploits the paradigm that sparse signals can be under sampled without losing relevant information. A coded aperture is placed in front of or behind a prism (or a diffraction grating) via collimation in the optical path of the imaging system; the coded aperture is used to encode spectral signatures, which are later used to reconstruct the compressive input into a complete hyperspectral image. Traditionally these systems are large and expensive due to additional elements such as collimating optics and coded masks, making them bulky and hard to handle in practice. Takatani [4] proposes to use faced reflectors on which colour filters are attached; the key idea is based on the principle that each of multiple reflections on the filters has a different spectrum, which allows to observe multiple intensities through different spectra; the technique can be implemented either by a coupled mirror or a kaleidoscope geometry. Computed tomography imaging spectrometry (CTIS) [5] uses a diffraction grating to split incident light rays into a number of spectral projections on an image plane; since multiple sub-images need to be captured with a single sensor, the effective spatial resolution of the reconstructed images is drastically reduced. Bodkin [6] proposed a particular Hyperpixel Array (HPA) optical processor, a special grating whose design allows the entire data cube to be mapped onto the focal plane and to be read as a normal two dimensional image.

However, none of these instruments have seen yet widely adoption in commercial and industrial application, due to high computational effort and manufacturing cost.

2.2 Spectral discrimination

2.2.1 Spectrograph

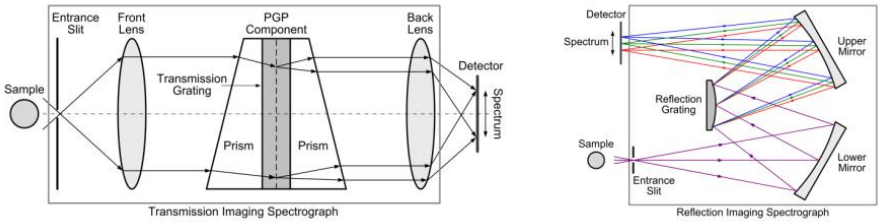


Figure 2.5. Wavelength dispersive imaging spectrographs; left: prism-grating-prism (PGP) transmission spectrograph; right: Offner reflection spectrograph.

The imaging spectrograph is the core component of line-scan hyperspectral imaging systems. It is an optical wavelength-dispersive device that spatially separates broadband light into different wavelengths. It acquires a line of spatial information from a sample via an entrance slit and projects a two-dimensional image where one dimension represents the spatial and the other the spectral axis. The two major types of diffraction gratings used in imaging spectrographs for wavelength dispersion are transmission and reflection gratings.

A transmission-grating-based imaging spectrograph is shown in Figure 2.5 left. Incoming light is first collimated by a front lens and then dispersed at a prism-grating-prism (PGP) component, where light propagation direction is dependent on wavelength. The dispersed light is projected onto a detector through a back lens, creating a special 2-D image [7].

A reflection-grating-based imaging spectrograph, based on an Offner configuration, is shown in Figure 2.5 right. This spectrograph includes a pair of spherical mirrors and a convex reflection grating. The lower mirror guides light from the entrance slit to the reflection grating, where the beam is dispersed into different wavelengths. The upper mirror then reflects the dispersed light to the detector, where a continuous spectrum is formed for each spatial point along a scanning line on the sample.

2.2.2 CZT detector for X-ray spectroscopy

A different example of device used for the discrimination of the energy (wavelength) of incoming radiation can be found for x-ray imaging.

In recent years, the semiconductors CdTe and CdZnTe (CZT) have emerged as the material of choice for room temperature detection of hard X-rays and soft gamma-rays [8](Iniewski2014). Through direct conversion of the energy of the detected photons into an electronic signal, the use of CZT detectors improves the energy, spatial and contrast resolution of imaging systems, in contrast to the indirect conversion used in conventional scintillator based detector instruments.

CZT detectors are used in two modes of operation: photon counting and spectroscopy. In photon counting mode the x-ray photons can be counted like in standard digital camera for visible light; this mode of operation can work at very high flux rate. Alternatively for each photon the system can detect precisely its energy; this mode of operation is called spectroscopy; as extracting energy information takes more processing time, the spectroscopic systems are inherently slower than the photon counting ones.

With the advance in the process technology of growing the crystals, CZT detector can be built as single pixel detector, as well as one dimension and two dimensions pixel

array, thus enabling the realization of hyperspectral imaging devices for a line scanning or snapshot imaging.

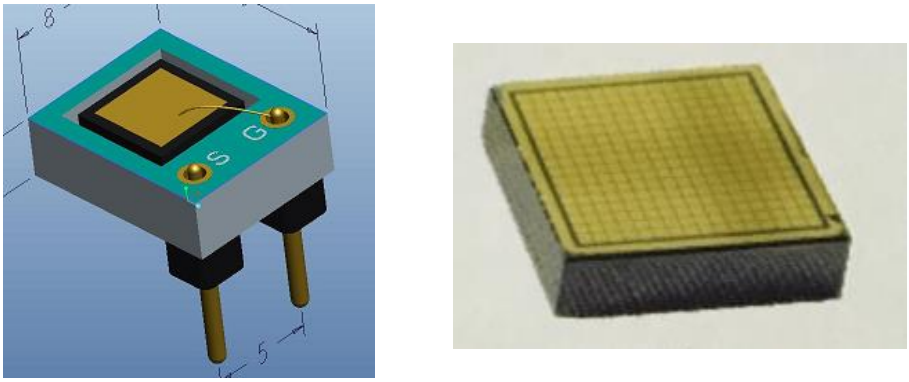


Figure 2.6. Example of single pixel and 2D array of a CZT detector.

2.3 Type of light-matter interaction

The nature of light-matter interaction is at the basis of optical imaging technologies. Figure 2.7 shows the different kinds of such interaction: the light incident on a turbid material can be back-reflected after absorption and multiple scattering events or via energy transfer from light to particles and back thru fluorescence or Raman scattering; light may also be transmitted through the material without being fully absorbed. These light-matter interaction processes correspond to different arts to implement imaging techniques (among them hyperspectral imaging) or spectroscopy: reflectance, transmittance, fluorescence or Raman scattering mode [1].

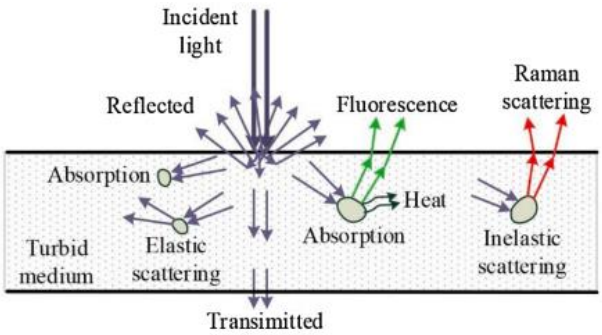


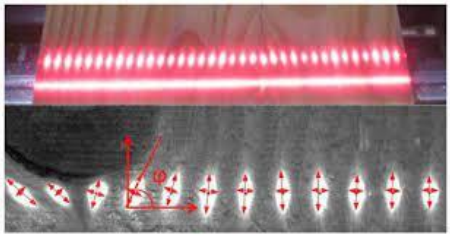
Figure 2.7. Representation of various types of light-matter interactions in a turbid material.

The principal method used throughout this project is reflectance, and some experiments were done with transmittance as well; on the other hand, fluorescence or Raman scattering were not considered.

In reflectance imaging, the incident light and sensor are positioned on the same side of the sample. The typical method to produce light is to use diffuse or uniformly distributed light sources, for example halogen lamps or LED modules mounted in one or multiple rows pointing towards the sample or in the opposite direction on a diffusing surface; the first method permits to use less light power but sometimes produces reflections if the material is lucid or glossy; the second method can be preferable to measure horticultural products, which are often pseudo spherical and glossy. An alternative method (Figure 2.8 left) is to use point lighting with a very narrow and intense light beam; this is achievable for example by using laser light sources or by focusing a lamp or LED light source on a small area, to obtain an illumination circle of about 1-2 millimetres. This creates a light scattering image at the surface of the sample and by acquiring and analysing this image it is possible to determine the light absorption and scattering coefficients. This is useful to assess information on fruit samples like texture (firmness) or flavour (for example sugar content) or on wooden solid samples like the fibre orientation.

Reflectance imaging probes only the superficial region of the sample, ranging from tenths of a millimetre to several millimetres depending on the optical properties of samples and on the lighting and sensing hardware. On the other hand, this technique works well on a wide range of wavelengths, making it possible to acquire a useful signal in the full visible and infrared range that are considered in this work.

In transmittance sensing, the incident light and sensor are positioned on opposite sides of the sample to measure. This configuration has the advantage, compared to reflectance, in detecting internal characteristics of the samples, like defects or quality parameters (for example sugar content) that are not evenly distributed in the sample body or just observable from the external surface only. The main difficulty in transmittance is the absorption of light in the observed body, requiring very high-power light sources and sensitive sensors, making it more difficult to implement. The optical arrangement requires special care to avoid that the tiny residual signal attenuated by the sample is shadowed by the larger signal that can scatter on the outside of the sample or on the measuring environment (see Figure 2.8 right for an example of setup with a light shield to avoid that the intense light coming from the top interferes with the small signal that is found at the bottom of the fruit). In practical terms, only the NIR range (700-1000 nm) is used, since biological tissues are relatively transparent in this area, while sorption in the visible range (below 700nm) or in the short-wave infrared (above 1000) is high.



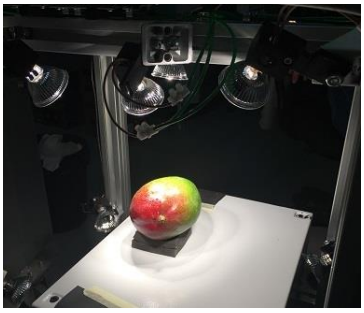


Figure 2.8. Left: example of point light sources used to estimate fiber deviation on wood. Right: setup for measure in transmittance.

2.4 Spectroscopy and spectrometry

Spectroscopy is the science of studying the interaction between matter and radiated energy. It is the study of absorption characteristics of matter, or absorption behaviour of matter, when subjected to electromagnetic radiation.

The study of spectrometry dates back to the 1600s when Isaac Newton first discovered that focusing light through glass split it into the different colours of the rainbow (known as the spectrum of visible light), and it has expanded to include interactions between electrons, protons and ions. Multiple scientific fields including chemistry, physics, and astronomy have grown as a result of spectroscopy.

On the other hand, spectrometry studies the practical study and measurement of a specific spectrum, i.e. is the method used to acquire a quantitative measurement of the spectrum. It is the practical application where results are generated, helping in the quantification of, for example, absorbance, optical density or transmittance.

In other words, spectroscopy is the theoretical science, and spectrometry is the practical measurement. While spectroscopy does not generate any results, because it is the theoretical approach to science, it is the application of spectroscopy that creates the results that can be assessed.

Spectrometry involves the splitting of light (or more precisely electromagnetic radiation) into its constituent wavelengths (a spectrum), which is done in much the same way as a prism splits light into a rainbow of colours. Old style spectroscopy was carried out using a prism and photographic plates, while modern spectroscopy uses diffraction grating to disperse light, which is then projected onto digital image sensors.

2.5 Chemometrics

Chemometrics [9] [10] is a chemical discipline established at the beginning of the 1970s by Wold, Kowalsky and Massart that uses mathematics, statistics and formal logic to design or select optimal experimental procedures, to provide maximum relevant chemical information by analysing chemical data and to obtain knowledge about chemical systems.

In contrast with the classical approach, that aims to understand effects, chemometrics give up the necessity to understand the relationships and effects and identify other aims like prediction or classification. The classical approach separates all factors, examining one factor at a time, and produce a theory from which a model is derived. On the other hands the chemometrics approach uses multivariate methods, all variables are considered at the same time, and try to find models to fit the data. As a result, the classical approach determines new causal relationships and discovers new natural laws, whereas chemometrics finds formal relationships. The classical approach has the advantage of being generally accepted and well based, and the relationships have a physical meaning; the disadvantage is that in nature factors are often correlated and cannot be always orthogonally separated. The advantage of chemometrics is that correlations between variables can be utilized, but the relationships do not have necessarily a physical meaning.

2.6 Analysis methods

The goal of spectroscopy and hyperspectral data analyses is to build a predictive or classification model. To exploit the high dimensional nature of the spectral data, the model should be multivariate [1]. Depending of the specific detection task, it can be quantitative to provide numerical prediction (for example to determine the concentration of chemical constituents), or qualitative to perform classification (for example for defect detection or surface segmentation). Both kind of models are computed based on learning from the given data, based on a chemometric approach. In general, separate datasets are required for training, validating and testing the model.

The typical methods for multivariate analysis use principal component analysis (PCA), multiple linear regression (MLR), principal component regression (PCR) and partial least squares (PLS) regression [11] [12] [13].

Principal component analysis (PCA) is used as a tool for screening, extracting and compressing multivariate data. PCA employs a mathematical procedure that transforms

a set of possibly correlated predictor variables (the measured spectra) into a new set of non-correlated variables, called principal components. PCA produces linear combinations of variables that are useful descriptors or even predictors of some particular structure in the data matrix.

Multivariate regression techniques aim to establish a relationship between the measured predictor variables X ("X-variables", the spectral matrix) and the observed response values ("Y-variables", attributes of interest or target).

In multiple linear regression (MLR) the response is approximated by a linear combination of the predictor values at every single wavelength. The regression coefficients are estimated by minimising the error between predicted and observed response values based on least squares. MLR models typically do not perform well because of the often high co-linearity of the spectra and easily lead to overfitting and loss of robustness of the calibration models.

Principal component regression (PCR) is a two-step procedure, which first decomposes the predictor X-variables by a principal component analysis (PCA) and then fits a MLR model, using a small number of principal components (PCs or latent variables) instead of the original variables as predictors. The advantage with respect to MLR is that the new predictor variables are uncorrelated, and that the noise is filtered. The main drawback of PCR is that the latent variables are ordered according to decreasing explained variance of the original predictor variables (the spectral matrix), which is not necessarily the most informative with respect to the target variable.

Partial least squares (PLS) regression is a generalisation of multiple linear regression (MLR) [14]. Unlike MLR, it can analyse data with strongly collinear (correlated), noisy and redundant predictor variables (X-variables) and efficiently model the response variables (Y-variables) at the same time. In PLS regression an orthogonal basis of latent variables is constructed one by one in such a way that they are oriented along the directions of maximal covariance between the predictors (spectral matrix) and the responses. In this way it is ensured that the latent variables are ordered according to their relevance for predicting the Y-variable. Interpretation of the relationship between X-data and Y-data (the regression model) is then simplified as it is concentrated on the smallest possible number of latent variables. PLS is commonly referred as PLSR (for regression, where the target variable Y is a continuous value) or PLS-DA (for discriminant analysis, where the target variable Y is a set of classes expressed as one-hot encoding).

PLS will be the main method used to perform the chemometric analysis in the tests that will be presented in this work.

2.7 Applications

Hyperspectral imaging originated initially from remote sensing and has been explored by NASA for various applications. The advantage of acquiring images across a wide range of electromagnetic spectrum has been applied to many other different areas; a few examples are agriculture, recycling, art and heritage, medicine and pharmaceutical, forensics.

2.7.1 Applications for fruit quality control

One important field where hyperspectral imaging has been applied is the quality and safety control for fruit, vegetables and horticultural products in general.

Many studies have been conducted on the use of spectroscopy for this purpose and even if this measure is lacking the spatial information, these studies are relevant for the data analysis part that can be borrowed in hyperspectral applications as well. Spectroscopy [15] is a method capable to analyse internal compounds in fruits and vegetables: it is non-destructive, inexpensive, rapid and reliable technique that has been used in food chemistry for the quantitative and qualitative determination of different compounds in fruits. The most used spectral range has been the near infrared (NIR) spectroscopy [16] [17] [18]. This technique has been used for the determination of soluble solid contents (SSC), firmness, acidity, dry mater, chemical substance (glucose, sucrose, citric acid, malic acid, starch or cellulose) in different fruits [19] [20] [21], for the determination of maturity index [22], internal quality index [23] [24].

The principal limit of spectroscopy is that it can measure a single point over the fruit sample. On the other hand, hyperspectral imaging (or imaging spectroscopy) extends spectroscopy with the ability to measure both the spatial and spectral information, thus it allows the study of the spatial distribution of properties, which is not possible with conventional spectroscopy.

These reviews [25] [26] [27] [1] very useful, as a thorough overview on a multiplicity of applications of hyperspectral imaging to the quality assessment on fruits and vegetables and its advancement in recent years.

Below a few examples of applications in recent years.

One of the first use of hyperspectral imaging was described by Martinsen et al. [28], where they used this approach for the determination of soluble solids in kiwifruit.

Menesatti et al. [29] studied the possibility to use hyperspectral imaging in SWIR range (900-1700 nm) to assess the correct harvest time of apples, by automatically imaging apples cut in halves, to automatize the traditional method through the starch-iodine test, where fruits halves are dipped into an iodine solution and patterns are visually evaluated and compared with reference charts by experts.

Mendoza et al. [30] used hyperspectral imaging for the inline prediction of firmness and SSC for different apple cultivar.

Lleo [31] applied hyperspectral imaging in the VisNIR range (400-1000nm) for the prediction of the maturity of peaches; they computed different maps of maturity indices and showed that the ripening was not uniform throughout the whole fruit.

Rajkumar [32] used hyperspectral imaging in the VisNIR region (400-1000nm) to study quality and maturity stage in banana fruits; quality parameters like moisture content, firmness and total soluble solids were determined and correlated with the spectral data using partial least squares analysis, while principal component analysis was used to select optimal wavelengths; the study was conducted at three different storage temperatures (20, 25 and 30°C).

Schmilovitch [33] studied intact bell peppers using hyperspectral imaging (550-850nm), obtaining high correlations by means of PLS regression for some internal compounds (SSC, chlorophyll, carotenoid and ascorbic acid); they also used the spatial resolution of the hyperspectral technique to obtain mappings of these parameters.

Munera [15] studied astringency of persimmon by using hyperspectral imaging (460-1020nm); they used PLS-discriminant analysis (PLS-DA) to classify fruits in a few different stages and were able to identify a reduced number of informative wavelengths using the Successive Projections Algorithm (SPA). Their hyperspectral imaging system was an industrial camera coupled with two liquid crystal tunable filters (LCTF).

Zhou [34] used both VisNIR and the SWIR spectral ranges (380-1000nm and 875-1700nm) to investigate the feasibility and potentiality of determining firmness, soluble solid content (SSC) and pH in kiwifruit. They used a linescan hyperspectral reflectance imaging configuration and a multitude of analysis methods to establish models, in order to obtain very accurate prediction models.

Li [35] used both VisNIR and SWIR wavelengths (600-1000nm and 850-1600nm) for the non destructive measurement of firmness, SSC and color components of plums; they used two line scanning (Push broom) hyperspectral imaging systems combined with a moving translation stage. In order to achieve a more accurate measurement, they considered two opposite sides of the fruits.

Concerning the topics analysed in this project, works were done on avocado [36] and on citrus [37] [38], studies that applied NIR spectroscopy (not hyperspectral imaging) to the estimation of various quality parameters on avocado and citrus fruits.

2.7.2 Application in wood sector

Hyperspectral imaging has been used in different studies for the qualitative and quantitative characterization of wood.

Wood is an essential material for many industrial segments, such as the building, furniture and paper industries. Wood is considered an important and complex raw material. It is formed by three major macromolecular components: cellulose, hemicellulose and lignin, with small amount of minor extractive compounds.

Infrared spectroscopy has proven to be a powerful tool for the analysis of the properties of wood over the past 20 years [39]. In the range of 800-2500 nm there are numerous overtone and combination bands of vibrational frequencies due to functional groups, such as C-H, O-H and N-H. It is possible to detect these differences on objects at wavelengths that are not visible to the naked eye. The overlapping appear non-specific and poorly resolved, but they can be evaluated well by multivariate calibration algorithms and statistical methods by means of chemometrics (as for example principal component regression (PCR) or partial least squares (PLS) analysis).

Below a few examples of applications in recent years.

Kelley et al. [40] used spectrometry in the range 500nm to 2400 nm to measure chemical and mechanical properties of solid wood of specie loblolly pine wood. They were able to estimate with a good correlation the chemical composition of wood (lignin, extractives, glucose, xylose, mannose, and galactose), its mechanical properties (stiffness or modulus of elasticity MOE, and modulus of rupture MOR) and the microfibril angle as well. They used a spectrometer (not images), and collected about 30 scans per sample that were averaged into a single spectrum.

Thumm et al. [41] used a hyperspectral imaging system in the SWIR range (900-1700 nm) from Specim to scan discs of wood to visualize the distribution and variation of lignin, galactose and glucose in the samples, which are related to the chemical composition of wood. The same author later extended the study [42].

They also used the same setup [43] to scan 30 shooks of radiate pines in order to detect resinous defects; their model was able to further well discriminate resin from other visually similar defects like sapstain, pith or kiln brown stain. Both studies used PLS as mathematical tool to compute their models.

Meder at al. [44] used a hyperspectral system to predict the severity of compression wood in samples and compared with a reference given by a subjective microscopic assessment, but without achieving good correlations.

Haddadi et al. [45] used a SWIR hyperspectral system (900-1700 nm) on small samples of wood (subalpine fir) cut out from boards to predict their moisture content and density (wood density and basic specific gravity). They used partial least squares (PLS) regression models, obtaining good results for moisture.

Ma et al. [46] used hyperspectral imaging in the SWIR range (1000-2100 nm) to identify the wood species. The measure setup was arranged in order to evaluate the light scattering patterns of an intense halogen light spot (1mm diameter) on the wood surface; this arrangement is able to separate between light scattering and absorption, thus achieving better results than simpler NIR spectroscopy.

In a later study, Kanayama et al. [47], applied data reduction with PCA and convolutional neural networks to the same kind of images, being able to accurately distinguish between 38 species of hardwood.

Mishra et al. [48] used hyperspectral imaging for digital phenotyping of whole plants in a non destructive manner; plant phenotyping is the measurement of the interaction of a plant with its surrounding environment, which is important in the science domain and in breeding programs to study the performances of plants exposed to different conditions and identify the best performing genotypes.

Kobori [49] used a hyperspectral push broom system in the Vis-NIR range to monitor the moisture content on small slices of wooden boards, during their natural drying process; they considered two wood species, European beech and Scots pine, obtaining very good correlations.

Colares et al. [50] applied SWIR (1200nm to 2500nm) hyperspectral imaging for the determination of the distribution of holocellulose (cellulose + hemicellulose), lignin and extractives on samples of Mahogany of very small size (approximately 1 cm per side).

2.8 Off-the-shelf components for hyperspectral systems

Although the hyperspectral imaging is a relative recent technology, there are already different vendors that offer off-the-shelf components, in terms of hardware equipment and software. In this section I present a non-complete list of those that were considered throughout the duration of this project.

Hardware products:

- Specim (Finland) is a supplier of hyperspectral imaging systems, offering complete camera solutions covering a board spectrum range (from 400nm to

12.4um, i.e. visible, near infrared, short-, mid- and long-wavelength infrared), spectrograph only and software.

- Innospec (Germany) is another supplier of complete hyperspectral imaging systems and components for industrial applications.
- IMEC (Finland) is a large research & development organization that, leveraging its background in semiconductor fab, equipment and process technology, offers the technology to deposit a number of interference-based optical filters on top of image sensors, with a single pixel granularity. This permits to realize snapshot multispectral sensors with 16 or 25 bands, as well as linescan sensors with some tens of bands.
- Eoptis (Italy) has developed a small multispectral device, made of 9 discrete image sensors coupled with different bandpass filters, suited for portable or aerial applications in precise farming or environmental monitoring applications.
- Glana (Sweden) is a startup that has developed a high spatial resolution hyperspectral sensors, by coupling an image sensor with a linear variable filter (LVF).

Software products

- Perception Park (Austria) offers a generic data processing solution enabling the industrial use of industrial cameras for high-speed tasks like sorting in different industries (food processing, recycling, mining, pharmaceuticals). They have developed the concept of Chemical Colour Imaging to simplify the management of the complex spectral information into simpler colour images.

Figure 2.9. Example of off the shelf products. From top left: Specim, Innospec, IMEC, Eoptis, Glana.

X-ray hyperspectral

A special mention is for some recent developments of the technology capable of measuring the spectrum of the absorbed x-ray radiation with the possibility to pixelize these sensors.

- DeeTee (Finland) develops x-ray detector solution for medical, security and industrial applications; one of their division is focused in the realization of multi-energy x-ray detector, optimized for security and industrial applications that require state-of-the-art material discrimination capability even when they have similar atomic compositions.

- Due2lab (Italy) develops x-ray spectrometers based on innovative materials, with knowhow in the CZT detector arrays fabrication and in the relative signal read out and elaboration electronic.
- X-Next (Italy) develops multi-energy x-ray detector and realizes inspection systems for food safety, security controls, material recycling and pharma safety applications.

3 Hardware

This chapter explains the laboratory setup for the data acquisition and the developments that have been done to industrialize a working prototype, in particular from the hardware point of view[1].

3.1 Initial setup

The first attempts I made to apply the hyperspectral techniques were based on a few measures done using a closed system that permitted to acquire and model data to obtain pseudo-colour maps like depicted in Figure 1.1.

After these first positive attempts on few data, a hyperspectral imaging camera sold by the German company Innospec was acquired to extend the measures. The system used a SWIR sensor with a resolution of 320x256 pixels and spectral range between 950 and 1700 nm. The illumination used a set of halogen lamps, in order to provide a high light flux and relatively white spectrum, but with the disadvantages of producing a lot of heat, and the need of a certain time to reach a stable light emission.

This setup was used to collect diverse kind of data, both on wood materials and on different fruits, to test and asses the validity of the technology and of the processing.

The most promising tests were:

- wood boards (hardwood, like eucalyptus or oak), detection of heart and sapwood
- wood boards (resinous wood, like pine or spruce), detection of stain, rotten, different fungi
- wood boards (resinous wood, like pine or spruce), detection of juvenile wood
- fruit (avocado), determination of dry matter
- fruit (citrus), determination of sugar content
- fruit (dates), measure of humidity

[1] This part of work was done in cooperation with other people. For section 3.2, I did the overall design, component selection and evaluation, development of part of firmware and the acquisition software. Andrea Gottardo and Andrea Azzalin made the design of the electronic boards and the development of firmware; Alberto Celin worked on firmware development; Johann Thaler helped for the design of the mechanical aspects. For section 3.3, I searched and selected the LEDs models, and designed the arrangements to build and test; Simone Faccini made the design of the LED electronic boards; Simone Faccini and Matteo Caffini did together with me the evaluation of the different developed solutions.

- fruit (raspberry), determination of firmness

In parallel to these feasibility studies, it was analysed the cost of such a system and the limitations of its integration in the existing scanners produced by Microtec.

Cost was an important point. The existing cameras are very expensive, the cost is easily some tens of thousands of euros, a price which is difficult to amortize in our systems. Counting on the promising results, and with the experience in designing custom electronic, the design of a proprietary camera based on existing sensors to be used as two dimensional sensor in the push broom setup was promising.

The integration in existing system was another important topic. Microtec produces scanners to measure the quality of wooden boards, which are based on a combination of different sensors working in the visible, infrared and x-ray spectral range, and on the data fusion of the images coming from them; see Figure 3.1 for an example of inline scanner and the different images that are acquired. Due to space limitation, these sensor works in the same area, thus it is necessary to time-multiplex them, by pulsing the illumination over time, or using band-pass filters, to avoid crosstalk between sensors. Because of this, the use of halogen light poses problem since it is difficult to pulse them at the required frequency of some kHz or to superimpose due to its broad spectral emission, which would disturb the existing optical sensors. Therefore the use of LEDs as light source was preferred to halogen lamps.

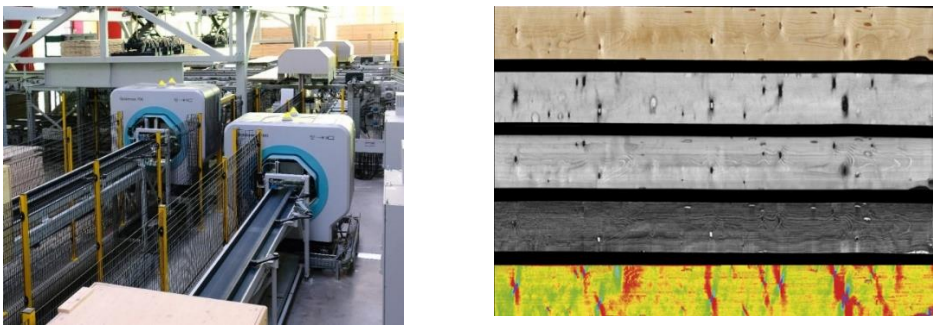


Figure 3.1. Left: two Microtec Goldeneye scanners installed in a sawmill. Right: multiple images taken from different imaging systems working in the same space.

The flexibility of the sensor was another limiting factor. As previously introduced, a push broom system uses a two-dimensional sensor array and a prism to imagine one spatial dimension and to split the wavelength axis on the other sensor's axis. In some application it is not necessary to use the whole wavelength range to solve the task, and it would be possible to use only a subset of it with the benefit of reaching a higher speed.

Unfortunately, the existing commercial solutions do not offer this flexibility, but have only the possibility to read out a single region of interest.

The above-mentioned reasons convinced me that the design of a proprietary hardware solution, was needed to solve these limitations and to reduce costs. Figure 3.2 shows the main components of a push broom hyperspectral system, that is spectrograph or prism, sensor and illumination, and my work considered in particular the camera and the illumination unit. These components to design need to cover the wavelength range in visible, infrared and SWIR, thus both the sensor and the illumination components need to be selected accordingly.

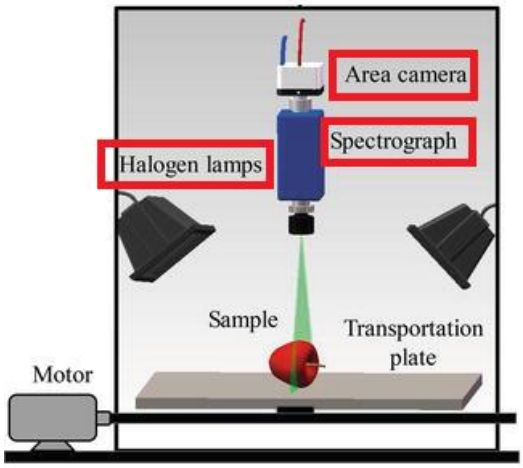


Figure 3.2. Elements of a hyperspectral push broom system, with the main components: camera, spectrograph, illumination.

3.2 Development of the camera

The first step for the development of a custom hyperspectral system consisted in a survey of existing camera. To cover the visible and near infrared range (400-1000 nm) was easy, both because there exist an immense number of available cameras and sensors. In addition Microtec had already developed some cameras based both on commercial but also on a proprietary developed CMOS sensor. The selection was driver by the need of having a low noise and high sensitive sensor; therefore the proprietary custom sensor, designed to reach a very high frame rate but with some compromise on noise was discarded and one sensor from manufacturer Sony, was preferred.

The selection of a sensor in the SWIR range (950-1700 nm) was harder. I did an extensive research on existing camera, but at the time of this survey in 2018, the number of available models was low and cost were still quite high, far beyond the 10K€. In addition to the camera survey I did also a survey of different sensor manufacturers. The list of manufacturers of these sensors is also limited, since the process of producing hybrid InGaAs sensor array is not as widespread as for silicon based sensor. The survey on InGaAs sensors was based on extensive internet research, contact with suppliers and visit of some trade shows. The most promising manufacturer were Chunghwa (Taiwan, with a distributor in Germany), Hamamatsu (Japan, with a distributor in Italy) and Sofradir/Linred (France).

The initial list of possible sensors was:

Name	Resolution (pixel)	Frame rate (fps)	Note
Chunghwa	320x256	330	Cooled and uncooled
Hamamatsu	128x128	280	Low resolution but lower cost
Hamamatsu	320x256	225	
Sofradir/Lynred	640x512	320	Uncooled

The chosen sensor was from Sofradir (the company rebranded itself as Lynred in 2020), since it was a good compromise between resolution, frame rate, low noise and the possibility to use it without cooling. Together with the hardware team at Microtec, we designed the electronic board, we wrote the firmware to control the sensor and the software drivers to communicate from a computer with Windows operating system. Figure 3.3 shows two pictures of the electronic boards.

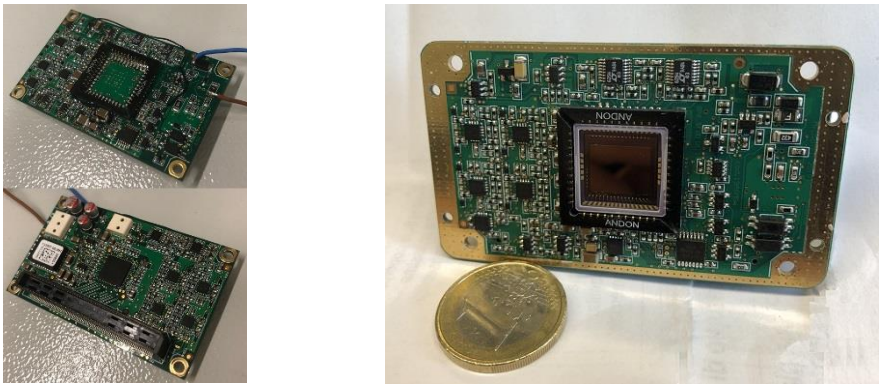


Figure 3.3. Picture of the electronic board, first and second prototype.

A particular care was put in the design of the firmware. A very important feature was the possibility to freely select a lower number of rows. This is especially useful since in a push broom system, the camera acquires frames where the spatial dimension is mapped on the camera horizontal axis and the spectral dimension is mapped on the vertical axis. For some application it is not necessary to use the full spectral range, but the relevant information could be represented in only a subset of important wavelengths. In addition, the sensor frame rate depends on the number of acquired rows, by reading a subset of the pixel it is in general possible to increase the frame rate. Therefore, a very useful feature is the possibility of reading a subset of not consecutive rows, with a corresponding increase in the framerate. Figure 3.4 depicts this concept; the red box shows the pixels that are read from the sensor matrix; on the left side the region of interest is a compact area, a feature which is common to many sensors; on the right the region of interest is made of different non-contiguous rows, and this feature is not generally available.

This possibility, while present in the Vis-NIR sensor (one can program a certain number of rectangular regions to access), is not present in the SWIR sensor (one can only access a single compact region). To overcome this, a special firmware to control the sensor was developed by my colleagues, which uses the trick of aborting the readout of the rows to skip, thus allowing to save the time to access the unnecessary pixels.

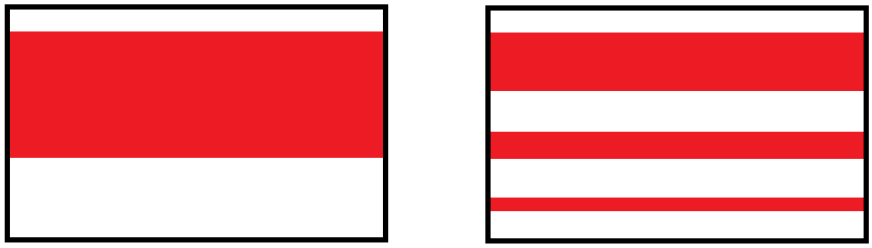


Figure 3.4. Partial sensor readout. Left: compact region of interest. Right: noncontiguous region of interest.

The SWIR camera uses an uncooled sensor, but needs a relative stabilized temperature in order to work well. To obtain this it, an enclosure with a Peltier element was designed. The Peltier element has the function to keep the sensor and electronic temperature of the sensor and the electronic stable, achieving a stability of +/-0.05°C independently from the outside temperature (Figure 3.5). The downside of this is that the final camera body is relative large and the long time required to reach a stable working point.

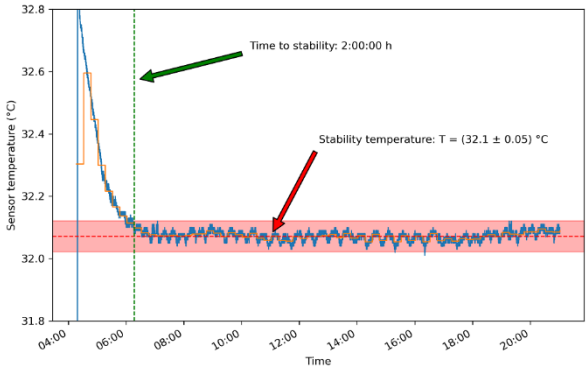


Figure 3.5. Plot showing the temperature stability obtained with the Peltier element.

Finally the lens for the SWIR hyperspectral system is optimized for working in the SWIR range, thus it is different from conventional lenses for the optical materials and coatings.

In parallel with the design of the two camera modules, I started the selection of the spectrograph, which is the optical component that acts as prism to split the wavelengths in order to build a push broom system. The first selection ended with a short list of possible suppliers, namely Specim (Finland) and Innospec (Germany). The two companies supply optics for different wavelengths ranges. I decided to use the prism in the Vis-Nir range from Specim and for the SWIR range from Innospec.

Figure 3.6 shows the CAD drawing of the final housing; from left to right: opening for the lens, spectrograph (in yellow), the controller for the Peltier unit (in front of the spectrograph), the sensor housing (orange), electronic housing (gray) and Peltier unit (below the orange housing).

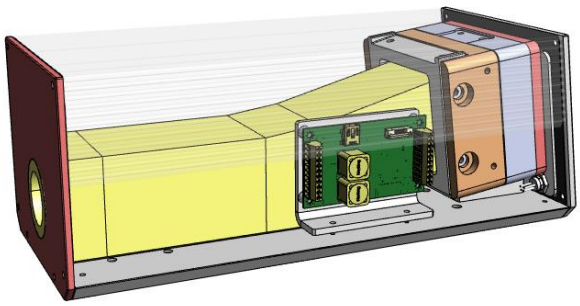


Figure 3.6. Technical drawing of the SWIR hyperspectral system, comprehensive of prism and Peltier cooling part.

The outcome of this development were two cameras, depicted in Figure 3.7, with all the characteristics needed for an industrial use. Table 3.1 summarizes the main technical specifications.

	Vis-NIR	SWIR
Wavelength range (nm)	400-1000	950-1700
Resolution (pixel)	2048x1536	640x512
Subsampling	yes	
Binning	yes	
Power supply	24V	
Size (mm)		
Interface	1 Gbit Ethernet	
Synchronization	over Ethernet, jitter ~10 usec	

Table 3.1. Technical specification of the two hyperspectral sensors.

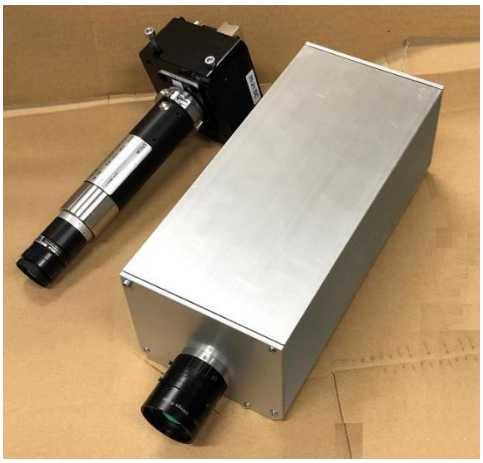


Figure 3.7. The two hyperspectral system, both comprehensive of camera, prism and lens: Vis-NIR (left) and SWIR (right).

3.3 Development of the illumination

The illumination unit in the initial prototype used halogen lamps, in order to be cheap, easy to build, have a relative white and flat spectrum and a very high flux. However, they also pose many constraints: the amount of produced heat is very high and in some environments (for example in sawmills where dry wood is processed and inflammable fine sawdust is present) there is the concrete risk of starting a fire; their average life time is not very long, in the range of 2000-4000 hours, therefore their use in industrial 24/7 environment is not ideal; their illumination spectrum is very wide, also beyond the useful used spectral range, and this makes more difficult to overlap multiple measurements in the same area; in addition it is difficult to use band-pass optical filter

to reduce the spectrum width (due to heat) or to pulse the illumination in order to use time-multiplexing (halogen lamps cannot be pulsed at high frequencies).

Because of these points I started to look for alternatives and LEDs were considered. LEDs (light emitting diodes) are evolving at a very fast pace in terms of emitted light flux and emitted spectrum (colour), and are replacing other light sources (halogen, fluorescent bulbs) in many fields.

In order to be used in a push broom hyperspectral configuration, the lamp should illuminate a long and narrow stripe, corresponding to the line scanned by the camera, and not a circular spot or a diffuse illumination.

I did an extensive research on producers of complete lamps, of LEDs and also chips to understand the current state of the art.

Since in Microtec we already use custom designed LED lamps in the visible and NIR range, I specifically searched for complete lamp solution in the SWIR range. The best solution we found was the solution from Metaphase, a company based in the US, which offered a solution in the SWIR range.

Regarding the single LEDs (not a complete assembled lamp), the following table shows the result of the survey, where the best options of discrete LEDs for the three wavelength ranges (visible, NIR and SWIR) are listed.

Company	Wavelength
Cree	Vis (400-650)
Luxeon	Vis (400-650)
Nichia	Vis (400-650)
Osram	NIR (700-1000)
Ledengin	NIR (700-1000)
Ushio	SWIR (950-1700)
Dowa/Epigap	SWIR (950-1700)

Table 3.2. List of LED suppliers, sorted by the emitted wavelengths.

Each LED has normally a band emission not very large, normally in the range of 50-100 nm. Therefore, in order to cover a large spectral band, it is necessary to combine multiple different LEDs, to sum each different spectrum to obtain a relatively flat wide spectrum. This is different in the visible wavelength range (400-700 nm), where a lot of work has been done in order to have a single illumination source able to cover the relatively large spectrum. Unfortunately, a similar solution is not yet available in the NIR and SWIR range; to my knowledge only a solution currently exists, from company Osram, of a broadband NIR LED, but the efficiency is still too low.

Figure 3.8 shows the typical spectrum in the visible range obtained by a single LED, the most common spectral emission wavelengths in the NIR and SWIR ranges.

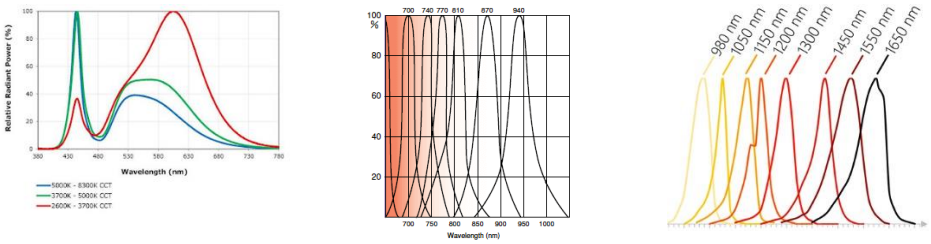


Figure 3.8. Typical emission spectrum of LED in the VIS, NIR and SWIR ranges.

In addition, in order to illuminate a long and narrow stripe, it is necessary to mount many equal LEDs in a row, or to use a particular optical element with an elliptical beam output. In some case it possible to have the optic already assembled in the LED package; with this, instead of having the normal radiation pattern with a large angle (typically of about 120°), it is possible to concentrate the emitted light in an angle of 10-20°. Figure 3.9 shows an example of package LED without optic and with optic, and the corresponding radiant emission plot.

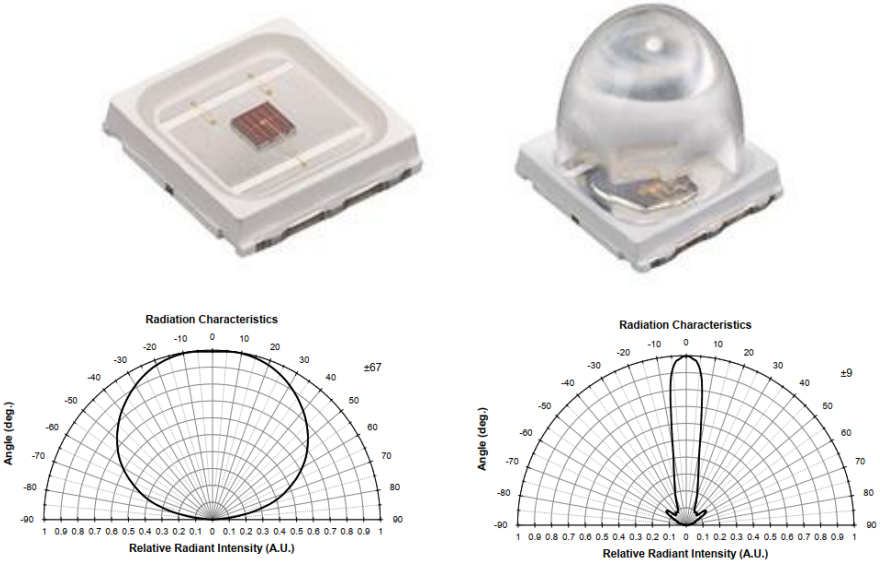


Figure 3.9. Picture of a led without and with primary lens (top), and the corresponding emission pattern.

We designed a set of small PCB (printed circuit boards), in order to assemble multiple LEDs in a small area, and to replicate this on a long stripe. Figure 3.10 shows one example of this design; on the left an initial prototype is shown, the board can mount six different packaged LEDs, and the special design of the lateral connectors allows to assemble on a line multiple boards; on the right the final lamp is visible; note that there are two rows of these boards in order to increase the total light flux, and that these are alternated, in order to reduce the intervals between bright spots.

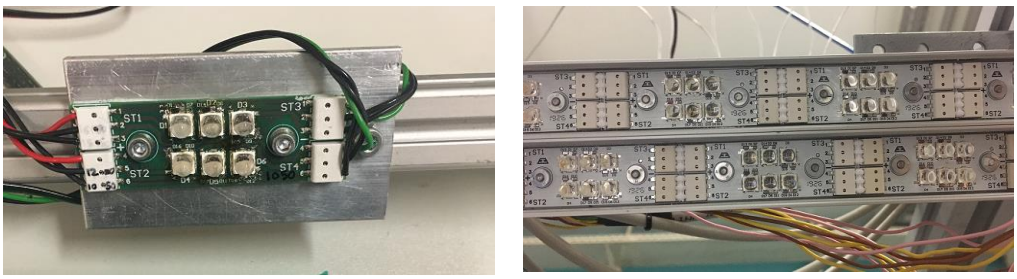


Figure 3.10. Picture of an early LED board and the finalized version mounted in a double stripes configuration.

This solution, while simple and effective, has the drawback that every emitting source with different colour is slightly shifted in space, thus the total emitting area is not punctual but has a surface of a few square centimetres (for example the above shown board has an emitting area of 20x13 mm). The drawback of this arrangement is that different points in space are illuminated by light with different colour coming from slightly different directions, and given that the illuminated material (in our case wood and fruit samples) are not perfect diffusive materials, it is more difficult to perform the calibration of the light source.

There are two solutions to this issue. The first one is to use a diffuser in front of each LEDs and a secondary lens to concentrate the light on a narrow stripe; this solution is used for example in lamps produced by the company Metaphase. The principle (Figure 3.11) is to have a row of emitting LEDs with a primary lens molded on them, followed by a first cylindrical lens to obtain parallel rays on one axis, by a diffuser used to spread and mix all the rays, and finally by a second cylindrical lens to exit again with parallel rays on one axis; the cylindrical lenses are needed in order to concentrate light on an elongated stripe.

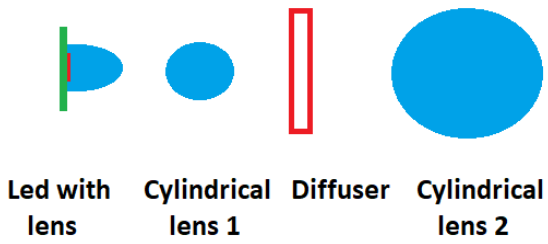


Figure 3.11. Basic principle of mixing light.

The second solution is to assembly multiple single chip in a single package; in this way the emitting area goes from a few centimetres to 1 or 2 square millimetres. For example Figure 3.12 shows a solution offered by the company Epigap that permits to combine 5 or 7 chips in a single package; the total package size is 3x3 millimetre, while the internal emitting area is about 1 square millimetre.

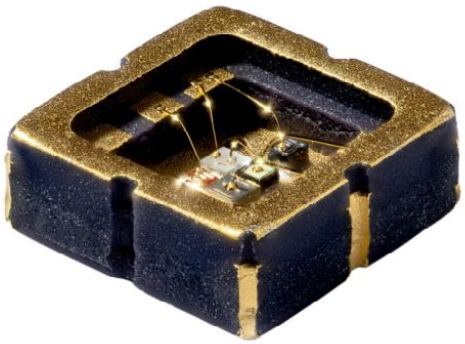


Figure 3.12. Multi wavelength chips assembled in a single package.

I've been in contact with the company Dowa Electronics Materials based in Japan, that develop and produce the wafers and the chip, in order to follow their developments, in particular concerning the number of different supported wavelengths and the emitted light flux, and with the linked company Epigap based in Germany that uses these chips and perform the assembly in a package, with the options of mounting a primary lens to concentrate the light and to assembly multiple chips in the same package.

Contacts are still ongoing in order to prepare a prototype of a lamp consisting of a continuous row of multi-chip led.

4 Software and Algorithms

This chapter explains the complete workflow of the hyperspectral application that have been developed, in particular from the software point of view[2].

In general terms, the workflow is divided in the steps shown in Table 4.1. The workflow is shown for the creation of a model and for its deployment at runtime.

Off-line (modelling)	On-line (runtime)
• Data acquisition • Pre-processing (dead pixel, smile, binning) • Normalization	• Data acquisition • Pre-processing (dead pixel, smile, binning) • Normalization
• Dataset creation (hdf5, hs cube, wavelength map, metadata) • Labeling (definition of roi) • Sampling of spectrum • Modeling	• Aggregation of spectrum (in case of value) • Model inference (spectrum -> class/value) • Post processing (spatial filtering)

Table 4.1. Workflow for modelling and runtime.

4.1.1 Data acquisition

This is the first step of the pipeline, that is to receive data from the sensor itself. Since we are using a push broom configuration, one sample corresponds to a two-dimensional image.

The two hyperspectral systems developed at Microtec use 1Gigabit Ethernet as interface. They implement a proprietary communication protocol, common to all sensors used in the company, that permits the integration of multiple sensors, the exact synchronization among multiple cameras and with digital input information (used to know the exact position of the objects under measurement), which is based on a high-

[2] This part of work was done in cooperation with other people. I did the overall design of the workflow process and wrote the code for the data acquisition and analysis; Simone Faccini and Matteo Caffini developed the software pipeline as well.

performance driver to avoid the overhead given by the operating system in case of high bandwidth.

A minimum set of configuration is necessary, the relevant parameters for these applications are integration time, analogue/digital gain, frequency at which frames are acquired, the selection of a region of interest or a pattern of subsampling/binning.

A unique software was written, to control and receive data from both type of sensors that was used during the project to setup the sensors, collect the stream of images to build the hypercube, apply the developed model and eventually to save the data to disk.

4.1.2 Pre-processing

The typical operations of preprocessing are dead pixel correction, smile correction and binning.

While nowadays the quality of processes used to produce standard sensors permit to practically completely avoid defects at pixel level in the case of silicon-based sensor, dead pixel are a common problem for InGaAs sensors used for the SWIR application.

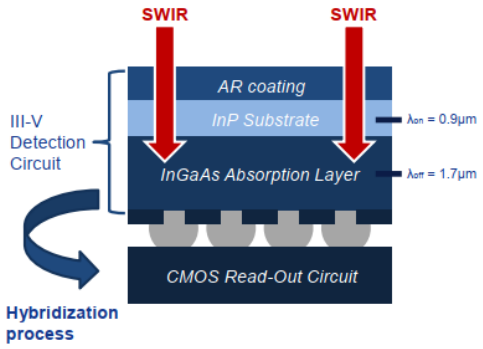


Figure 4.1. Block diagram of a hybrid CMOS- InGaAs assembly.

The production of these sensors requires a so called hybrid process, where quality and quantity of processed wafers are lower compared to standard processes. This causes that a few imperfections are tolerated, having as result that some pixels can be defective. For example in the chosen sensor, up to 0.1% of pixels (up to 300 pixels in the whole matrix) can have a responsivity that deviates more than 30% from the average. This high probability of bad pixels requires a treatment. The first step is to characterize each sensor and build a map of bad pixels; this is done by analysing the dark and white images (the same used later for the image normalization). The procedure of correcting these bad pixels is simple. At runtime, the value of these pixels is computed by the non-bad pixels in a neighbourhood of 3x3 or 5x5 pixels.

In a push broom setup, each image present one spatial axis and the spectral axis projected on the x-y plane. Ideally the two axes are perfectly orthogonal and linear, for example the spatial axis is projected on the x axis and the spectral axis on the y axis. When this is the case, the mapping from image to wavelength is a simple linear function on the pixel y coordinate. Smile is a particular kind of image distortion, where the peak position of a known wavelength is not a horizontal line but a curved one.

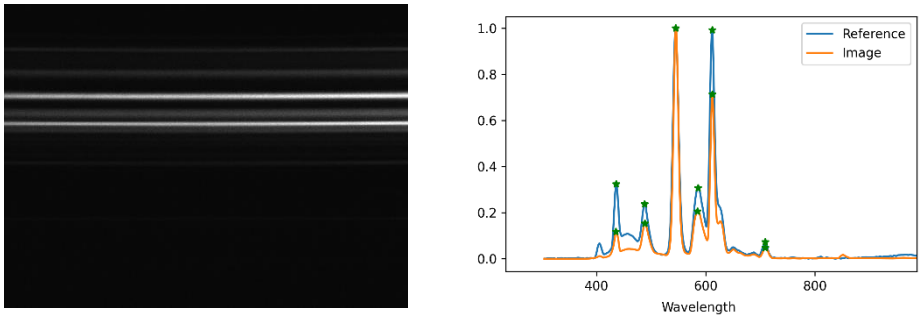


Figure 4.2. Wavelength calibration with a neon light source.

Smile correction is taken in account during the wavelength calibration, where a fluorescent light emitting a peculiar pattern of emission peaks is used, and a second order function is used.

Binning is used to reduce the resolution of the image sensor matrix and to reduce the pixel temporal noise (increasing the SNR). Taking as example the VisNIR system, the selected sensor has a resolution of 2048x1536 pixels, meaning that the spectrum contains 1536 samples; often this spectral resolution is not needed, since infrared spectroscopy deals with low-bandwidth spectra due the overtone smoothing effect.

4.1.3 Normalization

This step uses the two reference images of dark and white, to normalize the image values between 0 and 1. We have checked that the used sensors have a linear response, thus the used formula is shown in Equation 1, where I_{raw} is the raw image coming from the camera, I_{dark} is the image acquired in dark conditions and I_{source} is the image of a standard flat white diffusive target.

$$I_{norm}(x,\lambda) = -\frac{I_{darkraw}(x,\lambda) - I_{dark}(x,\lambda)}{I_{source}(x,\lambda) - I_{dark}(x,\lambda} \quad (1)$$

The two references images I_{dark} and I_{source} are obtained by averaging more images, to reduce effects due to the sensor temporal noise; we typically use 100 images. It is a good practice to allow some time for the sensor electronic and illumination source to reach a stable temperature. In laboratory conditions, or during the acquisition of data used in our experiments, dark and source references were taken at the beginning and end of the data collection. During operational industrial conditions, references are taken at the start of a shift and stored in files.

4.1.4 Dataset creation

We have chosen HDF5 as file format to store the hypercube data and additional metadata. This structured file format has many advantages: it handles in an efficient way very large files; it supports the main data types (integer with 8, 16, 32 bits, floating point numbers) as multidimensional array, which fits well with two-dimensional images and three-dimensional hypercubes; it has a simple and efficient C API, and it is well supported by environment like Python or Matlab; it implements a transparent lossless compression.

In our workflow, each file corresponds to a single sample, and it contains:

- thumbnail grayscale or color image of the sample;
- header file with information on the sensor (type of camera, resolution, parameters used during the acquisition, …);
- header file with information on the sample;
- wavelength calibration;
- hypercube data after normalization;
- [optional] hypercube as sensor raw signal
- [optional] dark and white image as averaged sensor raw signal

Figure 4.3 shows an example of meta-information stored for each sample, encoded as JSON string.

```
{
   "Device":{
      "Manufacturer":[
         "Microtec-Sony",
         "Specim"
      ],
      "Type":"VISNIR 400-1000",
      "SensorRows":1536,
      "SensorColumns":2048,
      "Bits":8,
      [...]
   },
   "Lens":{
      "Manufacturer":"Fujinon",
      "FocalLength":8,
      "Aperture":1.6
   },
   "Target":{
      "Distance":550,
      "Angle":0
   },
   "Light":{
      "Type":[
         "Goldene white",
         "Ushio 980"
      ],
      "Specs":[
         "2 x 2 x 8 Stripe Flickering",
         "2 x 2 x 8 (2 led) Stripe"
      ],
      "Distance":750,
      "Angle":30
   },
   "Algo":{
      "DeadPixelCorrection":true,
      "DarkSubtraction":true,
      "WhiteCorrection":true,
      "SmileCorrection":false,
      "XMirroring":false,
      "Thumbnail":"rgb"
   }
}
```

```
{
   "Notes":[
      "Microtec-Sony Specim",
      "halo"
   ],
   "Sample":{
      "Type":"Wood",
      "Variety":"Eucalyptus",
      "Origin":[
         "Urufor",
         "Uruguay"
      ],
      "Supplier":"Oliver Kier",
      "Pretreatment":"None",
      "Condition":[
         "Dry",
         "Raw"
      ],
      "Description":[
         "Heartwood",
         "Sapwood"
      ]
   },
   "System":{
      "Machine":"GLD600",
      "Location":"Brixen"
   },
   "FileVersion":"0.3",
   "Software":{
      "Name":"Python",
      "Version":[
         "3.7.6 (tags/v3.7.6:43364a7ae0",
         "Dec 19 2019",
      ]
   },
   "DateCreated":"2020-03-16T11:47:03.515516",
   "DateModified":"2020-03-17T10:37:21.375699",
   "UUID":"2020-03-17T10:37:21.375699@GLD600"
}
```


Figure 4.3. Example of header containing meta-information.

For every data file, it is possible to have one or more file containing labeling information. The type of information is related to the result of the model, that is to make a pixel-wise classification in classes or values (like the segmentation of wood boards in heartwood or sapwood, or the segmentation of fruits in different ranges of humidity), or to make an estimation of some global parameter (like the estimation of the dry matter or maturity level in fruits).

In the first case, the required information is contained in JSON files that contain regions in the spatial coordinate of the hypercube with the corresponding class or value.

In the second case, each sample corresponds to one of few numerical values, and is stored in tables as couple { filename, value(s) }.

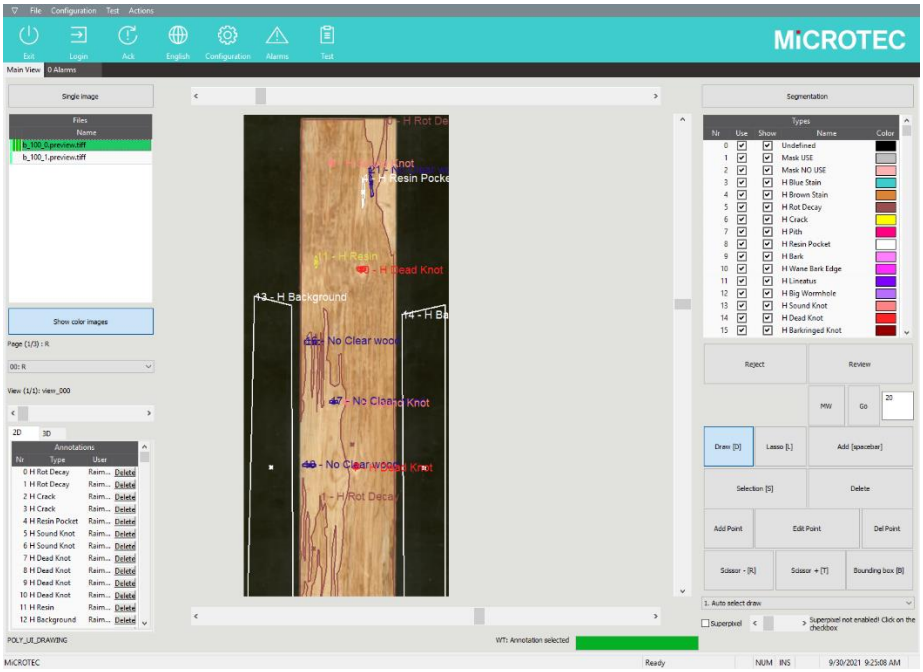


Figure 4.4. Screenshot of labeling software.

4.1.5 Sampling of spectrum

Working with hyperspectral data has the advantage of having many spectra, due to the fact that every pixel is in fact a spectrum. In facts, millions of spectra are available for every sample.

For practical reasons, since the spectral information is abundant, it is possible to subsample the spectra without performance degradation in the models with the advantage to reduce the computation time of the PLS algorithm. The following strategies were normally used:

- Fix % of spectra from the whole dataset;
- Fix number or % of spectra from every single labelled region;
- Fix number or % of spectra from every single sample, possibly containing more labelled regions;

The second strategy was normally used, since it permits to keep a more balanced distribution of the different classes.

4.1.6 Modelling

This step involves the creation of a model that convert one (or possibly more) input spectrum into a value (regression) or a class (classification). The chosen method is partial least squares (PLS). Wold et al. [14] gives a very useful review of this technique.

As mentioned in section 2.6, PLS modelling works by attempting to maximize the covariance in orthogonal space between x variables and y values, building a latent space with a lower dimensionality. From a practical point of view, having the set of measured spectra X, a (N*K) matrix of N observations of K variables (the predictor variables), and the set of target values Y, a (N*M) matrix of N observations of M target values (the response variables), the PLS method finds a new set of A (A<K) variables, the latent variables, to obtain the best approximation of Y; in the above notation, the N observations correspond to the two spatial axes in the hypercube, while the K variable corresponds to the spectral axis in the hypercube, that is the number of wavelengths. The important output of the method is the PLS regression coefficients B, a (K+1*M) matrix which is actually the trained model, that is used to compute an estimation of the target variables from the observations during the inference.

$$Y' = [1 \mid X] * B + Y_{residuals} \qquad (2)$$

An important assumption in model development is the number of PLS components selected to build the model; by increasing this number, at some point, the model will stop from describing the main features in the data and will start to try to describe the noise, causing over-fitting in the model.

This can be achieved using cross-validation when building the model. Common methods for internal cross validation are the leave one out (LOO) method and the k-fold cross validation (CV).

In k-fold cross validation the data are divided into k equally sized segments, commonly referred to as folds. Subsequently k iterations of training and validation are performed so that for each iteration, one fold of data is removed for validation while the remaining k-1 folds are used for training the model. The numbers of components that give the lowest error in the validation are considered optimal for the model.

After a model has been optimized, its performances are typically evaluated considering three parameters: coefficient of determination for cross validation (R^2cv), root mean square error of cross validation (RMSECV) and residual predictive deviation (RPD).

$$R^2{}_{cv} = 1 - \frac{\sum_{i=1}^{np} \left(y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{np} \left(y_i - \hat{Y}_m\right)^2} \quad (3)$$

$$RMSECV = \sqrt{\frac{\Sigma_{i=1}^{np} \cdot (\hat{y}_i - y_i)^2}{np}} \quad (4)$$

$$RPD = \frac{SD}{RMSECV} \quad (5)$$

Better models have higher R^2, because it explains the accuracy of model predictive capacity in cross validation, and lower RMSECV, because the values estimated by the model are on average very close to the measured values. RPD value indicates very good models (RPD>2), fair models (1.4 < RPD < 2) and non reliable models (RPD < 1.4).

4.1.7 Model inference

This step involves the inference of the model, in our case using the PLS method it simply consists in a simple matrix multiplication, see Equation (2).

The computation requires (N*K+1*M) multiplications. Taking the experiment reported in chapter 5.1 on heartwood/sapwood classification as a concrete example, the VIsNIR hyperspectral system uses a sensor configured to a resolution of 512x320 pixel and can run at 500 Hz. In a push broom hyperspectral configuration, this corresponds to N = 512*500 = 256.000 spectra per second, each spectrum containing M=320 samples, thus the amount of required multiplications for a PLS model to classify the two target classes is approximately 165 million of multiplications per second, a computational load that can be easily achieved by a modern CPU. Table 3.1 summarizes the computation for the VisNIR and SWIR hyperspectral system, in the typical configuration.

Note that the number of operations per second remains approximately constant when the camera framerate is increased by using the partial readout described in section 3.2 using a compact or a noncontiguous region of interest.

Description	Unit	VisNIR	SWIR
Resolution X (spatial axis 1)	Pixel	512	640
Resolution Y (spectral axis)	Pixel	320	512
Frame rate (spatial axis 2)	Hz	500	350
N (predictors)	#/s	256.000	224.000
K (Wavelengths)	#	320	512
M (target variables)	#	2	2
Multiplications	*1e6 / s	164	230

Table 4.2. Number of multiplications per second for PLS inference.

4.1.8 Aggregation of spectrum and post-processing

This step concerns some optional processing that can be applied to the single spectra before they are used for the inference of the model, or to the result of the inference.

Aggregation of spectra is relevant in the case of the estimation of a global variable, where a pixel-wise classification is not required; examples of this case will be presented in sections 5.2 and 5.3, where global quantities like dry matter or sugar content will be estimated in fruits. Instead of computing the simple average of multiple spectra (like a spectrometer would do), a method to exclude bad-spectra has been developed; bad-spectra are for example spectra too dark or saturated, due to the fact that the fruit surface is not smooth but can be wrinkled and can present local shadows or point of reflections. By eliminating these spectra before averaging, it is possible to obtain a better model correlation.

Post-processing is relevant in the case of pixel-wise classification, where the spectrum of each spatial pixel is used to compute a class or value. In some case, a standard 2D filter as the average or median is used to regularize the results and remove isolated outliers. It is also possible to get this result by applying the same filtering before the inference step, and to regularize by averaging the spectra.

5 Applications

This chapter reports different applications of the developed hyperspectral imaging technology[3]. The first application is an application of segmentation of different regions on wood boards. Then, other two applications on fruits for the estimation of internal quality parameters will be presented. Finally, other uses cases that were explored, using the hyperspectral technology but also other techniques, will be briefly presented.

5.1 Detection of heartwood / sapwood on Eucalyptus boards

5.1.1 Introduction

Wood is a heterogeneous material formed by a set of cells with specific properties to perform the main functions of water conduction, storage of biochemicals and mechanical support of the plant body.

Wood of most trees can be divided into two distinct regions in terms of their physiological activity: sapwood and heartwood. Sapwood, also called alburnum, is the outer, living layer of trees, necessary for the transport of water and minerals between the roots and the leaves. Heartwood, also called duramen, is the dead, central wood of trees, mechanically strong and resistant to decay. These regions are identified in many species, although their occurrence, properties and color can vary [51].

[3] This part of work was done in cooperation with other people. I was in charge for the design of all the tests, the evaluation methods and procedures. Simone Faccini and Matteo Caffini worked on the data acquisitions, labelling and evaluations of the data. Section 5.3 was part of a master thesis of Ilario Chini where I was coadvisor.

Figure 5.1. A section of a tree, showing the heartwood and sapwood.

There are many differences between the heartwood and sapwood, which may be relevant depending on the use of wood. Heartwood contains more extractives than the sapwood [52] (Miranda2006), higher lignin content [53] (Lachenbruch2011), and generally, less cellulose and holocellulose [54] [55]. The heartwood has a lower moisture content due to reduced physiological activity; the heartwood is less permeable and has more compact tissue than the sapwood [56](Burger and Richter 1991).

Wood density is affected due to the extractive contents of heartwood [57] [58] [59] [60], however even its strength is correlated with the density, sapwood does not differ structurally from heartwood [61]. The heartwood usually has a higher natural durability due to the absence of nutritious materials (carbohydrates, mainly in the form of starch), and especially to the presence of extractives [62] [63] [64].

These anatomical and functional differences involve different behaviour of heartwood and sapwood from both physical and chemical point of view. Wilkes [65] reports that while the sapwood is used for the production of pulp for paper (low content of extractives), heartwood is preferred in the furniture industry, where high requirements in finishing are necessary, due to its characteristics of greater natural resistance.

The differentiation of these regions is possible in some species by direct visual analysis of the wood due to pronounced differences in color between sapwood and heartwood caused by the accumulation of extractives. However, in certain species there are little or almost no visually detected differences between heartwood and sapwood colors. In such cases, identification requires the analysis of differences in the chemical level (different pH between the sapwood and heartwood [51] [66] [67] [68] [69] or the observation of anatomical features [70].

5.1.2 Material and methods

This study explores the classification of heartwood and sapwood on hardwood rough boards, cut from Eucalyptus trees grown in Uruguay.

A first explorative test considered a set of 13 small samples, with size of 40 cm in length, 3.5 cm in thickness and 10 cm in width. The main faces were analysed and divided in the two classes, heartwood and sapwood, using markers to facilitate the labelling phase in the acquired images. Figure 5.2 left shows a few examples of these boards, with the visible white markers. Looking at the picture, it is visible that in this specie of hardwood the colour information is not enough to provide a clear hint of the boundaries between the central heartwood from the lateral sapwood, since the wood colour does not have an evident difference and because of the presence of strongly coloured stains.

The measure was performed at the Microtec laboratory, using the two hyperspectral imaging systems in the SWIR (900-1700 nm) and in the VisNIR (400-1000 nm) wavelength range, in reflectance mode. The illumination used in this test consisted in two rows of halogen lamps, in order to have a relative white spectrum, perpendicular to the conveyor system.

The measured data were corrected with a dark and white reference and transformed in an absorbance spectrum; the images were then labelled, using polygons to describe the two classes of heartwood and sapwood; the spectra of the two classes were used to train a PLS classification modelling. The model was then applied on the full hypercube, giving the very satisfactory qualitative results reported in Figure 5.2 right.

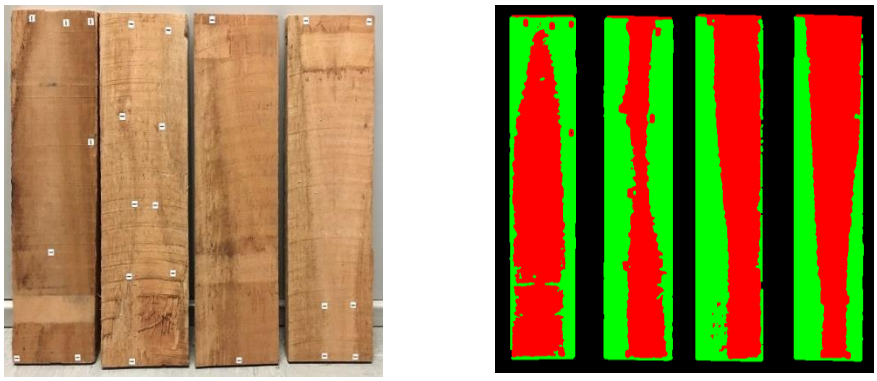


Figure 5.2. Example from the set. Left: picture of the rough boards, with the visible markers to guide the labeling. Right: qualitative results of the heartwood/sapwood classification using the SWIR wavelength range.

After the initial positive pre-study, a more consistent test was repeated. In this second test, a set of 27 Eucalyptus boards from Uruguay was considered. These boards were of bigger size, with an average length of 4 meters, thickness of 3.5 cm and width of about 30 cm. The wood surface was not planed (rough) and often coloured stained stripes was present.

Since the first study gave similar results for the two hyperspectral systems in the VisNIR and SWIR wavelength range, we decided to limit the second test to the 400-700 nm wavelength range only and to use a LED illumination because it was the easiest and simplest configuration to be added into the existing scanner for wood boards from Microtec. The measure was performed in the Microtec laboratory, where each board was measured on both sides with the hyperspectral imaging system in the VisNIR range using LED illumination in the visible range only.

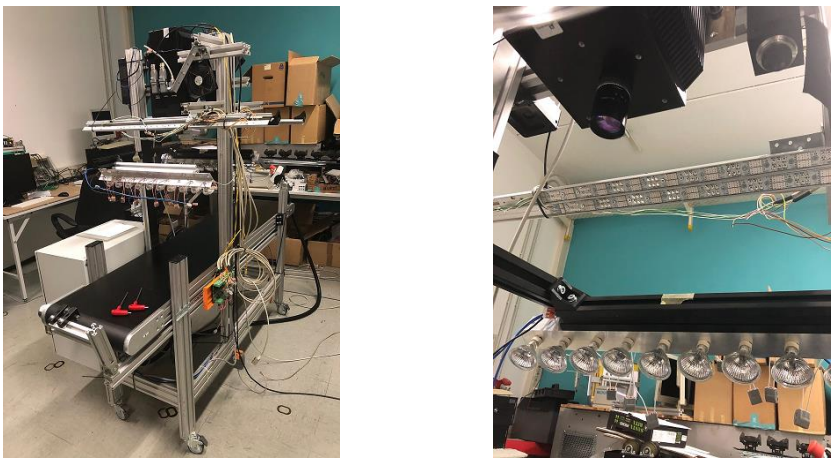


Figure 5.3. Picture of the acquisition setup.

With the help of an expert, 47 sides were labelled as densely as possible with polygons describing the two classes of interest, sapwood and heartwood. Where the transition between the two regions was not clear, a gap between the areas was left unmarked. Moreover, some regions like large knots were left unmarked.

Figure 5.4. Example of a labeled board.

Figure 5.4 shows an example of a labelled board, where red is used for heartwood and green for sapwood. One can note the unmarked region around a big know, and the presence of vertical stains and dirt in general on the surface.

To store the labels, we used a simple JSON file, containing the list of different classes (in this case: heartwood, sapwood and background), and the list of the labels, each containing the belonging class, the coordinates of points, the type of connectivity (sparse dots or closed polygon), and the name of the expert. Figure 5.5 shows a portion of the description corresponding to the same board in Figure 5.4.

```
{
  "tag": [
    { "name": "heartwood",
      "description": ""
    },
    { "name": "sapwood",
      "description": ""
    },
    {"name": "background",
      "description": ""
    }
  ],
  "roi": [
    {
      "tag": "sapwood",
      "x": [ 405, 356, 348, 350, 356, ... ],
      "y": [ 238, 246, 523, 923, 1285, ... ],
      ],
      "connection": "poly",
      "user": "mario rossi"
    },
    {
      "tag": "heartwood",
      "x": [ ... ],
      "y": [ ... ],
      "connection": "poly",
      "user": " mario rossi "
    }
  ]
}
```


Figure 5.5. Example of JSON label file.

The total number of labelled spectra is approximately 20.4 million of spectra. The distribution of sapwood and heartwood is not equal, in particular there is a majority of heartwood (83%) compared to sapwood (17%). The majority of boards, about 60%, consists of heartwood only, a few (10%) are mainly sapwood, and the rest have a mix of both regions.

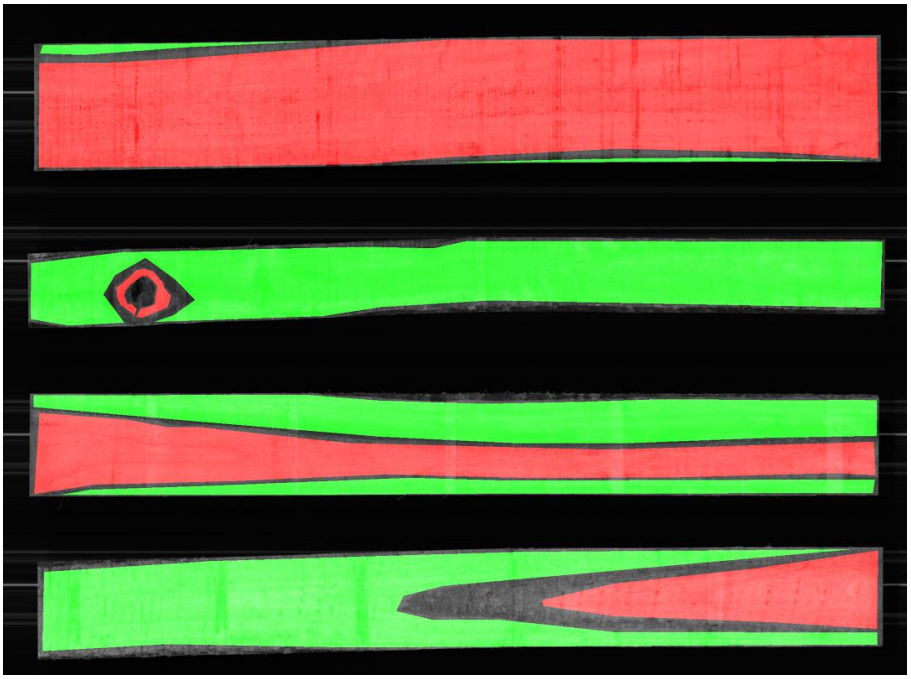


Figure 5.6. Example of heart/sapwood distribution on boards.

Table 5.1 summarizes the statistics of the dataset.

Number of wood faces	47
Hypercube size	512x2500x320
Size of dataset	20.6 GByte
Spatial resolution	512x2500 pixel
Spectral resolution	320 wavelengths
Number or labels	170
Number of labelled spectra	20.400.000 (17% sapwood, 83% heartwood)

Table 5.1. Statistics of data.

5.1.3 Data analysis

The spectra stored by the data acquisition software were converted in reflectance spectra by correcting them with a dark and white reference, which were taken at the beginning and end of the acquisition process.

Figure 5.7 shows the average and standard deviation for the two classes to distinguish, sapwood and heartwood. The two ends of the spectrum, where the signal is very low, are put to zero, to avoid the injection of noise.

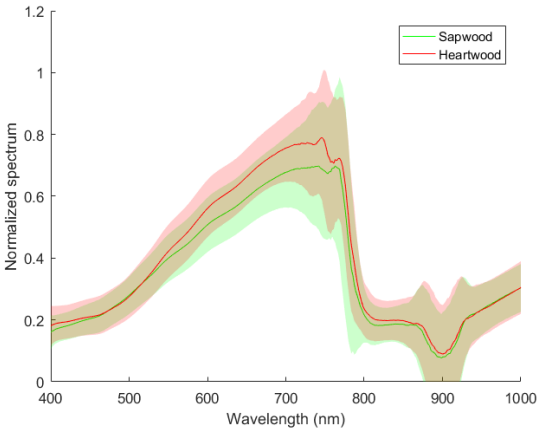


Figure 5.7. Spectra of all boards.

The spectra were then extracted from the hypercubes using the labels, but only a fraction of the spectra were used for the model calibration. Even if some millions of spectra are available, it has been observed that a few thousands are enough to compute a model.

A few different techniques were used to sample the spectra: uniform sampling of all spectra, sampling of a fix number of spectra for each board, sampling of a fix number of spectra for each labeled area, and sampling of a number of spectra for each label proportional to its area.

Each technique results in very similar results, thus it was decided to extract 1000 spectra randomly sampled from each labeled polygon, for a total of about of 120.000 spectra. Each spectrum is associated with the class of the original label.

The model was computed using the partial least squares discriminant analysis (PLS-DA), commonly used for the task of classification. The number of latent variables in the PLS was optimized between 1 and 20 using k-fold cross validation with k=10 and accuracy as performance indicator.

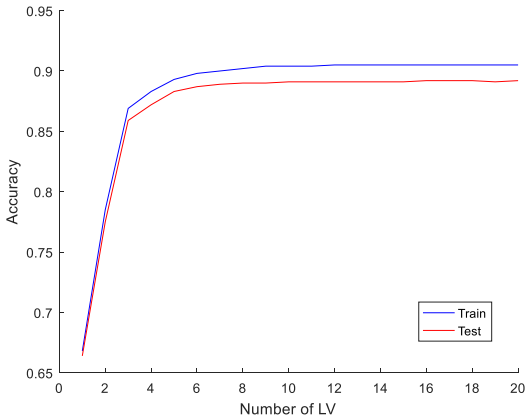


Figure 5.8. Accuracy vs Number of PLS latent variables.

At about LV=8 a plateau is reached, thus it was decided to use 10 latent variables. In addition it was observed that there is no overfitting, because the accuracy does not drop by increasing the number of variables in the PLS. This is probably due to the already very high number of spectra that are used for the model estimation.

5.1.4 Results and discussion

Table 5.2 shows the confusion matrix for the classification of heartwood/sapwood; the F1 score was F1=0.916.

		predicted	
		Sapwood	Heartwood
actual	Sapwood	91.97%	8.03%
	Heartwood	12.00%	88.00%

Table 5.2. Confusion matrix of the classification of heartwood/sapwood.

The model was applied to the entire set of spectra contained in the labeled regions (for training it was used only a subset of all spectra, as explained previously), obtaining the same performances. I tested a morphological filter to attenuate the local noise in the classification, both on the input spectra (a pre-processing step, where the spectra are averaged on a 3x3 or 5x5 pixels window before applying the PLS model), and on the output classes (a post-processing step, again on a 3x3 or 5x5 window), obtaining a marginal improvement.

The model was finally qualitatively evaluated on the full hypercube, taking in account the whole boards, in order to evaluate the discriminative power of the sapwood and heartwood regions outside the labeled areas.

Figure 5.9 shows the results of three boards. Each example shows the pseudo colour image obtained from the full spectrum, the labelled regions, and the result of the pixel-wise PLS classification. Here non pre/post processing was applied.

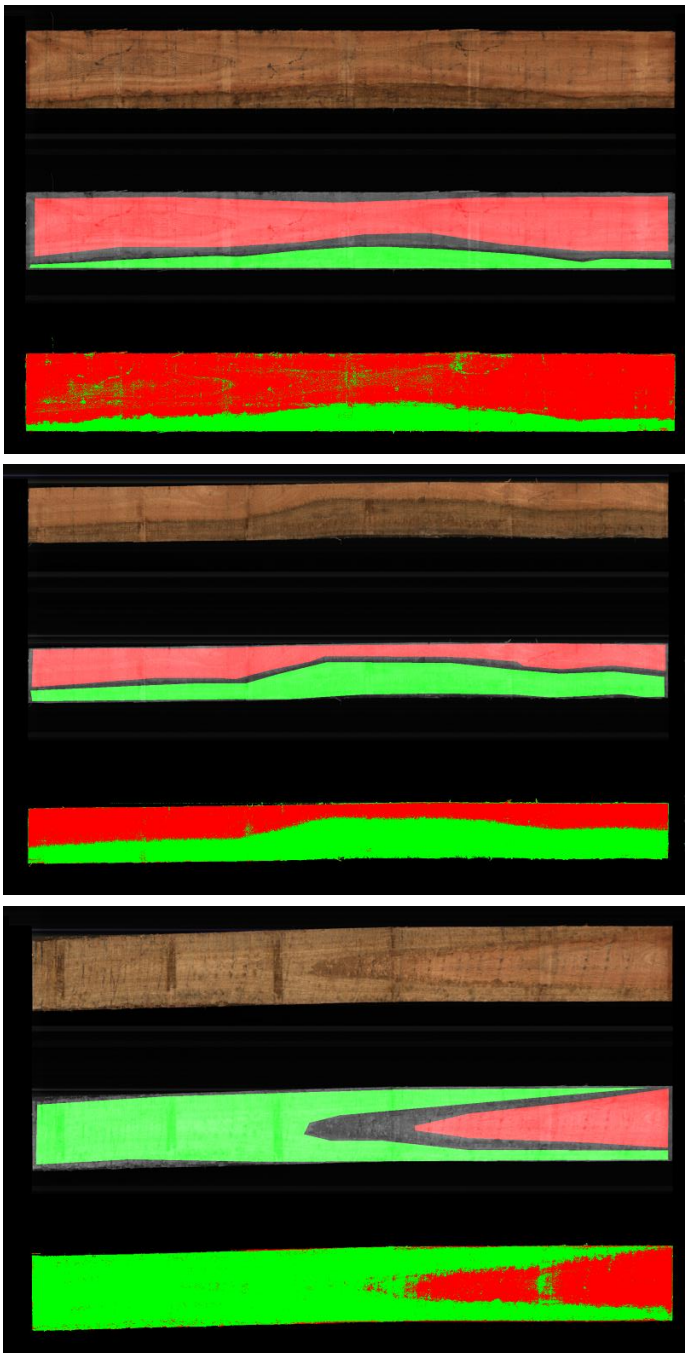


Figure 5.9. Classification of the whole board. Pseudo color image (top), ground truth labels (middle), result of pixel-wise PLS-DA classification (bottom).

A direct comparison with the literature is hard because of the lack of studies on the same subject. The most similar studies have been made by Thumm and Colares.

Thumm et al. [41] [42] [43] used a hyperspectral imaging system in the SWIR range (900-1700 nm) from Specim to scan discs of wood to visualize the distribution and variation of lignin, galactose and glucose in the samples, which are related to the chemical composition of wood. They also investigated shooks of radiate pines in order to detect resinous defects; their model was able to further well discriminate resin from other visually similar defects like sapstain, pith or kiln brown stain. All these studies used PLS as a mathematical tool to compute their models.

Colares et al. [50] applied SWIR (1200nm to 2500nm) hyperspectral imaging for the determination of the distribution of holocellulose (cellulose + hemicellulose), lignin and extractives on samples of Mahogany of very small size (approximately 1 cm per side).

However these reports do not report enough numerical values on the performances of the classification or segmentation tasks, thus a direct and precise comparison is not possible.

5.1.5 Installation in an industrial sawmill

Given the promising results, it was decided to install the hyperspectral equipment in a real industrial scenario. In particular Microtec installed a first scanner in Uruguay and six scanners in Germany, where two hyperspectral cameras and the illumination lamps based on LEDs were added to the existing sensor equipment, for the detection on the top and bottom surfaces of wood boards. All systems were used for the detection of sapwood and heartwood, on Eucalyptus wood (Uruguay) and Oak (Germany). Figure 5.10 and Figure 5.11 show pictures of the two installations.

Figure 5.10. Microtec scanner in Uruguay.

Figure 5.11. Two of Microtec scanners in Germany.

5.2 Estimation of Dry Matter in avocado fruits

5.2.1 Introduction

Avocado is a plant originally from southcentral Mexico; its fruit is botanically a large berry containing a single large seed [71].

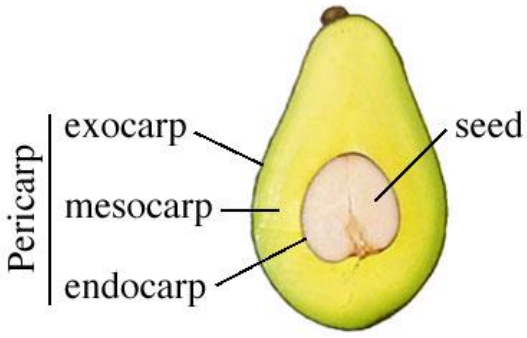


Figure 5.12. Avocado

Avocado is often called a superfood due to its nutritional density and health benefits. Avocados contain vitamins A, B, C, E, and K, plus 25 essential nutrients. The fruit unsaturated fats help the stability of heart and reduce cholesterol. The versatility of the good goes beyond the nutritious advantages, avocado is a great source of vitamin E which is ideal for skincare, for example. Hence, apart from the food industry, avocado is reaching the pharmaceutical and the cosmetic industry.

During the last years, the consumption of avocados in the world has been continuously growing. The global avocado production is about 7 million tons (which is worth of approximately 9 billion US dollars). Latin American countries are at the top in the production and exportation of avocado; Mexico is the top world producer of avocado (approximately 2.3 million of tons in 2019), followed by Peru, Dominican Republic and Colombia). The main importers are the US and Europe with 45% and 39%, respectively. In the Mediterranean region, Israel is the biggest producer with its 100.000 tons in the 2019, while Italy has a small production mainly in Sicily, Apulia and Calabria.

Avocado is a climacteric fruit (like apples, bananas, tomatoes, and others), that is it matures on the tree but ripens after being harvested. In order to ripen properly, avocados must be mature. Commercial growers aim to harvest fruit at "harvest maturity", which allows the fruit to ripen off the tree. The best tasting fruit would be ripened on the tree,

as they would have the longest time to accumulate sugars and starch, but would have no allowance for transport and shelf life. Because of this, growers try to optimise the balance between on-tree ripening and transport/shelf-life, which means achieving physiological maturity, picking fruits when they are hard and keeping in coolers until they reach the final destination for commercialization. At room temperature avocados ripen in one or two weeks (depending on the cultivar and on the maturity level), or faster if they are stored in the presence of ethylene gas (artificially or close to other fruits like apples or bananas). Therefore, [72] [19] maturity is an important quality index for avocado, since it is important to harvest mature fruits, to ensure that they will ripen properly and have acceptable eating quality.

Currently, commercial avocado maturity estimation is based on destructive assessment of the percentage dry matter (% dry matter or DM%) and sometimes percent oil, both of which are highly correlated with maturity [36] [73].

Different guidelines for the determination of export quality (OECD) recommend a minimum maturity standard for its growers of 23% dry matter (greater than 10% oil content) for 'Hass' avocados, although consumer studies indicate a preference for at least 25% dry matter [74].

A rapid and non-destructive system that can accurately and rapidly monitor internal quality attributes (in this case % dry matter) would allow the avocado industry to provide better, more consistent eating quality fruit to the consumer, and thus improve industry competitiveness and profitability.

NIRS has been demonstrated to be an accurate, precise, rapid and non-invasive alternative to wet chemistry procedures for providing information about relative proportions of C-H, O-H and N-H bonds. Analysis of NIRS absorption spectra aids in the qualitative and quantitative determination of many constituents and properties of horticultural produce, including oil, water, protein, pH, acidity, firmness, and particularly soluble solids content or total soluble solids of fresh fruits [75] [76] [77]. Moreover, NIRS has been used to estimate % dry matter in various horticultural products [78] [79] [80] [81] including avocados [36] [19] [82].

It is known [83] that in avocado fruits, the dry matter content is not evenly distributed but presents relevant differences. The fruit, while having a general symmetric shape and an elongated main axis (peroid), is not symmetric in terms of stem position, which is

not on top of the fruit but shifted laterally. In parallel with this asymmetry, we have observed that the dry matter content on the stem face is in average higher than the opposite face.

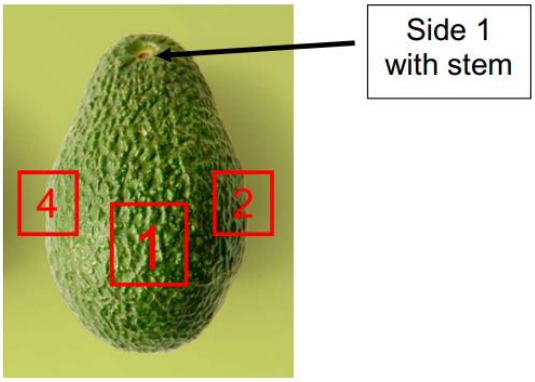


Figure 5.13. Definition of faces.

5.2.2 Material and methods

In this study, two sets of 25 Hass avocado fruits (50 fruits in total) provided by a fruit importer are considered. Between the two sets approximately 4 weeks passed (26/10/18 and 22/11/18). Each fruit was marked with an identification number, and the four sides were labelled according to the convention depicted in Figure 5.13: side 1 is where the stem leans more towards the lower part of the fruit, with the other sides 2,3,4 are east-to-west having the stem at the north pole.

The non-destructive measure was performed at the Microtec laboratory, using the hyperspectral SWIR (950-1750 nm) imaging system in reflectance mode. The illumination used in this test consisted in several halogen lamps, in order to have a relative white spectrum (compared to a LED system) and high flux; however in order to evaluate the possible utilization of LEDs, the spectra of a LED illumination lamp were simulated by cropping the full halogen lamp spectrum. Each fruit was passed 4 times in the measure system, having each side on top. Moreover 5 fruits were measured 5 more times on the same side, in order to have a measure of repeatability.

Table 5.3 summarizes the statistics of the dataset.

Number of fruit faces	200
Hypercube size	320x256x350
Size of dataset	11.6 GByte
Spatial resolution	320x350 pixel
Spectral resolution	256 wavelengths

Table 5.3. Statistics of data.

The destructive analysis was performed with the flesh peeling method, taken from the four sides on the equatorial region, with area of about 10 cm2, approximately 90° apart, and processed with the oven method at 103°.

Figure 5.14. Picture of destructive analysis.

The four different points of measurement around the equator were consistently labelled, and the corresponding dry matter DM% was individually evaluated. Figure 5.15 depicts the relative distribution. The internal variability in each fruit is likely up to +/-2%. This is again an indication that by measuring only one side of the fruit and by not taking in account the internal variability, the final measure will have a big deviation from the average value.

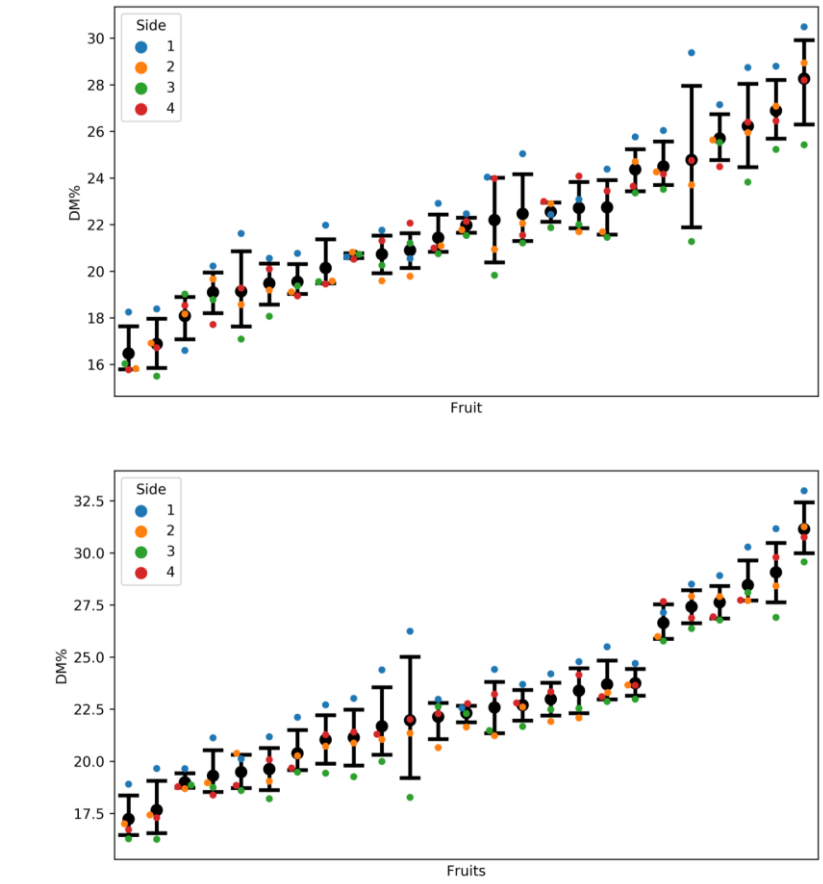


Figure 5.15. Dry matter distribution.

5.2.3 Data analysis

The analysis of the measured data consists of a first step where each spectrum is corrected with dark and white references and transformed to a reflectance spectrum. The fruit where then located in the obtained reflectance image, and a rectangular or circular region corresponding to a surface of approximately 10 cm^2 was selected (Figure 5.16). After the removal of saturated or very dark spectra (corresponding to specular reflections or under-illuminated areas), a single spectrum was obtained by averaging all valid spectra of the four sides of each fruit. These spectra were then used for the PLS regression modelling.

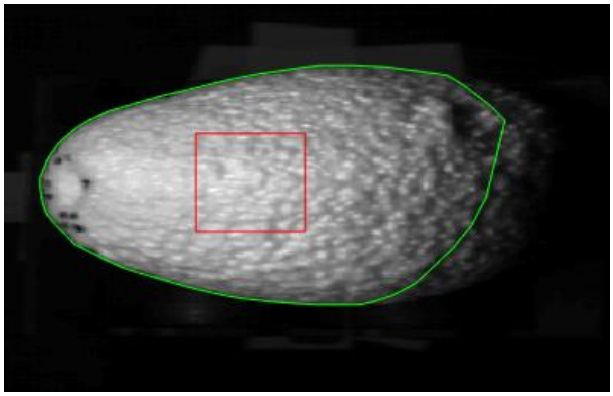


Figure 5.16. Segmentation of the fruit (green) and the region used for the evaluation (red).

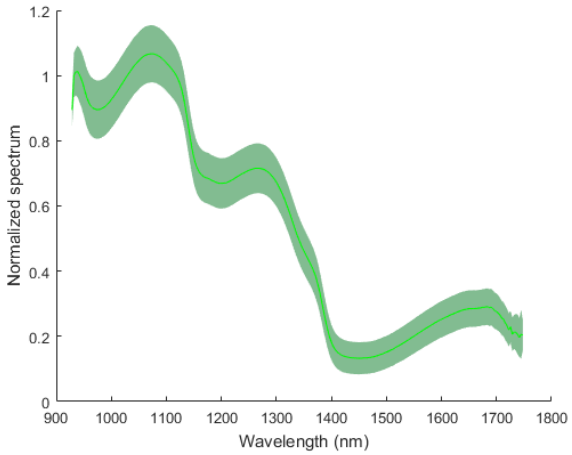


Figure 5.17. Distribution of spectra of one avocado fruit.

The PLS models went under a randomization test in order to evaluate the prediction capabilities. The test requires the dataset to be repeatedly and randomly split in training and test sets. At each iteration the model is calibrated on the training set and the prediction performances are evaluated against the test set. The metrics used are the coefficient of determination R^2 and the root mean square error of prediction RMSEP. Figure 5.18 depicts the output of the test, with the predictions from the different iterations collected and displayed together.

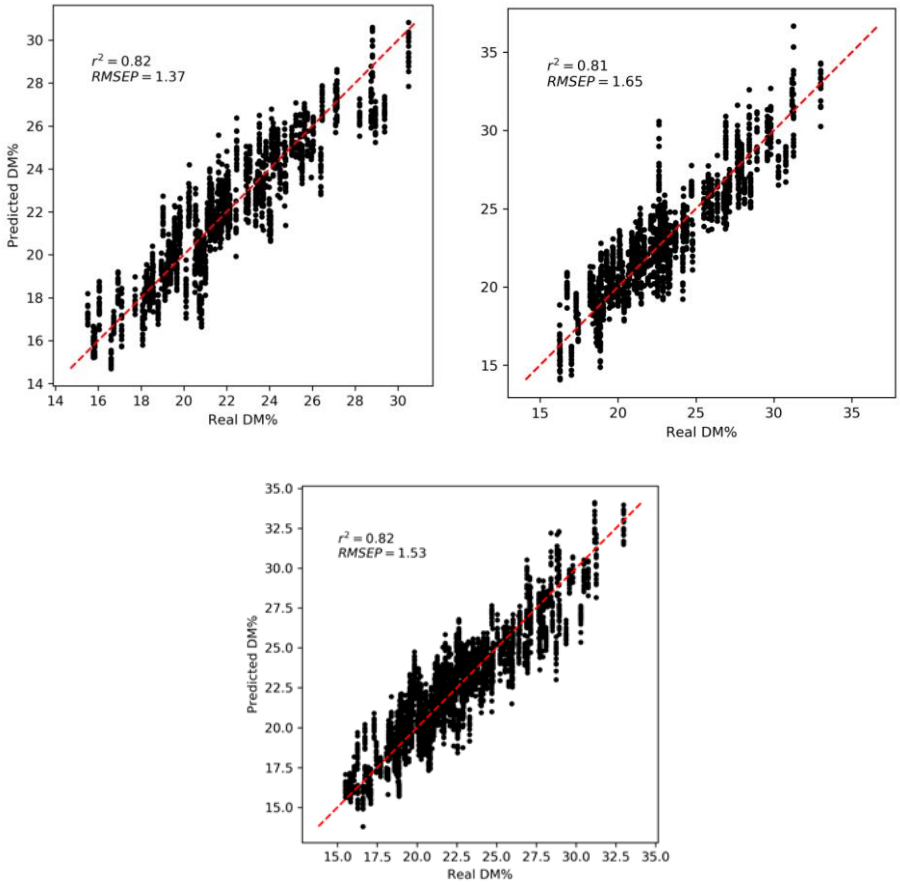


Figure 5.18. Top: results of the dry matter model, divided per batch. Bottom: same result with a unique model.

5.2.4 Repeatability

In order to assess the robustness and repeatability of the hyperspectral measure, 5 fruits were randomly sampled out of the second set and measured an additional 5 times on the same side, adding small variations in the position to introduce a small degree of variability in the measurement process. These additional measurements were pre-processed and computed using the PLS model on the second batch. The variability of the estimated dry matter values is reported in Table 5.4; the average deviation of 0.6 DM% is low enough and allows the use of this technique for a stable measure.

Fruit #	1	2	3	4	5
Stdev (DM%)	0.7	0.6	0.7	0.2	0.7

Table 5.4. Results of repeatability test.

5.2.5 Results and discussion

The results obtained in this project performs in a similar way as other studies in the literature. Table 5.5 reports an overview of the comparison.

It is interesting that the majority of the applications are based on spectroscopy, which cannot account for the inhomogeneity of dry matter on the fruit sides. Only recently ([84], [85]) there have been publications using the hyperspectral technology; however, these study were limited to laboratory conditions, while this study covers the implementation in industrial conditions. For example [84] analysed 21 avocado fruits on 4 cardinal points around the equatorial zone, in a similar way of this study; they used the average of all spectra obtained by each scan, thus obtaining a single spectrum from each measure, while in this study only the valid spectra (i.e. eliminating the too dark or ill-illuminated points) from the hyperspectral images are used for the average, thus giving some advantage compared to a pure spectroscopy method. In addition this study proposes a novel approach to measure the fruits on a 360° equatorial band, in order to strongly reduce the errors due to the uneven DM% distribution.

Method	Wavelength (nm)	R^2	RMSECV	Source
Hyperspectral	950-1700	0.82	1.53	This study
Spectroscopy	300-1140	0.88	1.80	[36]
Spectroscopy	300-1100	0.79	1.14	[82]
FT Spectroscopy	830-2500	0.76 0.93	1.53 1.48	[72]
Spectroscopy	400-2500	0.84	2.38	[86]
Hyperspectral	400-1000	0.96	1.35	[84]
Hyperspectral	400-1000	0.90	2.60	[85]

Table 5.5. Overview of the applications of hyperspectral imaging and spectroscopy to measure the dry matter content in avocado and comparison with the literature.

5.2.6 SWIR LED simulation

The present study uses halogen light as light source, because of its low cost, high flux and relatively white spectrum. However, halogen light has some drawbacks, like reduced lifetime, heat production, necessity of some time to reach stability. Therefore a simulation of LED light source was conducted. The complete spectra obtained with halogen light was weighted to functions having a shape similar to the spectra of real

commercial SWIR LEDs at different wavelengths, from 940 to 1650 nm. Various combinations of one, two or three wavelengths were used with a grid search, in order to find the best combination.

The best combination was to use wavelengths at 1050, 1200 and 1300 nm, which resulted in the regression shown in Figure 5.19. The general performances and the error prediction are very similar to the halogen light, which means that a LED light source could be used instead of LEDs.

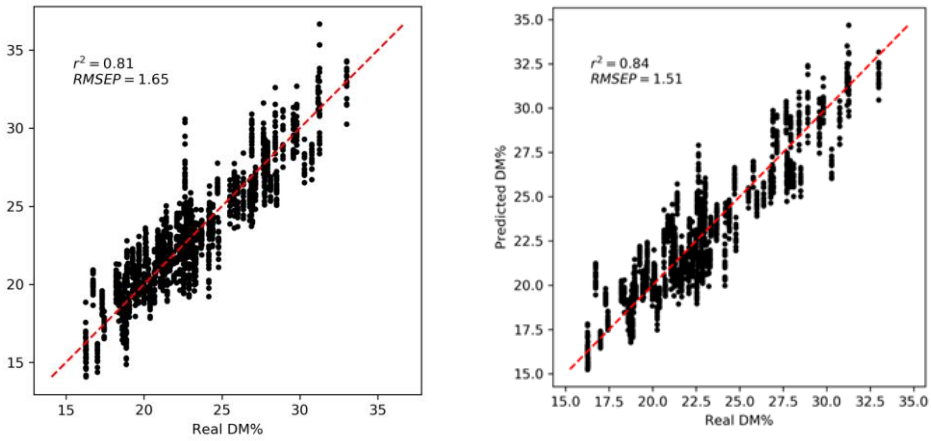


Figure 5.19. Left: results of the dry matter model on batch 2 with halogen light. Right: with simulation of LED illumination.

5.2.7 Real scenario implementation

In order to implement a hyperspectral imaging system for the determination of the dry matter content in a real case scenario, it is necessary to take in account the uneven distribution of dry matter that requires to scan the fruits on the whole surface with the constraints given by the mechanization of the sorting lines.

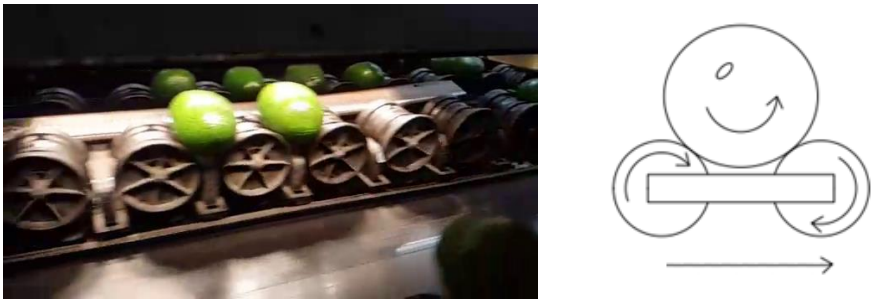


Figure 5.20. Picture of the transport line.

In sorting lines, the avocado fruits are moving in cups that move forward and have a mechanism that put the fruit in rotation; this is done by having a series of rolls pulled by a chain to obtain the forward motion and sliding over a second chain to obtain their rotation; by controlling the speed of the two chains it is possible to obtain the desired forward and rotation speeds. In practice these cups advance at about 1 meter per second or about 10 cups per second. In addition, the fruits make a complete rotation in about 0.5 meters (depending of their diameter).

In order to scan the whole surface, since the fruits are rotating, one could use a single shot hyperspectral system, or use multiple line scanning (push broom) systems, placed at a distance corresponding to a rotation of 90°; unfortunately, both methods are complex and very expensive.

To overcome these limitations, we have devised a configuration that uses a single line scanning system and is able to cover an equatorial stripe over the entire avocado diameter. This approach is described in patent [87].

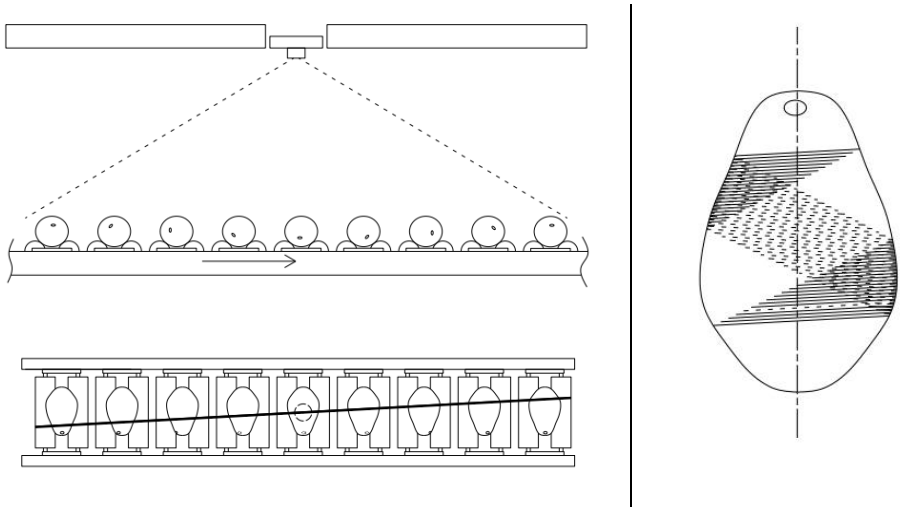


Figure 5.21. Left: lateral and top view of the patented arrangement. Right: the surface on a fruit covered with this arrangement.

The arrangement is shown in Figure 5.21 left: the line observed by the push broom system is aligned almost parallel to the fruit's movement; there is a small angle of about 5-10°, so that the combination of the forward motion and rotation scan an envelope that corresponds to Figure 5.21 right.

The coverage is not the whole surface, but still permits to cover the full 360° equatorial band, thus permitting to obtain an average measure, thus obtaining results that are comparable to the average of the four faces as in the laboratory setup.

5.2.8 Installation in an industrial sorting line

Considering the results obtained in laboratory, the good comparison with results reported in the literature, and the advancement for an industrial implementation given by the acquisition speed and the coverage of the complete surface of the fruit, this configuration has been built and a prototype was mounted at a packing house in the northern part of Israel. Testing was started in order to evaluate the long-term stability of the model for the determination of dry matter.

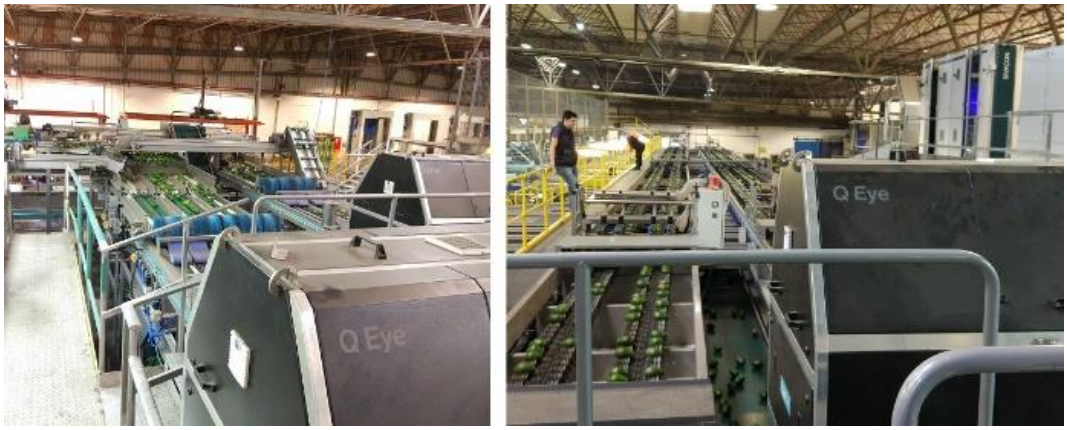


Figure 5.22. Picture of the prototype installed in Israel.

5.3 Estimation of soluble solids content (sugar) in Citrus

5.3.1 Introduction

Citruses are hesperidia produced by trees and shrubs belonging to the family Rutaceae. These fruits share a common structure composed of an exocarp called flavedo, a mesocarp called albedo made of a white spongy layer, and an endocarp divided into locules filled with juice vesicles [88].

Citruses are non-climacteric fruits; indeed, their ripening process stops after harvest, therefore they can only be picked and marketed when an adequate maturity has been reached [89]. In addition, their sugars and acids content is quite stable after harvest. The most important citruses for commercial relevance are oranges, mandarins, clementines, satsumas, tangerines, lemons and limes.

Citrus fruits have a recognised role in providing nutritive and therapeutic value, well known for their content of vitamin C and other nutraceutical phytochemicals such as carotenoids, flavanones, limonoids, and vitamin-B, responsible for antiallergenic, anticarcinogenic and anti-viral properties. The consumption of citruses also provides pectin and fiber, known to decrease the risk of heart attack when consumed regularly [90].

Citruses are nowadays the most produced group of fruits in the world. The global production was about 150 million tons in 2018, harvested in more than 80 countries. Asia is the most important citrus producing continent, its production, coming mainly from China and India covers 50% of world production and have faced a strong increase in the recent decades. America follows, producing 29% of world oranges, while Africa produces the 13% and Europe the 7%. In Europe the most important producing country is Spain, in which about 7 million tons are harvested every year, while Italy is the second most important producing region [91]. Concerning Italy, 62000 farms cultivate more than 100 varieties of citrus fruits on more than 129000 hectares, producing yearly about 2,5 million tons. Most of the Italian production comes from the south of the country, where citruses represent some of the most important agricultural products. As an example, Sicilian citrus fruit production in year 2018 was accounting for 630 million euro, corresponding to 13% of the whole regional agricultural production value [92].

The concept of quality is defined by ISO regulation 9000:2015 as the set of properties and traits of a product or a service linked to its capability to fulfil explicit or implicit needs. The quality concept is not univocal among entrepreneurs and final

consumers, but it assumes different meanings for different players of the supply chain, which apply different acceptability criteria for the same relevant parameters. For this reason, a correct quality definition must refer to a target group. Quality control process is based on the setting of quality standards, which are observed by quality inspectors to determine grading and marketability, which are essential parameters to set fruit price. In order to fulfil consumer quality requirements, food industry needs precise, effective and sustainable tools to define and evaluate customer-oriented product quality features [93]. The currently used quality parameters are many and distinguished according to different ways of classification. Quality evaluation methods involve the assessment of physical, chemical and physiological parameters. Physical parameters apply physic principles on fruits to measure responses. Measurements belonging to this group are for instance total soluble solids content (SSC), fruit firmness and fruit weight. Chemical parameters are quality attributes measured using chemical reactions involving fruit chemical compounds. Examples of chemical quality attributes are total titratable acidity (TA), ascorbic acid level and pH. In the end, physiological parameters like rate of respiration and ethylene evolution indicate physiological processes occurring inside the fruit. The assessment of quality attributes follows both subjective and objective methods. Subjective methods involve the response of human senses, while objective methods use instruments. A further classification distinguishes between destructive and non-destructive quality evaluation ways, depending on the fact that the fruit is damaged or remains intact during the analysis [88].

From the final consumer point of view, the quality of fruit can be assessed during purchase only considering external aspects, like size, colour, shape and absence of defects. Human minds link by means of experience, the appearance of a fruit to its flavour, which is a combination of taste and aroma and depends mainly upon internal chemical parameters. Citruses contain a set of biochemical compounds divided into primary and secondary metabolites, the most important are proteins, polyamines, amines, organic acids, carbohydrates, lipids, phenols, terpenoids, flavonoids, aromatic compounds, mineral elements, vitamins and hormones. All of them play a relevant role in the metabolism and physiology of citrus fruits, moreover many of these compounds are highly nutritive for human beings and are also responsible for fruit sensory properties. Respiration and photosynthesis are primary metabolic processes that produce and allocate into the fruit basic chemical compounds, which are partially stored and partially transformed into secondary metabolites. For example, carbohydrates, acids and amino acids are utilized to synthetize sugar derivatives, proteins, and volatiles [88].

According to the literature, the most important parameters used to assess citrus quality are SSC, TA, soluble solids to acids ratio (SSC/TA), brix minus acids (BrimA) and texture [94] [95]. Important correlations can be found in the literature between SSC level of citrus fruits and maturity stage, as well as consumer acceptance of the product [96]. For this reason, SSC is considered the most important parameter to assess overall quality of oranges and marketing standards [97] [98]. The total soluble solids content is not only related to soluble sugars, but also to all the other molecules soluble in fruit juice. The contribution of soluble sugars account in oranges for the 75-85% of SSC value, this predominant fraction is a mixture composed primarily of the monosaccharides D-glucose, D-fructose and the disaccharide sucrose, and secondly of xylose, mannose and various heptuloses. The remaining 15-25% of the value is affected by acids, proteins, fructans, minerals, dissolved vitamins, pigments and phenolic compounds [94]. The most common and practice way to determine SSC is with refractometry. TA is a quality parameter corresponding to the content of organic acids into the fruit. In oranges, organic acids are dissolved in cell sap, either free or combined with glycosides, esters and salts. Citruses contain also high amounts of cations like magnesium, potassium and calcium, which combine with acids and form other salts. In oranges, the most important and abundant acid is citric acid, and accounts for the 80-95% of overall acid content. The remaining 5-20% is mainly composed of Malic and Succinic acids. Titratable acidity of oranges and other citruses is expressed as percentage of anhydrous citric acid by weight or by volume. Its assessment is carried out by titration with NaOH 0,1 N till the equilibrium point is reached. Organic acids are weak acids, therefore, when a titration with a strong base is carried out, the equivalent point results slightly basic, due to the presence of salts [88]. SSC/TA ratio, the so-called maturity index, is another key parameter for citruses acceptability. It describes the blending of sugar and acids into the pulp, which is linked to fruit juiciness and ripeness degree, and is widely used as a maturity criterion for non-climacteric fruits [99] [100]. The higher is this ratio, the lower is the acid relative content of the fruit. Usually, when a fruit has a ratio higher than 19-20, it is not well accepted by consumers due to an excessive sweet and flat flavour. This ratio is highly informative only when presented together with TSS value, as the same ratio may result from different combinations of values [88].

A precise quality assessment of citruses must consider their specific characteristic, in fact citrus fruits are anisotropic fruits, since their physical and chemical properties change according to the direction of measurement. For example, SSC is on average

higher in the distal apex of fruit and decreases going towards the proximal one, while the opposite happens for titratable acidity (TA) [101].

Rapid, precise, accurate and cost-effective technologies are needed by fruit industry to sort products in order to satisfy the minimum acceptance level of every consumer segment [102].

5.3.2 Material and methods

In this study 120 oranges of variety Navel were considered. The fruits were bought in a local supermarket; their origin was from Italy; all fruits were without visible external defects and commercially mature.

Figure 5.23. Picture of the oranges used in the test.

The aim of the study was to analyse different methods to realize a machine for the real-time in-line non-destructive assessment of quality parameters of citrus fruits, in particular the soluble solids content (SSC) and total titratable acidity (TA).

The methods evaluated in the study were:

- Vis-NIR hyperspectral imaging in reflectance (wavelength range 400-1000 nm)
- SWIR hyperspectral imaging in reflectance (950-1700 nm)
- Vis-NIR spectroscopy in transmittance (400-1100 nm)
- SWIR spectroscopy in reflectance (950-1700 nm)

The motivation to use additional methods (spectroscopy in the Vis-NIR and SWIR range) was to test simpler methods in addition to hyperspectral. A spectroscopy device returns a mono-dimensional vector for each measure (compared to a two-dimensional image of a hyperspectral push broom device), thus the data amount is strongly reduced. On the other hand, this requires that the measure is optically aligned and collimated, and

does not permit more elaborated processing to eliminate not valid measures due to specular reflections, shadows or under-illuminated samples.

One important note is that, while the quantities to measure are internal parameters (the SSC and TA of the internal pulp of oranges), three of the methods measures only the surface of the fruit; the only technique that directly measure the inner fruit pulp is the Vis-NIT transmittance spectroscopy.

The non-destructive measurement took place at Microtec, where all fruits were measured with the four methods over two days (separated by one week). During the idle periods, the oranges were stored in a fridge at 6°C.

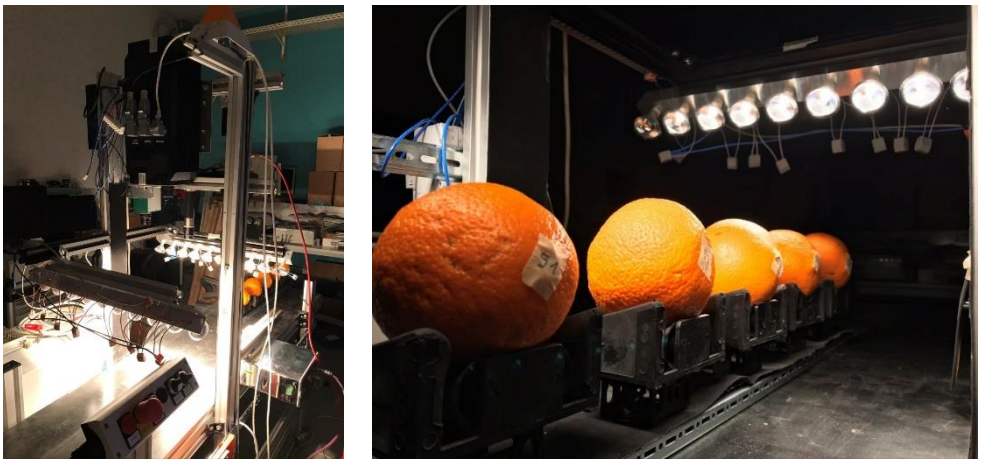


Figure 5.24. Left: picture of the acquisition setup. Right: details of the cups used to transport the oranges.

The measurement with the two hyperspectral imaging methods consisted in a dynamic acquisition of the light reflected by the fruits while moving on a conveyor. The light source was composed of two rows of halogen lamps working at 2950K with aluminium reflector, to preserve the infrared component of light; the two rows were perpendicular to the conveyor and tilted of approximately 45° toward the camera scanline. The two cameras used were developed at Microtec, and used a CMOS image sensor for the Vis-NIR system and an InGaAs (indium gallium arsenide) sensor in the SWIR system.

The fruits were measured eight times, always having the equatorial side facing up, by manually rotating the fruit of 45° at every measure. No measures were taken at the poles region (calix and pedicel) because in real inline condition fruits rotate while moving forward and tend to align the axis between the poles on a horizontal direction, thus leaving the equatorial region on top, where the acquisition device is located.

The measurement with the spectroscopy methods consisted in a static measure, where the fruit where manually placed in front of the acquisition device. In the case of NIR spectroscopy, the light sources were located above the fruit, while the detector was below the fruit, and a set of collimators was used to collect only the transmitted light and remove any light diffused from the fruit surface or the environment. In the case of SWIR spectroscopy, the light sources and detector were on the same side and contained in a single box. For both setups, we acquired equatorial measurements as well as polar measure, on the calix and pedicel end.

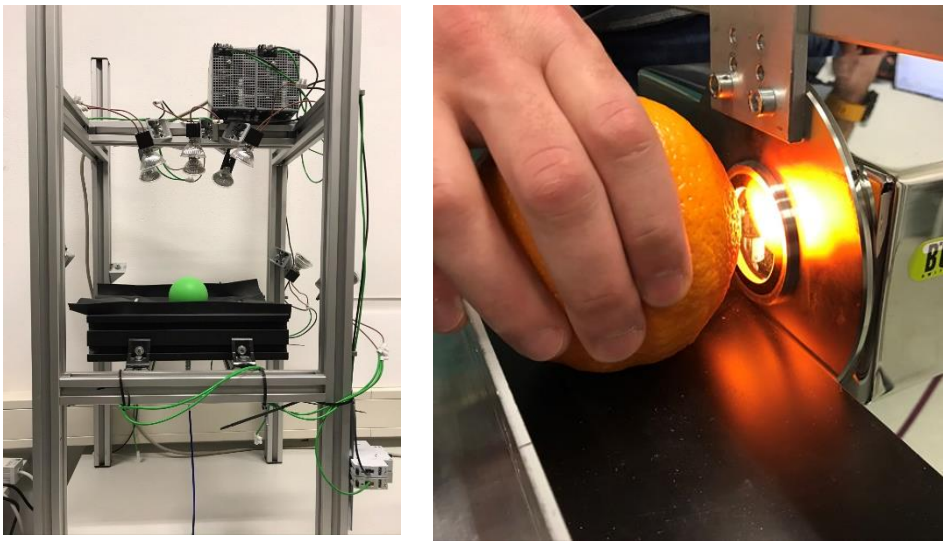


Figure 5.25. Picture of Vis-NIR spectroscopy in transmission (left) and SWIR spectroscopy in reflection (right).

Table 5.6 summarises the conditions of the various methods.

Measure	Method	Condition	Equatorial	Polar
Hyperspectral Vis-NIR	Reflectance	Dynamic	8	0
Hyperspectral SWIR	Reflectance	Dynamic	8	0
Spectroscopy Vis-NIR	Transmittance	Static	8	4
Spectroscopy SWIR	Reflectance	Static	4	2

Table 5.6. Summary of the various non-destructive measures, methods and conditions on oranges.

The destructive analysis was made the day following the non destructive, at the laboratories of the University of Bolzano. Every fruit was cut in half in the equatorial region and squeezed with a citrus juicer. The juice was poured in a beaker and homogenized with a disposable plastic pipette. A small quantity of juice was analyzed with the pocket refractometer pal-3 Atago to assess the SSC. For each juice, two measurements were taken separately and averaged, to obtain a more accurate value. Then titratable acidity was measured on small samples of homogenized juice, following

the same procedure described by Ncama et al. [38], using the automatic titrator Titro Line 5000 SI Analytics. This procedure was repeated for each fruit, keeping all the instruments carefully cleaned with water and deionized water between each analysis.

Total soluble solids content and acidity values measured with the destructive procedures are summarized in Table 5.7.

Parameter	Samples	Repetitions	Min	Max	Avg	Stdev
SSC (°Bx)	120	2	8.05	12.1	10.01	0.86
TA (%c.a.)	120	1	0.267	0.713	0.495	0.088

Table 5.7. Summary of the various destructive analysis on oranges.

5.3.3 Data analysis

The two sets of hyperspectral acquisitions have as output a three dimensional hypercube. Every image was processed with dark signal subtraction and compensation with the white source signal normalization, to obtain a reflectance spectrum. Averaging over the wavelength axis, a two-dimensional image of the scan was obtained, where the fruit position was located and a region of interest was selected around the central part of the fruit. After removing all spectra that were saturated or under-illuminated, the average was computed, in order to obtain a single spectrum for each side.

In the case of NIR spectroscopy the preprocessing was simply a dark signal correction, while no source reference normalization was applied (due to the transmission configuration and the lack of a reference transparent white body).

In the case of SWIR spectroscopy, the used instrument outputs a calibrated reflectance spectrum, after dark correction and source normalization.

After having obtained the set of spectra, other pre-processing techniques were applied:

- Binning
- Standard normal variate (SNV)
- Spectral band reduction (mask)
- Savitky-Golay filter (SG)

By trying the different preprocessing techniques, it was observed that band reduction was always beneficial, while binning and SNV were beneficial only in some cases. The use of Savitky-Golay filtering did not make any particular difference.

These spectra were then correlated with the target variables of SSC and TA using the common partial least squares regression (PLSR) method. PLSR is a very common method, successfully used in spectral data treatment, as shown also by the results obtained in many similar studies [101] [103] [17] [38] [104].

The number of PLS components was selected in a range between 5 and 25 with k-fold cross validation (k=5).

The performance of the model was evaluated considering three parameters, the coefficient of determination for cross validation (R^2cv), root mean square error of cross validation (RMSECV) and residual predictive deviation (RPD).

5.3.4 Results and discussion

The non-destructive prediction of SSC gave satisfactory results. Models built on distal measurements performed worse than those built on equatorial ones (confirming what is reported in [100]).

Models built on average equatorial measurements performed systematically better, but results were coherent with those obtained using the individual measurements for the respective technique. This confirms the assumption that multiple measurements around the whole fruit are more informative than single spot acquisitions. For reflectance analysis, the RMSECV value decreased by 0,1-0,15 °Bx going from single to average equatorial measurements. Considering transmittance analysis, the reduction of RMSECV value was also observable, but with an average extent of 0,05°Bx, confirming a lower relative impact of multiple measurements for transmittance acquisition mode.

The best estimation is given by Vis-NIR transmittance spectroscopy. The models developed for Vis-NIR and SWIR hyperspectral data had lower predictive performances and were very similar. The least performant analysis was SWIR reflectance spectroscopy, however its scores were similar to the hyperspectral models' scores.

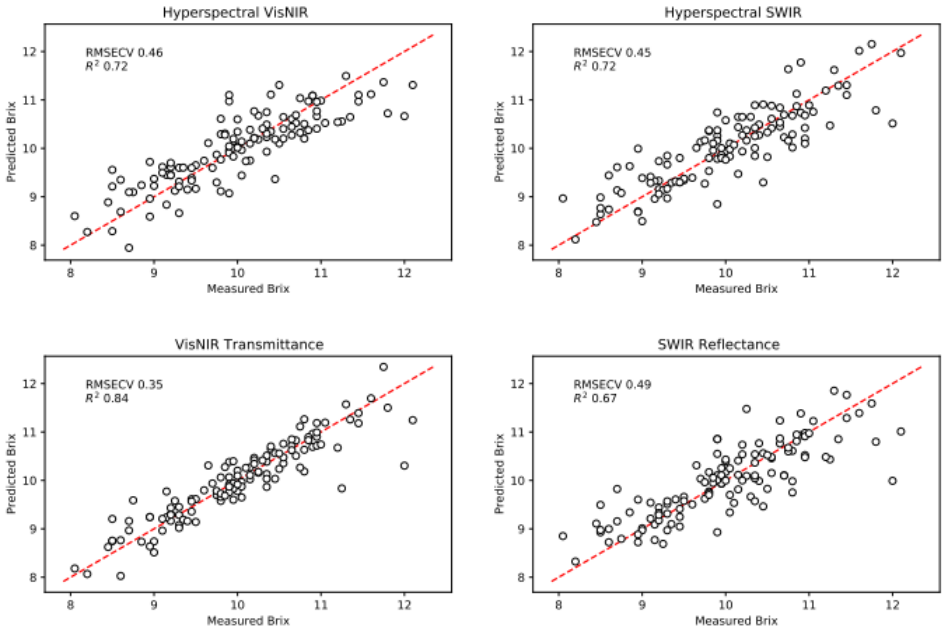


Figure 5.26. Results of the four measures.

The results obtained in this project performs in a similar way as other studies in the literature, in some cases with better results, for example in the case of spectroscopy in transmittance. The following table reports an overview of the comparison.

An interesting note is that the vast majority of the studies applied spectroscopy in reflectance, while only one study was found where the hyperspectral imaging was used.

Method	Configuration	Wavelength (nm)	R^2	RMSECV	Source
Spectroscopy	Transmittance	400-1000	0.84	0.35	This study
Hyperspectral	Reflectance	400-1000	0.72	0.46	This study
Hyperspectral	Reflectance	950-1700	0.72	0.45	This study
Spectroscopy	Reflectance	950-1700	0.67	0.49	This study
Spectroscopy	Reflectance	450-1700	0.87	0.44	[98]
Spectroscopy	Reflectance	350-1800	0.94	0.45	[38]
Spectroscopy	Reflectance	450-2500	0.93	0.28	[17]
Spectroscopy	Reflectance	578-1840	0.91	0.51	[102]
Spectroscopy	Reflectance	350-1050	0.90	0.361	[103]
Spectroscopy	Reflectance	1000-2500	0.93	0.55	[100]
Spectroscopy	Reflectance	350-2500	0.91	0.60	[101]
Spectroscopy	Transmittance	400-980	0.77	0.46	[104]
Hyperspectral	Transmittance	325-1100	0.92	0.38	[105]

Table 5.8. Overview of the applications of hyperspectral imaging and spectroscopy to measure the SSC in citrus and comparison with the literature.

On the other hand, for the total titratable acidity TA, which is the second quality parameter analyzed, no satisfactory prediction could be achieved with any technology applied. This result is in line with what is reported in other studies [102] [101], that confirm the difficulties in the prediction of TA with spectroscopic techniques and hyperspectral imaging, the main reason being probably the relatively low level of organic acids with respect to total SSC in citruses.

The result of this work shows the good opportunity of the spectroscopy and hyperspectral imaging techniques for the assessment of SSC in citrus fruits like oranges. Considering only the predictive performance, the best method came from transmittance Vis-NIR spectroscopy. However, taking in account the potential implementation in sorting lines in an industrial environment, transmittance spectroscopy presents some implementation issues, mainly concerning the need of completely sealing the detector from the light source and from the environment. On the other hand, the methods based on hyperspectral imaging in the Vis-NIR and SWIR wavelength ranges, even with a slightly worse performance, are the most easily transferable to an industrial line. Moreover, the use of a hyperspectral configuration allows -with the arrangement

described in the previous section- to sample the fruit over the complete equatorial area in order to average the measurements on multiple sides.

5.4 Other applications of the hyperspectral technology to wood and fruit

The hyperspectral imaging technology has been tested on many other use cases. Here I shortly report the qualitative results on some of these cases.

5.4.1 False heartwood on log face of Beech wood

The workflow developed for the classification of wood boards was applied to log faces for the classification of false heartwood from clearwood. False heartwood is a defect that largely influences the quality of beech wood and is formed by air penetration into mature wood. In the process of sawing logs into boards, it is important to know in advance the boundary between the false heartwood and clearwood, in order to minimize the amount and volume of products containing it [106] [107].

In this pre-study, five samples were collected from a sawmill in Central Germany. Discs of thickness between 5 and 10 centimeters were cut from logs, having a diameter of about 40 centimeters. One side of the discs was roughly sawed and exhibits deep saw marks, the other had a flatter surface; both sides were scanned. Wood was still in very wet conditions, with large areas containing water in the sapwood region. The exact boundary between these two regions was marked with the help of experts.

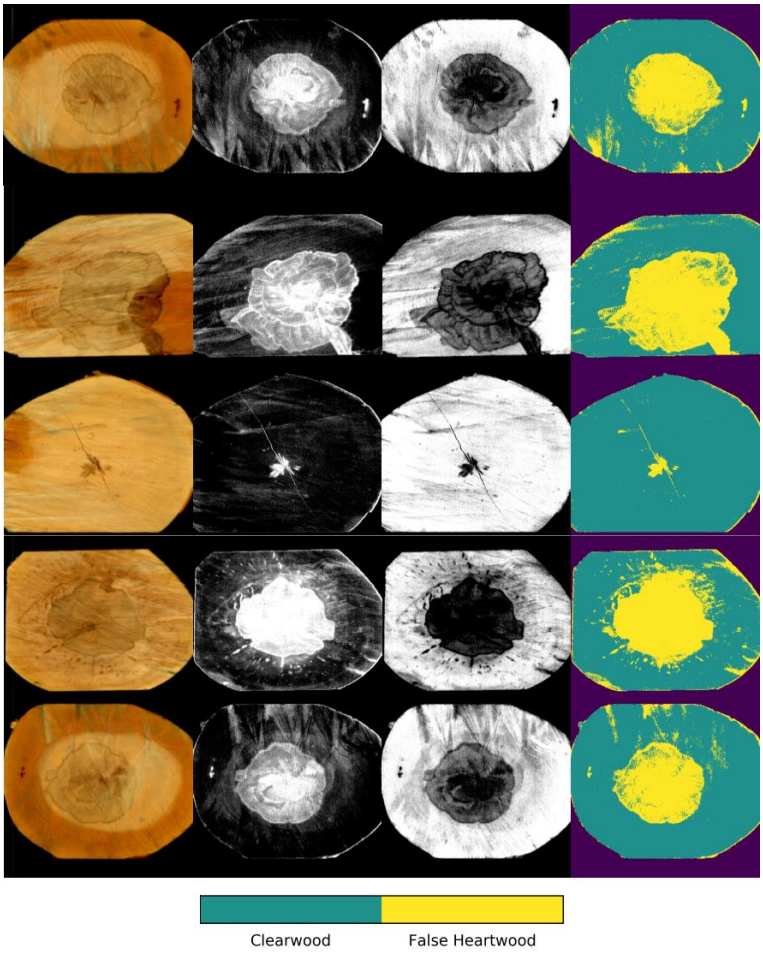


Figure 5.27. Classification of false heartwood on beech wood; from left to right: picture of log face, score map of class heartwood, score map of class clearwood, final classification.

Figure 5.27 left shows the appearance of false heartwood in traditional color RGB images; the darker regions are related to sapwood and to its high content of water; on the other hand, the lighter color in the middle is not all false heartwood, therefore the classification based on pure color images is not trivial. The hyperspectral system permits to differentiate well between the two classes of interest, without being affected much by water.

5.4.2 Moisture content of dates

The distribution chain of the dates requires that these fruits reach a precise level of water content. To obtain this, dates are put in special ovens to dry them; this process is not always accurate, because the initial humidity in dates is not always known, and since the process is relatively rough for speed reasons. A method of detecting to wet fruits (to

be put again in the drying process) or to dry ones (to be rehydrated again) could be useful.

In this study, dates of variety Medjool at different moisture level were collected from two packing houses in the southern deserted part of Israel.

Figure 5.28. Picture of the measured dates.

Figure 5.28 show a subset of these dates from the first packing house with normal, dry and wet samples, and from the second house with only normal and wet samples; note that while in the first case the wet fruits are optically more turgid and it is possible a visual sorting of the wet samples, this is not true in the samples from the second case.

The same workflow used for the classification of wood boards was used. A subset of fruits was marked as belonging to one of the three classes (normal, wet, dry), a model using the PLS method was computed and applied to the remaining of the samples.

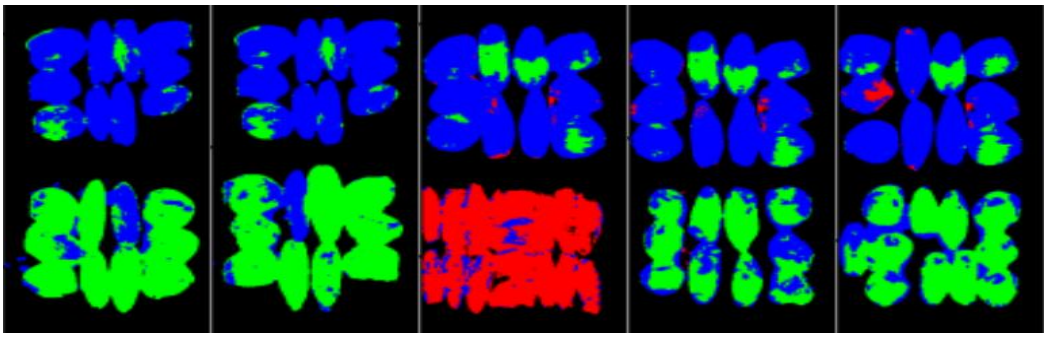


Figure 5.29. Result of classification of dates in three classes according to moisture content: blue=normal, green=wet; red=dry.

Figure 5.29 shows the classification results of the hyperspectral method. Colors is used to mark normal (blue), wet (green) and dry (red) samples. While in some cases the fruits are not classified with the same humidity range, the differences between the groups are clearly visible.

5.4.3 Firmness of blueberries

Firmness is an important quality factor in many fruits and vegetables, and is commonly used to determine the quality and level of fruit maturity, or as indirect measurement of ripeness.

For this test, blueberries were bought from a local store and stored partly in a refrigerator at 4°C and partly at ambient temperature for some days, to produce two different levels of firmness. After a few days, the blueberries reached a very different level of firmness, with the ones stored at ambient temperature that were very soft.

Two groups of fruits were formed, with 28 fruits belonging to the soft class and 26 to the hard class. A measure with the SWIR hyperspectral system using halogen light as illumination was taken, and the complete workflow with the PLS-DA method was applied in order to classify each pixel.

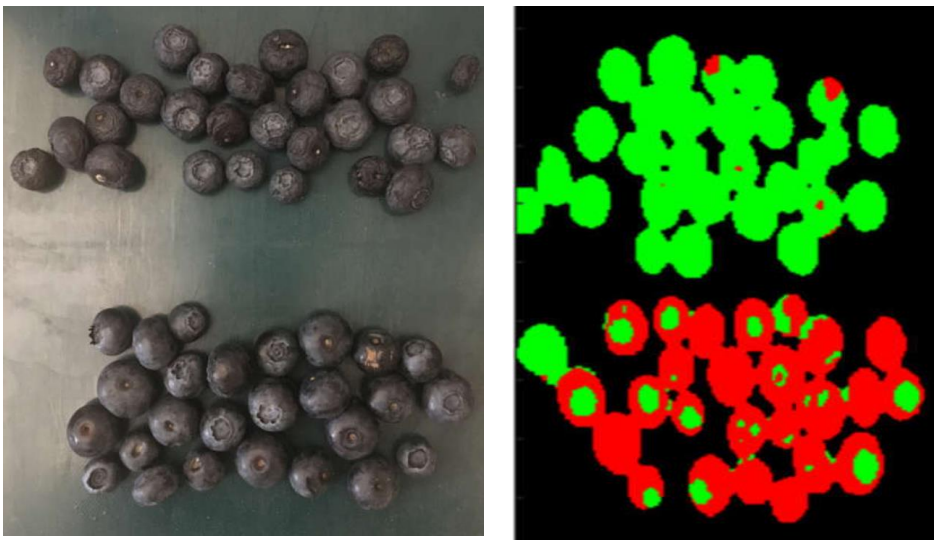


Figure 5.30. Firmness of blueberries; left: picture of the two groups, soft on top and hard on bottom; right: result of the pixel-wise classification.

Figure 5.30 left shows a picture of the blueberries, with the soft ones on to and the hard on the bottom. The result, obtained with LV=10, is reported in Figure 5.30 right, where red denotes hard and green the soft class.

5.5 Application of other methodologies to wood and fruit

During the project, other technologies in addition to hyperspectral imaging were tested. One example was described in the section on the estimation of sugar content and acidy in citrus, where a comparison between spectroscopy and hyperspectral in Vis-NIR and SWIR range were considered. Here I shortly report the qualitative results for two other cases.

5.5.1 Dry matter of avocado with MRI (magnetic resonance imaging)

The importance of dry matter estimation in avocado was described in a previous section. During the study on this application, several methods were tested in addition to hyperspectral imaging:

- Spectroscopy in reflection on the external surface of avocado with and without skin; this included the external layers of pulp as well as internal layer. Spectroscopy in transmission on thin slices was tried as well.
- A thermal camera was used to measure the temporal evolution of temperature over the surface, with the idea that a different quantity of water in the fruit pulp could imply a change in the heat capacity.
- X-ray tomography was tested to check if the different amount of water could be seen as variation in the pulp density. A test was carried out by cutting small cylinders of pulp that were scanned with a high resolution CT scanner and the measured with the dry oven method.
- Dual energy x-ray radiography was tested, in order to measure slices of pulp with known thickness, by exploiting the capacity of dual-energy to resolve the density of composite samples (here we tried to separate the effect of the pulp and water in the x-ray absorption).
- A magnetic resonance imaging scanner was used to scan avocado fruits and to make a volumetric reconstruction.

None of the above-mentioned methods brought to useful results, with the exclusion of the last.

In this test, 43 avocado fruits of variety Haas were considered. Each sample was scanned with a medical device (built from Paramed) working at 0.32T and the volumetric reconstruction was performed. The measure technique was fat and water separation (Dixon), in order to compute two images with intensity proportional to the

density of water and fat (Figure 5.31); since the measure of dry matter considers the percentage of water content in the pulp, and the aim is to measure the oil content, this seamed a good method.

Every scan required approximately 3 minutes. The scan resolution was approximately 1x1 mm with slices every 4.5 mm on the fruit longitudinal axis.

The volumetric reconstructions were averaged, and a simple linear model with two variables (mean intensity of water and oil) was constructed.

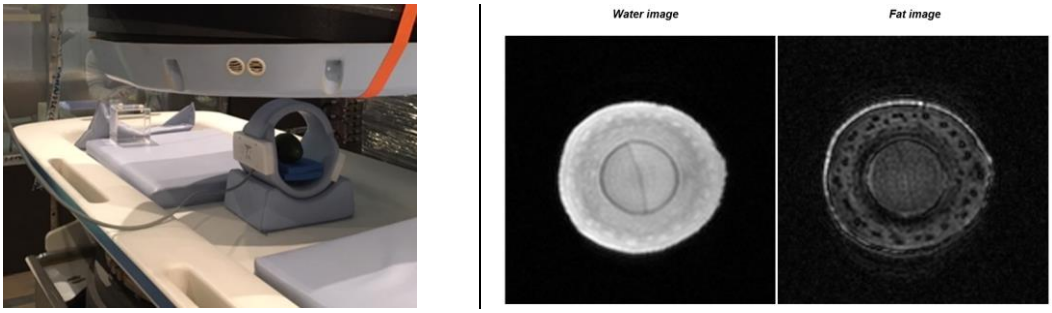


Figure 5.31. Left: the measure setup. Right: one slice from the volumetric reconstruction.

The model showed a good correlation (Figure 5.32 left), notwithstanding a few outliers which presented not representative values. The reasons of these outliers was not clear and would need a better understanding, but after removing them the performances increased (Figure 5.32 right). Further tests were needed, but due to the cost of such an implementation, the complexity to adapt a similar medical device for an industrial use and the time required for a single scan, no further actions were performed.

However initial discussions were started in order to realize a much faster scanning pattern, where no volumetric scanning was performed, but a single measure for the whole fruit volume was considered.

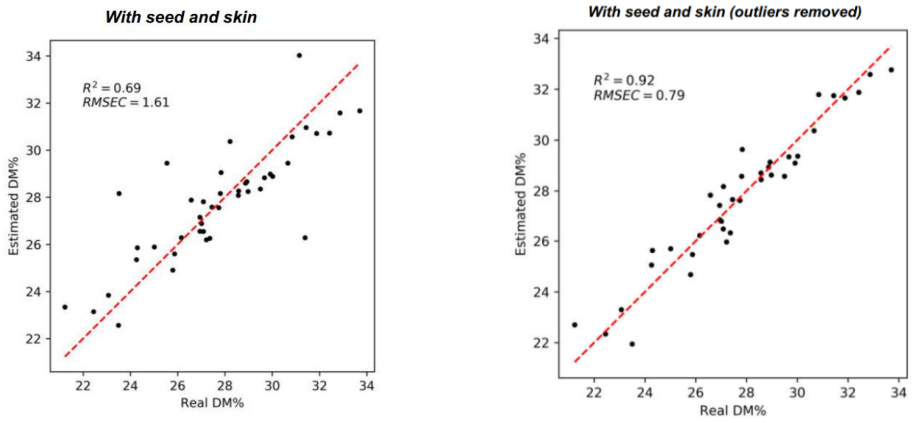


Figure 5.32. Left: linear model results, with all samples. Right: linear model after removal of 5 outliers.

5.5.2 Internal colour of apples with spectroscopy

Modern agriculture is always looking for new varieties than can attract interest of customers and create new niches. One of these examples in the fruit sector is a recent variety of apple, called Red Moon, which has as its main characteristic the red coloring both outside and inside the fruit.

A quality index for this peculiar variety is the percentage of internal colored pulp volume and its color intensity; growers would like to ensure that the product sold maintains the promise of having almost all the internal pulp colored with a dark red intensity. See Figure 5.33 left for an increasing scale of good apples.

We received a set of these apples from Fondazione Edmund Mach (FEM), grown in a local vineyard in Trentino, Italy.

16 apples were chosen, measured with different non-destructive methods, and finally cut in half and sorted into three classes of color. The non-destructive methods were transmittance NIR spectroscopy, hyperspectral Vis-NIR and SWIR in reflectance. Four spectra were acquired with every method, and the single and averaged spectra were correlated with the target class.

While both hyperspectral methods did not achieved good results, the results of NIR spectroscopy are shown in Figure 5.33 right. The spectra were modeled using the PLS method, to compute a continuous color index. A perfect division between the three classes was not possible, but a clear distinction between class 1 (light coloration) and class 3 (intense coloration) is visible.

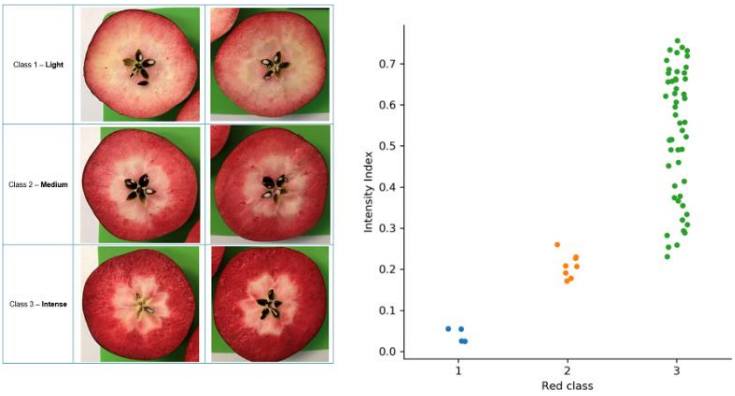


Figure 5.33. Left: examples of different levels of internal red pulp. Right: color index from destructive measure versus color index obtained from the model.

5.6 Conclusions

The following table summarizes the experiments explained in the previous sections.

Material	Application	Method	Configuration	Wavelength range
Wood boards	Heartwood/sapwood	Hyperspectral	Reflection	Vis-NIR
Avocado	Dry matter	Hyperspectral	Reflection	SWIR
Citrus	Brix° (sugar)	Hyperspectral	Reflection	Vis-NIR or SWIR
Wood logs	False heartwood	Hyperspectral	Reflection	Vis-NIR
Dates	Moisture content	Hyperspectral	Reflection	SWIR
Blueberries	Firmness	Hyperspectral	Reflection	SWIR
Avocado	Dry matter	Magnetic resonance	-	-
Apple	Pulp color	Spectroscopy	Transmission	NIR

Table 5.9. Summary of all reported tests.

This table shows that the hyperspectral imaging technology has been proven to be effective for solving a wide range of tasks in the automatic sorting and grading of wooden and organic products.

This technology was tested on wood boards to segment the heartwood from the sapwood, and on wood log slices to segment the false heartwood from the sapwood, tasks that with other methods based on other technologies, like color or x-ray sensors, could not be tackled. In the quality sorting of wooden products, there are many other parameters that were only quickly explored, like the presence of fungi, brown or soft rot or the juvenile wood, where this technology could be further applied.

Moreover, the same technology was successfully used to estimate the dry matter content in avocado, soluble solid content in oranges, moisture in dates and firmness of blueberries. These same parameters (dry matter, sugar contents, moisture, firmness) are relevant for other fruits as well, therefore it is likely possible that the applicability can be further increased.

Although other methods were more effective in solving other tasks, like spectroscopy in transmission for the determination of the sugar content or magnetic resonance for the dry matter estimation, they pose more issues from a practical point of view in terms of integration in industrial environments or extremely high cost, which might limit their diffusion.

On the other hand, the hyperspectral imaging technology is more friendly in terms of the potential implementation on conveyor lines, a common configuration in industrial application: it is a contactless and non-destructive measure, it can cope with the typical speed of the industry, it does not require special mechanical arrangements since it uses a reflection configuration.

6 Output

This chapter, in addition to list the published papers and patents during the PhD work, shows additional projects and activities that I started and are linked to the thesis. These projects are mainly cooperation outside the company Microtec, involving the university or other research centers.

6.1.1 Patents and papers

Marco Boschetti, Giancarlo Zane, Simone Faccini, "Metodo e apparecchiatura per eseguire un esame iperspettrale della superficie esterna di prodotti vegetali", Italian requested Patent IT102020000022354, 2020

Marco Boschetti, Simone Faccini, Enrico Ursella and Giancarlo Zane, “Method and apparatus for non-destructive inspection of fruits having an axis of rotational symmetry”. European Patent EP3521812B1, 15 07 2020.

Enrico Ursella, Marco Boschetti, “Tunnel CT Scanner”. European patent request EP3690429A1, 05 08 2020.

Marco Boschetti, Simone Faccini, Matteo Caffini, "Imaging iperspettrale su tavole di legno", Italian patent request 2021

Phyu Phyu Htun, Marco Boschetti, Attaullah Buriro, Roberto Confalonieri,
Boyuan Sun, Ah Nge Htwe, and Tammam Tillo, “A lightweight approach for wood hyperspectral images classification”, Conference on Intelligent Computation in Manufacturing Engineering, Gulf of Naples, Italy, 2021

Nicola Busatto, Lorenzo Vittani, Brian Farneti, Marco Boschetti, Fabrizio Costa, Simone Faccini, Matteo Caffini, "Physiological and molecular characterization of the late ripening stages in Mangifera indica cv Keitt", Postharvest Biology and Technology, 2021 (submission)

E. Ursella, F. Giudiceandrea and M. Boschetti, “A Fast and Continuous CT scanner for the optimization of logs in a sawmill,” in 8th Conference on Industrial Computed Tomography, Wels, Austria, 2018.

A. Gasparetto, D. Ressi, F. Bergamasco, M. Pistellato, L. Cosmo, M. Boschetti, E. Ursella and A. Albarelli, “Cross-Dataset Data Augmentation for Convolutional Neural Networks Training . Int. Conf. on Pattern Recognition,” in Int. Conf. on Pattern Recognition, (ICPR2018), 2018.

6.1.2 “H2I”, FESR project

The goal of this project is the development of an hyperspectral SWIR camera and illumination, of methods based of deep learning to analyze the hypercubes, and the creation of knowhow and dataset specific of the local region. This project is a cooperation between the University of Bolzano and the company Microtec. I made the initial concept of this project, that was created together with prof. Tillo Tammam from the University and my colleague Andrea Ciresa from Microtec.

The project has been accepted and granted with 352.310€.

A. Proposta progettuale

1. Informazioni generali

Acronimo		H2I	
Titolo del progetto (IT)		Immagini iperspettrali per l'ispezione del legno e della frutta	
Titolo del progetto (DE)		Hyperspektrale Bilder für Holz- und Obst Inspektionen	
Durata del progetto	**Inizio**	01/01/2019	**Numero mesi**
	Fine	30/06/2021	29
Asse prioritario		1 - Ricerca e Innovazione	
Obiettivo tematico		1 - Rafforzare la ricerca, lo sviluppo tecnologico e l'innovazione	
Priorità investimento del PO		1.b - Rafforzare la ricerca, lo sviluppo tecnologico e l'innovazione provvedendo a promuovere gli investimenti delle imprese in R&I	
Risultato atteso		1.1 - Incremento dell'attività di innovazione delle imprese	
Bando di riferimento		Ricerca e Innovazione - 4° Bando	
Azione		1.1.4 - Sostegno alle attività collaborative di R&S per lo sviluppo di nuove tecnologie sostenibili, di nuovi prodotti e servizi	
Area di specializzazione S3		ICT e automation	
Codice del progetto		FESR1111	
Beneficiari > 1		Si	
Creato		17/07/2018 - Tammam Tillo	

Figure 6.1. H2I project summary.

Abstract:

Hyperspectral imaging permits to study the composition of objects in a contactless manner, by analyzing the reflected spectrum of every single pixel.

Hyperspectral imaging is already widely used in remote sensing applications in astronomy, but is gaining attention for archeology, agriculture and food industries as well.

There are two main challenges with this technology: first, the required hardware is very expensive and complex; second, the extraction of useful information from the hyperspectral images is complicated, because in addition to the spatial two dimensions, there exists the spectral dimension.

In this project we will design a hyperspectral detection platform, based on a configurable InGaAs sensor and a dedicated light source with a tunable spectral profile.

In addition, we will develop a plurality of deep learning based methods to extract the spatial and spectral information at the same time and develop training procedures which will be particularly effective for hyperspectral data.

The proposed platform could be used in a multitude of use cases, ranging from the inspection of defects on wooden materials, to the estimation of the degree of maturity and ripeness or the identification of any internal defects of fruits and vegetables, or to check the conditions of archaeological artifacts.

6.1.3 Master thesis of Ilario Chini with University of Bolzano.

The goal of this master thesis with title Use of non-destructive technologies to evaluate the quality parameters soluble solid content and acidity in oranges", was to make a comparison between different techniques to estimate internal properties of citrus; the examined techniques were spectroscopy and hyperspectral imaging both in NIR and SWIR range.

Abstract:

Currently, about 50 million tons of oranges, distinguished in non-blood oranges, blood oranges, navel oranges and bitter oranges are harvested every year all over the world. Orange processing industry requires reliable and cost-effective systems for individual fruit grading, compatible with in line implementation. This research work studied the performances of Vis/NIR spectroscopy, SWIR spectroscopy, Vis/NIR hyperspectral imaging, SWIR hyperspectral imaging and a simulated Vis/NIR-SWIR hyperspectral imaging to assess SSC and TA in Lane Late oranges, reasoning also about their possible implementation on devices for industrial in-line quality assessment. This study highlighted the effectiveness of all studied non-destructive technologies for the assessment of SSC with variable scores, while no satisfactory prediction of TA was achieved. Vis/NIR transmittance spectroscopy delivered the best prediction capability

for SSC parameter, with achieved model scores of $R^2cv = 0,84$; RMSECV = 0,35°Bx and RPD = 2,46. However, despite prediction performances, this technology was assessed incompatible for in-line implementation due to technical requirements. Better perspectives are seen for SWIR and HSI industrial implementation, even if these technologies would require technical improvements to enhance their prediction capability for SSC.

6.1.4 "Fruit 2020" project

The goal of this project was to study the measure of quality of fruits products and to enhance the performance of Microtec scanner in measuring both internal and external quality. Special effort was put to analyze novel sensing techniques (NIR spectrometry in transmission, microwaves at 3-5 GHz, magnetic resonance, imaging in the SWIR range, dual energy x-ray), making pilot tests on samples and collecting know how on the sensors aspects. An activity was related to enhance the computation by using edge devices or using small clusters of computers. I made the initial concept of this project, and prepared it with my colleague Andrea Ciresa.

This project has been granted by the Province of Bolzano with 133.000€.

Abstract:

The modern agri-food market is focusing more and more on high quality products, aesthetically perfect and with a defined and pleasant taste, in order to meet the needs of customers.

The fruit industry, therefore, in the world as well as in the province of Bolzano, is investing in fruit quality measurement systems that take into account two important aspects:

- The measure of the perceived goodness of the fruit
- Improvement of measures (especially on types of defects that are particularly difficult to detect with current technologies)

The measurement of perceived goodness is closely linked to the way in which the fruit is handled before reaching the consumer: in fact, most of the fruit is harvested from the unripe tree and is subsequently artificially ripened (either during transport or in warehouses, usually located near the places of consumption).

Fruits that are able to autonomously continue the ripening process even after harvest (therefore after being physically detached from the tree) are called "climacteric".

When these fruits are taken into consideration, we try to analyze two particularly important processes:

1) The maturity, which is the development process that allows the fruit to achieve minimum quality characteristics after harvesting,

2) The ripening which is the process of reaching the edibility stage.

For producers of technology serving the agri-food industry such as Microtec, it is therefore necessary to know these physical/chemical processes in detail and therefore to devise systems capable of measuring, for example quantitatively, the state of ripening.

This would make it possible to schedule the "time-to-consumer" of the individual fruit in an innovative way, correctly managing the thermal and controlled atmosphere characteristics to reach at a specific time (precisely the one in which the fruit will be displayed on store shelves) the optimum edibility stage. To be able to do this, one of the important indicators to take into consideration is the so-called dry matter, i.e. the weight of the material that remains after the drying of the pulp (which, as it is easy to imagine, in many fruits is closely linked at maturation).

In addition to maturity, ripening and dry matter (which are currently difficult to estimate or evaluate with current technologies or by not carrying out destructive tests of the fruit that would compromise its marketing) it is necessary to be able to measure indicators related to the taste-olfactory perception of the consumer.

This approach seeks to shift the determination of quality from mere aesthetic indicators (uniform color, regular shape and absence of stains) to indicators more related to taste, such as sugar content, softness of the pulp, quantity of juice, acidity. The determination of a single parameter of goodness is highly subjective and territorial (some territorial areas prefer sweet and juicy apples, other apples more acidic and crunchy), but the single objective parameters can be measured and used to differentiate the product offered to the consumer.

Being able to achieve, through an in-line machine, a certain result compatible with the timing and needs of the industry would already lead to a great goal since there are no machines on the market that are able to extrapolate these indicators.

6.1.5 Determination of maturity of mango fruit project

The goal of this activity was the quantitative determination of the maturity level of mango fruits and the study of the correlation with non-destructive techniques. This is cooperation with Fabrizio Costa and Nicola Busatto from the Edmund Mach Foundation FEM). This activity was unfortunately blocked due the Covid pandemic.

7 Conclusions

Industrial processes always require more the exact and precise knowledge of the material being processed, in order to refine, automate and optimize the production flow, to maximize the final product's value and quality, and to maximize the resource utilization in a sustainable way. These systems are based commonly on vision cameras and light sources, but there is a trend to enhance the capabilities by adding and combining a diversity of sensors, working in the infrared, UV or x-ray wavelength ranges.

Hyperspectral imaging systems are a recent technology which has proven to be useful to improve the detection capability of industrial scanners for automatic sorting according to quality parameters. So far the introduction of such systems is still limited due to technical and economic reasons. In this thesis I have outlined the work that I have supervised and directly conducted at Microtec to design a suitable system for the application in industrial environment, from the hardware and software point of view.

In addition I have shown many applications that have been tested, validated and in some case implemented in real use scenarios. These use cases were in the industry of wood and food processing, the field where Microtec is active. The use cases considered wooden boards and logs of different species (eucalyptus, oak, beech), and fruits of different varieties (avocados, citrus, apples, blueberries); I considered also different type of parameters, both external and internal.

The research I carried on has already brought concrete results that have led to applications in real industrial installations. In addition it could be regarded as a starting point for several future developments. First we need to consolidate the results obtained so far by monitoring the installations and possibly increasing the number of in-field implementations for the cases that were tested in laboratory only. Then there are improvements on the hardware side as well: technology changes at a fast pace, and new sensor are now available that are sensitive to the full visible, near infrared and SWIR range (400nm to 1700nm) that could permit to fuse the two developed systems into a single one; work is already ongoing to realize a longitudinal compact multi-chip LED lamp to overcome the limitations of wavelength overlap. Finally, the hyperspectral technology could be tested on additional challenging tasks in laboratory setups, in order to find other useful applications and increase its additional value in the future.

8 Bibliography

[1] Y. Lu, W. Saeys, M. Kim, Y. Peng and R. Lu, “Hyperspectral imaging technology for quality and safety evaluation of horticultural products: A review and celebration of the past 20-year progress,” *Postharvest Biology and Technology,* vol. 170, p. 111318, 2020.

[2] J. C. George Joseph, “Fundamentals of Remote Sensing,” 2017.

[3] I. Renhorn, D. Bergström, J. Hedborg, D. Letalick and S. Möller, “High spatial resolution hyperspectral camera based on a linear variable filter,” *Optical Engineering,* 2016.

[4] T. Takatani, T. Aoto and Y. Mukaigaw, “One-shot Hyperspectral Imaging using Faced Reflectors,” *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2017.

[5] W. R. Johnson, D. W. Wilson, W. Fink, M. D. Mark S. Humayun and G. H. Bearman, “Snapshot hyperspectral imaging in ophthalmology,” *J. of Biomedical Optics,* 2007.

[6] A. Bodkin, A. Sheinis, A. Norton, J. Daly, S. Beaven and J. Weinheimer, “Snapshot hyperspectral imaging: The hyperpixel array camera,” *Proceedings of SPIE,* 2009.

[7] J. Qin, M. S. Kim, K. Chao, D. E. Chan, S. R. Delwiche and B.-K. Cho, “Line-Scan Hyperspectral Imaging Techniques for Food Safety and Quality Applications,” *Applied Sciences,* 2017.

[8] K. Iniewski, “CZT detector technology for medical imaging,” 2014.

[9] K. Héberger, “Chemoinformatics, multivariate mathematical–statistical methods for data evaluation,” 2008.

[10] D. L. Massart, B. G. M. Vandeginste, S. N. Deming, Y. Michotte and L. Kaufman, Chemometrics: a textbook, Elsevier, 2003.

[11] D. Cozzolino, W. U. Cynkar, N. Shah and P. Smith, “Multivariate data analysis applied to spectroscopy: Potential application to juice and fruit quality,” *Food Research International,* 2011.

[12] W. Saeys, N. N. Do Trong, R. Van Beers and B. M. Nicolai, “Multivariate calibration of spectroscopic sensors for postharvest quality evaluation: a review.,” *Postharvest Biol. Technol.,* 2019.

[13] Y. Lu, W. Saeys, M. Kim, Y. Peng and R. Lu, “Hyperspectral imaging technology for quality and safety evaluation of horticultural products: A review and celebration of the past 20-year progress,” *Postharvest Biology and Technology,* vol. 170, p. 111318, 2020.

[14] S. Wold, M. Sjöström and L. Eriksson, “PLS-regression: a basic tool of chemometrics,” *Chemometrics and Intelligent Laboratory Systems,* 2001.

[15] S. Munera, C. Besada, J. Blasco, S. Cubero, A. Salvador, P. Talens and N. Aleixos, “Astrincency assesment of persimmon by hyperspectral imaging,” *Postharvest Biology and Technology,* 2017.

[16] B. Nicolaï, K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K. Theron and J. Lammertyn, “Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review,” *Postharvest Biology and Technology,* 2007.

[17] L. S. Magwaza, U. L. Opara, H. Nieuwoudt, P. J. R. Cronje, W. Saeys and B. Nicolaï, “NIR Spectroscopy Applications for Internal and External Quality Analysis of Citrus Fruit—A Review,” *Food and Bioprocess Technology,* 2012.

[18] A. Lopez, S. Arazuri, I. García, J. Mangado and C. Jare, “A Review of the Application of Near-Infrared Spectroscopy for the Analysis of Potatoes,” *Journal of Agricultural and Food Chemistry,* 2013.

[19] A. M. A. H. H. E. Schmilovitch and Y. Fuchs, “Determination of mango physiological indices by near-infrared spectrometry,” *Postharvest Biology and Technology,* 2000.

[20] P. Theanjumpol, G. Self, R. Rittiron, T. Pankasemsu and V. Sardsud, “Selecting Variables for Near Infrared Spectroscopy (NIRS) Evaluation of Mango Fruit Quality,” *The Journal of Agricultural Science,* 2013.

[21] M. Nagle, B. Mahayothee, P. Rungpichayapichet, S. Janjai and J. Müller, “Effect of irrigation on near-infrared (NIR) based prediction of mango maturity,” *Scientia Horticulturae,* 2010.

[22] S. N. Jha, P. Jaiswal, K. Narsaiah, R. Kumar, R. Sharma, P. Kaur and R. Bhardwaj, “Determination of Maturity Parameters of Indian Mango Cultivars Based on Eating Quality during Natural Ripening,” *Journal of Agricultural Engineering,* 2013.

[23] C. O. V.Cortés, J. B. N.Aleixos and P. T. S.Cubero, “A new internal quality index for mango and its prediction by external visible and near-infrared reflection spectroscopy,” *Postharvest Biology and Technology,* 2016.

[24] T. J. Nagy Attila, “Sweet cherry fruit analysis with reflectance measurements,” *Analele Universităţii din Oradea,* 2011.

[25] D.-W. S. Yao-Ze Feng, “Application of Hyperspectral Imaging in Food Safety Inspection and Control: A Review,” *Critical Reviews in Food Science and Nutrition,* 2012.

[26] B. Zhang, W. Huang, J. Li, C. Zhao, S. Fan, J. Wu and C. Liu, “Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review,” *Food Research International,* 2014.

[27] Y.-Y. Pu, Y.-Z. Feng and D.-W. Sun, “Recent Progress of Hyperspectral Imaging on Quality and Safety Inspection of Fruits and Vegetables: A Review,” 2015.

[28] P. Martinsen and P. Schaare, “A near-infrared imaging spectrometer,” *Symp. Proceedings of Image and Vision Computing New Zealand,* 1996.

[29] P. Menesatti, A. Zanella, S. D’Andrea, C. Costa, G. Paglia and F. Pallottino, “Supervised Multivariate Analysis of Hyper-spectral NIR Images to Evaluate the Starch Index of Apples,” *Food Bioprocess Technol,* 2009.

[30] F. Mendoza, R. Lu, D. Ariana, H. Cen and B. Bailey, “Integrated spectral and image analysis of hyperspectral scattering data for prediction of apple fruit firmness and soluble solids content,” *Postharvest Biology and Technology,* 2011.

[31] R. J. M. H.-L. A. D.-I. B. &. B. P. Lleó, “Comparison of multispectral indexes extracted from hyperspectral images for the assessment of fruit ripening,” 2011.

[32] W. N. E. G. R. G. S. V. &. G. Y. Rajkumar, “Studies on banana fruit quality and maturity stages using hyperspectral imaging,” *J. Food Eng.,* 2012.

[33] Z. Schmilovitch, T. Ignat, V. Alchanatis, J. Gatker, V. Ostrovsky and J. Felföldi, “Hyperspectral imaging of intact bell peppers,” *Biosystems Engineering,* 2014.

[34] X. Zhou, Z. WenYou, Y. LiGuo, Z. Chao and W. Qin, “Study on the clarification technology and stability of kiwi dry wine through whole fruit fermentation,” *Food Research and Development,* 2017.

[35] L. Li, W. Zhao, X. Feng, L. Chen, L. Zhang and L. Zhao, “Changes in Fruit Firmness, Cell Wall Composition, and Transcriptional Profile in the yellow fruit tomato 1 (yft1) Mutant,” *Agric. Food Chem.,* 2018.

[36] C. J. Clark, V. A. McGlone, C. Requejo, A. White and A. B. Woolf, “Dry matter determination in ‘Hass’ avocado by NIR spectroscopy,” *Postharvest Biology and Technology,* 2003.

[37] R. Cronje, W. Saeys and B. Nicolai, “NIR Spectroscopy Applications for Internal and External Quality Analysis of Citrus Fruit—A Review,” *Food Bioprocess Technology,* 2012.

[38] K. Ncama, U. L. Opara, S. Z. Tesfay, O. A. Fawole and L. S. Magwaza, “Application of Vis/NIR spectroscopy for predicting sweetness and flavour parameters of ‘Valenciaorange (Citrus sinensis) and ‘Star Ruby’ grapefruit (Citrus paradisi Macfad),” *Journal of Food Engineering,* 2017.

[39] T. Ma, T. Inagaki and S. Tsuchikawa, “Calibration of SilviScan data of Cryptomeria japonica wood concerning density and microfibril angles with NIR hyperspectral imaging with high spatial resolution,” *Holzforschung,* 2017.

[40] S. S. Kelley, T. G. Rials, R. Snell, L. H. Groom and A. Sluiter, “Use of near infrared spectroscopy to measure the chemical and mechanical properties of solid wood,” *Wood Science and Technology,* 2004.

[41] A. Thumm, M. Riddell, B. Nanayakkara, J. Harrington and R. Meder, “Near infrared hyperspectral imaging applied to mapping chemical composition in wood samples,” *Journal of Near Infrared Spectroscopy,* 2010.

[42] A. Thumm, M. Riddell, B. Nanayakkara, J. Harrington and R. Meder, “Mapping within-stem variation of chemical composition by near infrared hyperspectral imaging,” *Journal of Near Infrared Spectroscopy,* 2016.

[43] A. Thumm and M. Riddell, “Resin defect detection in appearance lumber using 2D NIR spectroscopy,” *European Journal of Wood and Wood Products,* 2017.

[44] R. Meder and R. R. Meglenb, “Near infrared spectroscopic and hyperspectral imaging of compression wood in Pinus radiata D. Don,” *Journal of Near Infrared Spectroscopy,* 2012.

[45] A. Haddadi, J. Burger, B. Leblon, Z. Pirouz, K. Groves and J. Nader, “Using near-infrared hyperspectral images on subalpine fir board. Part 1: Moisture content estimation,” *Wood Material Science and Engineering,* 2015.

[46] T. Ma, T. Inagaki, M. Ban and S. Tsuchikawa, “Rapid identification of wood species by near-infrared spatially resolved spectroscopy (NIR-SRS) based on hyperspectral imaging (HSI),” *Holzforschung,* 2019.

[47] H. Kanayama, T. Ma, S. Tsuchikawa and T. Inagaki, “Cognitive spectroscopy for wood species identification: Near infrared hyperspectral imaging combined with convolutional neural networks,” *Analyst,* 2019.

[48] P. Mishra, S. Lohumi, H. Ahmad Khan and A. Nordon, “Close-range hyperspectral imaging of whole plants for digital phenotyping: Recent applications and illumination correction approaches,” *Computers and Electronics in Agriculture,* 2020.

[49] H. Kobori, N. Gorretta, G. Rabatel, V. Bellon-Maurel, G. Chaix, J.-M. Roger and S. Tsuchikawa, “Applicability of Vis-NIR hyperspectral imaging for monitoring wood moisture content,” *Holzforschung,* 2013.

[50] C. J. G. Colares, T. C. M. Pastore, V. T. R. Coradin, L. F. Marques, A. C. O. Moreira, G. L. Alexandrino, R. J. Poppi and J. W. B. Braga, “Near infrared hyperspectral imaging and MCR-ALS applied for mapping chemical composition of the wood specie Swietenia Macrophylla King (Mahogany) at microscopic level,” *Microchemical Journal,* 2016.

[51] W. E. Hillis, Heartwood and tree exudates, Springer Verlag, 1987.

[52] “The influence of irrigation and fertilization on heartwood and sapwood contents in 18-year-old Eucalyptus globulus trees,” *Canadian Journal of Forest Research,* 2006.

[53] B. Lachenbruch, J. R. Moore and R. Evans, “Radial Variation in Wood Structure and Function in Woody Plants, and Hypotheses for Its Occurrence,” *Tree Physiology,* 2011.

[54] C.-L. Chen, “Lignins : Occurrence in wood tissues, isolation, reactions and structures,” *Wood Structures and Composition,* 1991.

[55] T. F. Shupe, C. Y. Hse, E. T. Choong and L. H. Groom, “Differences in some chemical properties of innerwood and outerwood from five silviculturally different loblolly pine stands,” *Wood and Fiber Science,* 1997.

[56] R. H. G. Burger ML, Anatomy of wood, Nobel, 1991.

[57] Z. Kai, Y. Haihong and G. Li, “Theoritical analysis of wood comprehensive utilization system optimization of jilin province,” *Journal of Jilin Forestry University,* 1991.

[58] R. Singleton, D. S. DeBell and B. L. Gartner, “Effect of extraction on wood density of western hemlock (Tsuga heterophylla (Raf.) Sarg.),” *Wood and Fiber Science,* 2003.

[59] M. Grabner, U. Müller, N. Gierlinger and R. Wimmer, “Effects of heartwood extractives on mechanical properties of larch,” *Iawa Journal,* 2005.

[60] T. D. PKC Pillai and K. V. Sankaran, “Wood density and heartwood proportion in eucalyptus trees from intensively-managed short-rotation plantations in Kerala, India,” *Journal of Tropical Forest Science,* 2013.

[61] A. J. Panshin and C. de Zeeuw, Textbook of Wood Technology, McGraw-Hill Book, 1980.

[62] R. K. Bamber, “Sapwood and heartwood,” *Technical Papers,* 1981.

[63] A. M. F. Oliveira, A. T. d. Lelis, E. S. Lepage, G. A. C. Lopez, L. C. S. d. Oliveira, M. D. Cañedo and S. Milano, “Agentes destruidores da madeira,” *Manual de preservação de madeiras,* 1986.

[64] A. C. Wiedenhoeft, Structure and Function of Wood, 2012.

[65] J. Wilkes, “Heartwood development and its relationship to growth in Pinus radiata,” *Wood Science and Technology,* 1991.

[66] R. J. McKimm, “Characteristics of the wood of young fast-grown trees of Eucalyptus nitens Maiden with special reference to provenance variation. I. Variations in growth, strain and density associated with provenance,” *Australian Forest Research,* 1985.

[67] A. G. Campbell, W.-J. Kim and P. Koch, “Chemical Variation In Lodgepole Pine With Sapwood/Heartwood, Stem Height, and Variety,” *Wood and Fiber Science,* 1990.

[68] J. E. Winandy and J. J. Morrell, “Relationship between incipient decay, strength, and chemical composition of Douglas-fir heartwood,” *Wood and Fiber Science,* 1993.

[69] C. Clarke, D. C. F. Garbutt and J. Pearce, “Growth and wood properties of provenances and trees of nine Eucalypt species,” *Appita Journal,* 1997.

[70] K. F. R. Bamber, “Sapwood and heartwood: a review,” *Chemistry,* 1985.

[71] W. B. Storey, “What Kind of Fruit is the Avocado?,” *California Avocado Society Yearbook,* 1973.

[72] C. W. B.B. Wedding, R. D. W. S.Grauf and P. .. G. B.Tilse, “Effects of seasonal variability on FT-NIR prediction of dry matter content for whole Hass avocado fruit,” *Postharvest Biology and Technology,* 2013.

[73] U. F. Amos Mizrach, “Nondestructive ultrasonic determination of avocado softening process,” *Journal of Food Engineering,* 1999.

[74] F. R. Harker, S. R. Jaeger, P. Hofman and C. Bava, “Australian consumers' perceptions and preferences for Hass avocado,” *Horticulture Australia,* 2007.

[75] J. A. Abbott, “Quality measurement of fruits and vegetables,” *Postharvest biology and technology,* 1999.

[76] P. Butz, C. Hofmann and B. Tauscher, “Recent developments in noninvasive techniques for fresh fruit and vegetable internal quality analysis,” *Journal of food science,* 2005.

[77] C. Scotter, “Use of near infrared spectroscopy in the food industry with particular reference to its applications to on/in-line food processes,” *Food Control,* 1990.

[78] G. S. Birth, G. G. Dull, W. T. Renfroe and S. J. Kays, “Nondestructive spectrophotometric determination of dry matter in onions,” *Journal of the American Society for Horticultural Science,* 1985.

[79] R. Hartmann and H. Bijning-Pfaue, “NIR determination of potato constituents,” *Potato Research,* 1998.

[80] S. K. VA McGlone, “Firmness, dry-matter and soluble-solids assessment of postharvest kiwifruit by NIR spectroscopy,” *Postharvest Biology and Technology,* 1998.

[81] J. Z. &. Y. L. Xiaobo, “Selection of the efficient wavelength regions in FT-NIR spectroscopy for determination of SSC of ‘Fuji’ apple based on BiPLS and FiPLS models.,” *Vibrational Spectroscopy,* 2006.

[82] K.B. Walsh and C. V. Greensill, “Sorting of Fruit Using near Infrared Spectroscopy: Application to a Range of Fruit and Vegetables for Soluble Solids and Dry Matter Content,” *Journal of Near Infrared Spectroscopy,* 2004.

[83] A. Schroeder C., “Physiological Gradient in Avocado Fruit,” *California Avocado Society,* 1985.

[84] D. Girod, J. A. Landry, G. Doyon, J. A. Osuna-Garcia, S. Salazar-Garcia and R. J. & Goenaga-Portela, “Evaluating hass avocado maturity using hyperspectral imaging,” *Caribbean Food and Crops Society Proceeding,* 2008.

[85] V. D. J. Jairo, S. A. A. Piedad and R. Z. D. Viviana, “Prediction of dry matter content of recently harvested Hass avocado fruits using hyperspectral imaging,” *Journal of the Science of Food and Agricolture,* 2020.

[86] O. Omoniyi, I. Bertling and L. Magwaza, “Non-destructive evaluation of avocado fruit maturity using near infrared spectroscopy and PLS regression models,” *Scientia Horticulturae,* 2016.

[87] M. Boschetti, G. Zane and S. Faccini, “Metodo e apparecchiatura per eseguire un esame iperspettrale della superficie esterna di prodotti vegetali,” *Italian requested Patent IT102020000022354,* 2020.

[88] S. Ladaniya M, “Citrus Fruit: Biology, Technology and Evaluation,” *Elsevier Academic Press,* 2008.

[89] C. B. Watkins, “Postharvest ripening regulation and innovation in storage technology,” *Acta Horticulturae,* 2008.

[90] K. Y. Filatova I.A., "The significance of flavonoids from citrus juice in disease prevention," *Pishcheviaya Promyshlennost,* 1999.

[91] M. Shahbandeh, *World production of citrus fruits in 2018, by region,* 2020.

[92] A. Bonaccorsi, *Settore agrumicolo in Italia: i numeri del comparto, tra bio e innovazione,* 2020.

[93] M. Alleri, "Applicazione di tecniche non distruttive per la valutazione delle proprietà fisico-meccaniche dei frutti," *PhD Thesis,* 2017.

[94] U. L. Magwaza L S, "Analytical methods for determination of sugars and sweetness of horticultural products—A review," *Scientia Horticulturae,* 2015.

[95] E. Arendse, A. Fawole O, S. Magwaza L and L. Opara U, "Non-destructive prediction of internal and external quality attributes of fruit with thick rind: A review.," *Journal of Food Engineering,* 2018.

[96] Y. Liu, X. Sun and A. Ouyang, "Nondestructive measurement of soluble solid content of navel orange fruit by visible–NIR spectrometric technique with PLSR and PCA-BPNN," *Food Science and Technology,* 2010.

[97] A. Kader A, "Postharvest Technology of Horticultural Crops," *Technology & Engineering,* 2002.

[98] Y. Liu, X. Sun, J. Zhou, H. Zhang and C. Yang, "Linear and nonlinear multivariate regressions for determination sugar content of intact Gannan navel orange by Vis–NIR diffuse reflectance spectroscopy," *Mathematical and Computer Modelling,* 2010.

[99] P. J. Fellars, "The relationship between the ratio of degrees Brix to percent acid and sensory flavor in grapefruit juice," *Agris,* 1991.

[100] C. Liu, X. Yang S and L. Deng, "Determination of internal qualities of Newhall navel oranges based on NIR spectroscopy using machine learning," *Journal of Food Engineering,* 2015.

[101] W. C. Cayuela, "Intact orange quality prediction with two portable NIR spectrometers," *Postharvest Biology and Technology,* 2010.

[102] A. Cayuela J, "Vis/NIR soluble solids prediction in intact oranges (Citrus sinensis L) cv. Valencia Late by reflectance," *Postharvest Biology and Technology,* 2008.

[103] S. Jamshidi B Minaei, E. Mohajerani and H. Ghassemian, "Reflectance Vis/NIR spectroscopy for nondestructive taste characterization of Valencia oranges," *Computers and Electronics in Agriculture,* 2012.

[104] S. Jie, G. Li, X. Yang, X. Liu and L. Xie, "Rapid Analysis of Soluble Solid Content in Navel Orange Based on Visible-Near Infrared Spectroscopy Combined with a Swarm Intelligence Optimization Method," *Spectrochimica acta,* 2020.

[105] Z. Hailiang, B. Zhan, F. Pan and W. Luo, "Determination of soluble solids content in oranges using visible and near infrared full transmittance hyperspectral imaging with comparative analysis of models," *Postharvest Biology and Technology,* 2020.

[106] M. D. Roland Hörnfeldt and L. Woxblom, “False heartwood in beech Fagus sylvatica, birch Betula pendula, B. papyrifera and ash Fraxinus excelsior - an overview,” *Ecological Bulletins,* 2010.

[107] I. C. Vladimír Racko, “Which of the factors do significantly affect beech false heartwood formation,” *Conference: Hardwood Science and Technology,* 2010.

[108] Y. H. Yuzhen Lu and R. Lu, “Innovative Hyperspectral Imaging-Based Techniques for Quality Evaluation of Fruits and Vegetables: A Review,” *Applied Science,* 2017.

[109] M. D. Wilson, R. Cernik, H. Chen, C. Hansson, K. Iniewski, L. L. Jones, P. Seller and M. C. Veale, “Small pixel CZT detector for hard X-ray spectroscopy,” *Nuclear Instruments and Methods in Physics Research,* 2011.

[110] D.-W. Sun, Hyperspectral Imaging for Food Quality Analysis and Control, D. Sun, Ed., Academic Press / Elsevier, 2010.

[111] J. Sugiyama and M. Tsuta, “Visualization of Sugar Distribution of Melons by Hyperspectral Technique,” *Hyperspectral Imaging for Food Quality Analysis and Control,* 2010.

[112] W. S. Ståhle L, “Partial least squares analysis with cross-validation for the two-class problem: A monte carlo study.,” *Journal of Chemometrics,* 1987.

[113] E. Sjostrom, Ed., Wood Chemistry, 2nd ed, Academic Press, 2013.

[114] L. R. Schimleck, F. Antony, C. Mora and J. Dahlen, “Comparison of whole-tree wood property maps based on near-infrared spectroscopic calibrations utilizing data at different spatial resolutions,” *Holzforschung,* 2019.

[115] A. Sandak, J. Sandak, A. Zitek and B. Hinterstoisser, “Industry 4.0 in forestry – NIR spectroscopy and hyperspectral imaging input,” *Conference: VIII Simposio Italiano di Spettroscopia NIR (NIRITALIA 2018),* 2018.

[116] A. Ruano, A. Zitek, B. Hinterstoisser and E. Hermoso, “NIR hyperspectral imaging (NIR-HI) and µXRD for determination of the transition between juvenile and mature wood of Pinus sylvestris L.,” *Holzforschung,* 2019.

[117] R. M. Rowell, Ed., Handbook of Wood Chemistry and Wood Composites, CRC Press, 2012.

[118] R. L. Schuler and C. A. Plese, “Preliminary observations on the ability of hyperspectral imaging to provide detection and visualization of bloodstain patterns on black fabrics,” *J. Forensic Sci.,* 2012.

[119] C. R. Mora, L. R. Schimleck, S.-C. Yoon and C. N. & Thai, “Determination of Basic Density and Moisture Content of Loblolly Pine Wood Disks Using a near Infrared Hyperspectral Imaging System,” *Journal of Near Infrared Spectroscopy,* 2017.

[120] P. Mishra, M. S. M. Asaari, A. Herrero-Langreo, S. Lohumi, B. Diezma and P. Scheunders, “Close range hyperspectral imaging of plants: A review,” *Biosystems Engineering,* 2017.

[121] K. Maraphum, K. Saengprachatanarug, K. Aparatana, Y. Izumikawa and E. Taira, “Spatial mapping of Brix and moisture content in sugarcane stalk using hyperspectral imaging,” *Journal of Near Infrared Spectroscopy,* 2020.

[122] T. Ma, T. Inagaki and S. Tsuchikawa, “Non-destructive evaluation of wood stiffness and fiber coarseness, derived from SilviScan data, via near infrared hyperspectral imaging,” *Journal of Near Infrared Spectroscopy,* 2018.

[123] W. R. M.L. Barker, “Partial Least Squares For Discrimination, Journal of Chemometrics,” *Journal of Chemometrics,* 2003.

[124] E. U. M. Boschetti and G. Zane, “Method and apparatus for non-destructive inspection of fruits having an axis of rotational symmetry,” *European Patent EP3521812B1,* 2020.

[125] G. Lua and B. Fei, “Medical hyperspectral imaging: a review”.

[126] H. Liang, “Advances in multispectral and hyperspectral imaging for archaeology and art conservation,” *Applied Physics,* 2011.

[127] T. Inagaki, K. Mitsuib and T. Satoru, “Visualisation of degree of acetylation in beechwood by near infrared hyperspectral imaging,” *Journal of Near Infrared Spectroscopy,* 2015.

[128] A. Haddadi, B. Leblon, J. Burger, Z. Pirouz, K. Groves and J. Nader, “Using near-infrared hyperspectral images on subalpine fir board. Part 2: Density and basic specific gravity estimation,” *Wood Material Science and Engineering,* 2015.

[129] A. Gowen, C. P. O'Donnell, P. J. Cullen, G. Downey and J. M. Frías, “Hyperspectral imaging—an emerging process analytical tool for food quality and safety control,” *Trends in Food Science & Technology,* 2007.

[130] J. K. et al., “Using VIS/NIR and IR spectral cameras for detecting and separating crime scene details,” *Proc. SPIE,* 2012.

[131] G. J. E. et al., “Hyperspectral imaging for non-contact analysis of forensic traces,” *Forensic Sci.,* 2012.

[132] G. ElMasry, D. W. Sun and P. Allen, “Near-infrared hyperspectral imaging for predicting colour, pH and tenderness of fresh beef,” *J. Food Eng,* 2012.

[133] M. B. E. Ursella, “Tunnel CT Scanner,” *European patent request EP3690429A1,* 2020.

[134] Dimitris Manolakis and G. A. Shaw, “Hyperspectral Image Processing for Automatic Target Detection Applications,” *Lincoln Laboratory Journal,* 2003.

[135] G. W. Dietrich Fengel, Ed., Wood: Chemistry, Ultrastructure, Reactions, De Gruyter, 1983.

[136] W. R. O. D. B. Malkoff, “Hyperspectral imaging applied to forensic medicine,” *Proc. SPIE,* 2000.

[137] S. G. Cherelli, M. M. P. Sartori, A. G. Próspero and A. W. Ballarin, “Heartwood and sapwood in eucalyptus trees: non-conventional approach to wood quality,” *Anais da Academia Brasileira de Ciências,* 2018.

[138] I. K. C. Fischer, “Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications,” *Studies in Conservation,* 2006.

[139] M. Boschetti, S. Faccini and M. Caffini, “Imaging iperspettrale su tavole di legno,” *Italian requested Patent,* 2021.

[140] J. M. J. X. J. P. Armando Fernandes and P. Melo-Pinto, “Measurement of intra-ring wood density by means of imaging VIS/NIR spectroscopy (hyperspectral imaging),” *Holzforschung,* 2013.